RSA Based Encrypted Data Embedding Using APPM

Fathima Nizar, Fathima Latheef, Alfina Jamal

Department of Information Technology
Amal Jyothi College of Engineering
M G University
Kottayam, India

Abstract— This paper introduces a new encrypted data hiding method based on Adaptive Pixel Pair Matching and RSA encryption algorithm. The basic idea of Pixel Pair Matching is to use the values of pixel pair as a reference coordinate, and search a coordinate in the neighbourhood set of this pixel pair according to a given message digit. The pixel pair is then replaced by the searched coordinate to conceal the digit. The text data to be hidden undergoes a pre-processing step which encrypts the data using RSA encryption before it is embedded into the cover image.

Keywords— Adaptive Pixel Pair Matching (APPM), Diamond Encoding (DE), Pixel Pair Matching (PPM), Rivest-Shamir-Adleman (RSA).

I. INTRODUCTION

Due to the wide development of multimedia and network, communication security over the Internet is becoming more important. The two main fields of research which is used to enhance the communication security are cryptography and information hiding both are applied to protect secret messages. The various cryptography methods are DES and RSA, are exclusively used to encrypt, which is the process of converting plaintext into cipher-text [1]. After data encryption, the secret data appears to be meaningless bits. Encryption avoids unauthorized user to decrypt or destroy it. Data hiding is a process to hide secret data such as message, image, information, etc. into another digital media such as text, image, audio or video streams. The concealing media is known as cover media. If the cover media is a digital image, it is called a cover image, and the altered cover image containing the secret information is called a stego-image. The embedding capacity and invisibility are the two major factors that we use in data hiding schemes [1]. Many approaches of information hiding have been proposed for diverse applications, such as patent protection, top secret transmission, tampering exposure, and figure authentication [2].

The most common encryption technique is Rivest-Shamir-Adleman (RSA) encryption algorithm [3].The RSA algorithm has both private key and public key. The RSA is a public key cryptographic algorithm that is used to ensure data communication security. The algorithm is based on two main cryptographic processes. In the first cryptographic process it uses a public key and in the second process, it uses a private key. When the cryptographic process uses a public key, it converts an input data called the plaintext into an unrecognizable encrypted output called cipher text and the process is known as encryption process. It is impossible to recover the original plaintext without the encryption password in a reasonable amount of time.

In the second process, the RSA then converts the unrecognizable data back to its original form, which is known as decryption process. Now a days, it is used in web browsers, email programs, mobile phones, virtual private networks, secure shells etc.

Diamond encoding [4] is an efficient data hiding method. The basic concept of DE is based on Pixel Pair Matching. In diamond encoding, we first partition the cover image into different blocks of consecutive pixels. The blocks of consecutive pixels must be non-overlapping. Then we need to transform the secret data into a sequence of digits. The diamond encoding method produces a Diamond Characteristic Value (DCV) of the pixel-pair block, and after the embedding procedure the value of DCV get varied. Some distortions occur when the DCV value changes and this can be minimized using diamond encoding. The less the distortion, the more image quality. DE is used to conceal the secret digit in N-ary notational system into pixel pair where $N = 2k^2+2k + 1$ when $k \geq 1$, where k is embedding parameter. The difference between the cover-block of the cover image and the stego-block of the stego image is not greater than k, and the embedding capacity of a block equals $\log_2 (2k^2+2k+ 1)$. The experimental results have proved that this hiding method is capable of hiding more data when compared with existing methods.

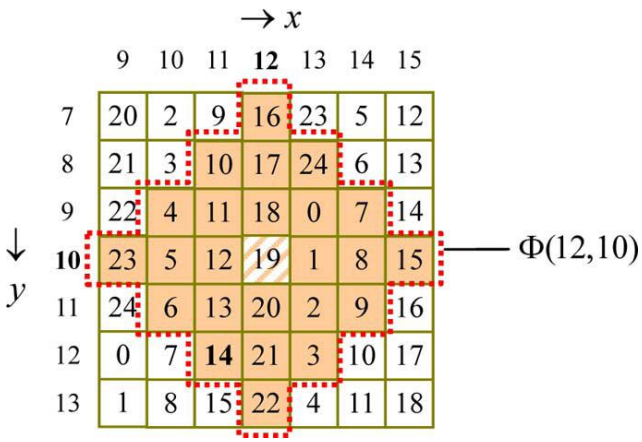

Fig.1 Neighbourhood set ϕ (12, 10) for k=3

II. RELATED WORKS

The diamond encoding scheme can conceal $(2k^2 + 2k + 1)$-ary digit into a cover pixel pair where k is the embedding parameter [5]. Steganography is the art or practice of concealing a message, image, or file within another message, image, or file. Data hiding, a form of steganography, embeds data into digital media for the purpose of identification, annotation, and copyright. In this section, Diamond Encoding, Steganography and Data Hiding Methods will be briefly reviewed.

A. Diamond Encoding

In 2009, Chao *et al.* proposed a DE method [4] based on PPM. This method conceals a secret digit in a B-ary notational system into two pixels, where $B=2k^2+2k+1$, $k \geq 1$. The payload of DE is $(1/2) \log_2 (2k^2+2k+1)$ bpp. The DE method is briefly described as follows.

Let the size of m bits cover image be M x M, message digits be S_B, where the subscript B represents S_B is in a B-ary notational system. First, the smallest integer k is determined to satisfy the equation 1.

$$\left\lfloor \frac{M \times M}{2} \right\rfloor \geq |S_B| \quad (1)$$

where $|S_B|$ denotes the number of message digits in a B–ary notational system. To conceal a message digit s_B into pixel pair(x, y), the neighbourhood set $\phi(x, y)$ is determined by the equation 2.

$$\phi(x, y) = \{(a, b) \,|\, |a-x|+|b-y| \leq k\} \quad (2)$$

where $\phi(x, y)$ represents the set of the coordinates (a, b)'s whose absolute distance to the coordinate (x, y) is smaller or equal to k . A diamond function f is then employed to calculate the DCV of (x, y), where $f(x, y) = ((2k+1)\, x + y) \bmod B$. After that, the coordinates belong to the set $\phi(x, y)$ are searched and DE finds a coordinate (x', y') satisfying $f(x', y') = s_B$, and then (x, y) is replaced by (x', y'). Repeat these procedures until all the message digits are embedded. In the extraction phase, pixels are scanned using the same order as in the embedding phased. The DCV value of a pixel pair (x', y') is then extracted as a message digit.

Here is a simple example. Let $k = 3$ and $(x, y) = (12, 10)$, then $B= 2 \times 3^2 + 2 \times 3 + 1= 25$. The neighbourhood set ϕ (12, 10) and its corresponding DCV values are shown in Fig. 1 [4]. If a digit in a 25-ary notational system 14_{25} needs to be embedded, then in the region defined by ϕ (12, 10), we find the DCV value of $(x', y') = (11, 12) =14$. Therefore, we simply replace (12, 10) by (11, 12) and the digit 14_{25} is embedded. To extract the embedded digits, we calculate

$f(x', y') = f(11,12)=(7 \times 11+12) \bmod 25 =14$; the calculation result 14 is then the embedded digit.

B. Steganography

Steganography is a secret Communication to hide the secret Data. It is an invisible communication that hides data like text, image, and audio, video etc. The secret message is inserted into the image files. The image files can use stego-key to hide the data and the resultant image is called as stego-image. Steganography [6] plays an important role in defence.

1) Kinds of Steganography

Steganography [7] can be used for main categories of file formats like text, images, audio/video and protocol (Fig.2).

a) Text Steganography

Text as a cover medium is the oldest techniques used in early days. Secret message is detected by taking the first letter of every word present in the file. The process is continued by considering the various positions of the letters. The amount of message that is hidden in this type is very less and easily recoverable by the frequency of letters. The following are some of the methods used frequently by steganographers.

- Text hiding in Mark-up Languages (HTML)
- Semantic methods
- Character encoding
- Line shifting method
- Word shifting
- Text steganography in specific characters in words
- Open spaces

b) Image Steganography

Pictures are attractive to human rather than text. Internet pages are very popular for its pleasant pictures. With this concept images are used to hide the secret information. Now a day's people are using various calculations. Randomly selecting the pixels in the image and replacing the ASCII values of the text are highly unbreakable algorithm cryptography and

steganography shakes its hands together to make the image steganography robust.

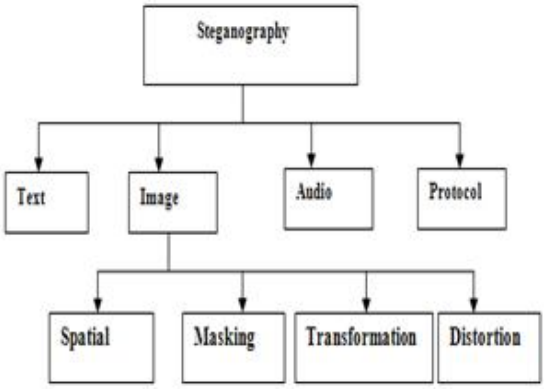

Fig.2 Classification of steganographic Techniques

c) Audio Steganography

Digital wave files are used to hide message. People are hearing the music in their day today life. Free music downloads from internet through PDA, mobile phones and PC makes the music files popular. Steganographers pay their attention in these audio files their secret message.

To embed data secretly onto digital audio there are few techniques introduced;

- Parity coding
- Spread spectrum
- LSB coding
- Phase coding

d) Protocol Steganography

A set of rules used to govern the communication is known as protocol. TCP, IP, UDP are the some of the protocols used for communication purpose. Steganographers use this protocol for hiding their secret data. Some unused parts of the protocol like packet header are efficiently used for message hiding [8].

C. Data Hiding Methods

Several methods are available in literature for hiding data, they are Least Significant Bit (LSB), Pixel Value Differencing (PVD), Gray Level Modification (GLM), Parity Checker Method (PCM), Diamond Encoding Method (DEM), Optimal Pixel Adjustment Process (OPAP), Exploiting Modification Direction (EMD), and Adaptive Pixel Pair Matching (APPM) (fig.3).

1) Least Significant Bit Method (LSB)

This method replaces least bit significant bits with the message to be encoded. It is a frequent technique used so far when dealing with images [9]. It is a simple, susceptible to lossy compression and image manipulation method.

2) Pixel Value Differencing Method (PVD)

This method can successfully use both embedding and outstanding gradual of the stego-object [10]. The differencing method may be divided in to original image into acceptable blocks having two joining pixels and changes the pixel difference in every one block for data can be embedded.

3) Gray Level Modification Method (GLM)

This method is used to map data by changing grey level of the image pixel. GLM steganography is a technique of diagrammatic representation of the data by changing the gray level values of the image pixel. Gray level steganography uses the sum of 0's and 1's value to map within an image. This is directly mapped between the binary values and the suitable pixels in an image. The set of image pixels are based on a mathematical function [11]. Gray level values of that pixel are determined based on the similarity with the bits stream that is to be mapped into the image.

4) Parity Checker Method (PCM)

This method uses 0 and 1 parity mechanism. In a simplified manner this method even value can be inserted at a pixel position to identify pixel has 1(odd) parity bits. It can be identical odd value insert at a pixel; if the pixel should be 0 (even) parity [12]. If the close similarity parity do not exist at a pixel position for odd or even, then the pixel location can be added and subtracted such that the change in the image quality will not be Visible (to the human visual system).

5) Exploiting Modification Direction Method (EMD)

This method uses only one pixel pair which is modified to one gray scale unit message digits in a five-ary system for embedding [4]. This is not efficient for high payload applications.

6) Diamond Encoding Method (DE)

This method is contest position in pixel pair. It is improved by much more payload capacity, meanwhile able to accept stego image with excellence quality [4]. Still two problems need to be overcome such that the payload should be in a most suitable notational system. The arbiter value is not selected in a system.

7) Optimal Pixel Adjustment Process Method (OPAP)

This easily understood and effort pixel adjustment process makes minor distortion by the least significant bit replacement [4]. However distortion will be there by LSB replacement of adjusting pixel values.

8) Adaptive Pixel Pair Matching Method (APPM)

This method uses for pixel pair which competes each other. It is used to refer different elements and find element in the area surrounding a particular pixel pair position in corresponding

message digit. This method has low distortion for different payload.

III. RSA ENCRYPTION METHOD

The RSA algorithm [13] was publicly described in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman at MIT; the letters RSA are the initials of their surnames.

RSA is an algorithm for public-key cryptography that is based on the presumed difficulty of factoring large integers, the factoring problem. RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman, who first publicly described the algorithm in 1977. Clifford Cocks, an English mathematician, had developed an equivalent system in 1973, but it wasn't declassified until 1997.

RSA users and publishes the product of two large prime numbers with an additional value, as their public key. The prime factor must be secretly kept by user. To encrypt a message the user can use the public key but with currently published methods. The message can be probably decoded only with the knowledge of prime factor. Whether breaking RSA encryption is as hard as factoring is an open question known as the RSA problem.

The system includes a communications channel coupled to at least one terminal having an encoding device and to at least one terminal having a decoding device. A message-to-be-transferred is enciphered to cipher text at the encoding terminal by encoding the message as a number M in a predetermined set. That number is then raised to a first predetermined power (associated with the intended receiver) and finally computed. The remainder or residue, C, is... computed when the exponentiated number is divided by the product of two predetermined prime numbers (associated with the intended receiver).

A. Key Generation

Public key and Private Key are the two keys involved in RSA algorithm**.** For encrypting the message use public key, that can be known for everyone. The encrypted message is decrypted using private key. The keys for the RSA algorithm are generated the following way:

1. Choose two distinct prime numbers p and q.

   ◦ For security purposes, the integers p and q should be chosen at random, and should be of similar bit-length. Prime integers can be efficiently found using a primality test.

2. Compute $n = pq$.

   ◦ n is used as the modulus for both the public and private keys. Its length, usually expressed in bits, is the key length.

3. Compute $\varphi(n) = \varphi(p)\,\varphi(q) = (p-1)(q-1)$, where φ is Euler's totient function.

4. Choose an integer e such that $1 < e < \varphi(n)$ and $\gcd(e, \varphi(n)) = 1$; i.e. e and $\varphi(n)$ are coprime.

   ◦ e is released as the public key exponent.

   ◦e having a short bit-length and small Hamming weight results in more efficient encryption – most commonly $2^{16} + 1 = 65{,}537$. However, much smaller values of e (such as 3) have been shown to be less secure in some settings.

5. Determine d as $d^{-1} \equiv e \pmod{\varphi(n)}$,
i.e., d is the multiplicative inverse of e (modulo $\varphi(n)$).

   ◦ This is more clearly stated as solve for d given $de \equiv 1 \pmod{\varphi(n)}$.

   ◦This is often computed using the extended Euclidean algorithm.

   ◦ d is kept as the private key exponent.

By construction, $d \cdot e \equiv 1 \pmod{\varphi(n)}$. The **public key** consists of the modulus n and the public (or encryption) exponent e. The **private key** consists of the modulus n and the private (or decryption) exponent d, which must be kept secret. p, q, and $\varphi(n)$ must also be kept secret because they can be used to calculate d.

B. Encryption

Alice transmits her public key (n, e) to Bob and keeps the private key secret. Bob then wishes to send message M to Alice. He first turns M into an integer m, such that $0 \leq m < n$ by using an agreed-upon reversible protocol known as a padding scheme. He then computes the cipher text c corresponding to

$$c \equiv m^e \pmod{n}$$

This can be done quickly using the method of exponentiation by squaring. Bob then transmits c to Alice.

C. Decryption

Alice can recover m from c by using her private key exponent d via computing

$$m \equiv c^d \pmod{n}$$

Given m, she can recover the original message M by reversing the padding scheme.

IV. ADAPTIVE PIXEL PAIR MATCHING

The basic idea of the PPM-based data-hiding method [14] is to use pixel pair (x, y) as the coordinate, and searching a coordinate $(x`, y`)$ within a predefined neighborhood set $\phi(x, y)$ such that $f(x`, y`) = s_B$, where f is the extraction function and s_B is the message digit in a B-ary notational system to be concealed. Data embedding is done by replacing (x, y) with $(x`, y`)$.

For a PPM-based method, suppose a digit s_B is to be concealed. The range of s_B is between 0 and B-1, and a coordinate $(x`, y`)$ element of $\phi(x, y)$ has to be found such that $f(x`, y`) = s_B$.

Therefore, the range of $f(x, y)$ must be integers between 0 and B-1, and each integer must occur at least once. In addition, to reduce the distortion, the number of coordinates in $\phi(x, y)$ should be as small as possible.

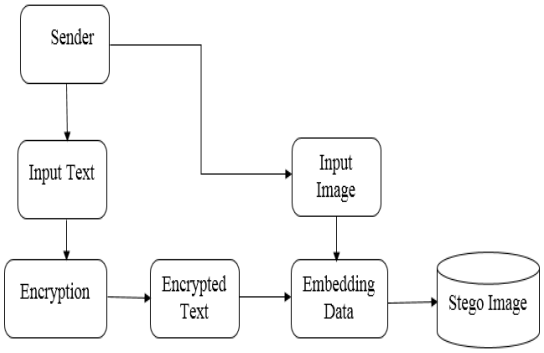

Fig.3 Flow Chart of embedding part

A. Embedding Procedure

Suppose the cover image is of size M x M, is the message bits to be concealed and the size of S is $|S|$. First we calculate the minimum such that all the message bits can be embedded. Then, message digits are sequentially concealed into pairs of pixels. The detailed procedure is listed as follows.

Input: Cover image I of size M x M, secret bit stream S, and key Kr.

Output: Stego Image $I`$, c_B, $\phi_B(x, y)$, and Kr.

1. Find the minimum B satisfying $[M \times M/2] \geq |S_B|$, and convert S to a B-ary notational system S_B.
2. Solve discrete optimization problem to find c_B and $\phi_B(x, y)$.
3. In the region defined by $\phi_B(0, 0)$, record the coordinate (xi, yi) such that $f(xi, yi) = i$, $0 \leq i \leq B$-1.
4. Construct a nonrepeat random embedding sequence Q using a key Kr.
5. To embed message bits s_B, two pixels (x, y) in the cover image are selected according the embedding sequence Q, and calculate the modulus distance, $d = (s_B - f(x, y)) \bmod B$ between s_B and $f(x, y)$, then replace (x, y) with $(x+x_d, y+y_d)$.
6. Repeat step 5 until all message bits are embedded.

B. Extraction Procedure

To extract the embedded message digits, pixel pairs are scanned in the same order as in the embedding procedure. The embedded message digits are the values of extraction function of the scanned pixel pairs.

Input: Stego image I', c_B, $\phi(x, y)$, and Kr

Output: Secret bit stream S.

1. Construct the embedding sequence Q using Kr.
2. Select two pixels $(x`, y`)$ according to the embedding sequence Q.
3. Calculate $f(x`, y`)$, the result is the embedding digit.
4. Repeat steps 2 to 3 until all digits are extracted.
5. Convert the digits to bit stream.
6. Finally, the message bits can be obtained by converting the extracted message digits into a binary bit stream.

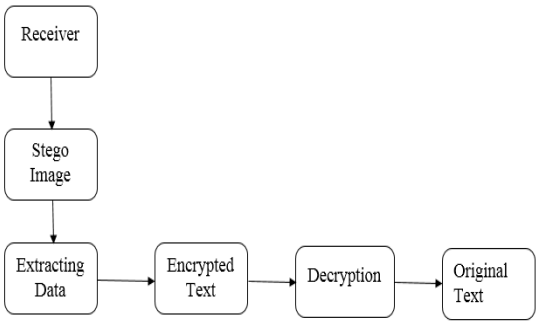

Fig.4 Flow Chart of extracting part

V. PROPOSED SYSTEM

In the proposed system, mainly we have two parts, the embedding part and the extraction part. First of all, we encrypt the text data using public key of RSA algorithm, before it is embedded into the cover image. After embedding the encrypted data into the cover image what we obtain is the stego image, which is the output of the embedding part.
In the extraction part, the output of the embedding part that is the stego image, becomes the input, which is given for performing extraction process. After extracting the encrypted data from the stego image, it is decrypted using the private key of the RSA algorithm to obtain the original text.

A. Embedding Part

The flow chart of the embedding part of the proposed system is shown in Fig.3 and the steps are:

1. Choose the cover image to hide the data.
2. Read the data to be hidden.
3. Encrypt the data using public key of RSA algorithm.
4. Encode or embed the data into the image using diamond encoding.

B. Extraction Part

The flow chart of the extracting part of the proposed system is shown in Fig.4 and the steps are:

1. Read the stego image that contains the data in it.
2. Decode or extract the data from the stego image.
3. Decrypt the data using private key to get the original message.

VI. CONCLUSION

This paper proposed a simple and efficient encrypted data embedding method based on Pixel Pair Matching. Two pixels are scanned as an embedding unit and a specially designed neighbourhood set is employed to embed message digits with a smallest notational system. The encrypted data hiding method allows users to select digits in N-ary notational system for data embedding, and thus achieves a better image quality. Moreover, because APPM produces no artifacts in stego images and the steganalysis results are similar to those of the cover images, it offers a secure communication under adjustable embedding capacity. The main advantage of encrypted data hiding method is security against attackers and high data payload can be used.

REFERENCES

[1] Ruey-Ming Chao, Hsien-Chu Wu, Chih-Chiang Lee, and Yen-Ping Chu,"A Novel Image Data Hiding Scheme with Diamond Encoding".

[2] P.Ramesh Babu, Y.ChittiBabu, Dr.P.Harini,"A Narrative Data Embedding Method Using Adaptive Pixel Pair Matching", International Journal of Computational Engineering Research||Vol, 03||Issue, 11

[3] Implementation of the RSA algorithm and its cryptanalysis Chandra M. Kota and CherifAissi University of Louisiana at Lafayette, College of Engineering Lafayette , LA 70504, USA.

[4] Wien Hong and Tung-Shou Chen ,"A Novel Data Embedding Method Using Adaptive Pixel Pair Matching", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 7, NO. 1, FEBRUARY 2012.

[5] R.M. Chao, H. C. Wu, C. C. Lee, and Y. P. Chu, "A novel image data hiding scheme with diamond encoding," *EURASIP J. Inf. Security*, vol. 2009, 2009, DOI: 10.1155/2009/658047, Article ID 658047.

[6] J.Fridrich, Steganography in Digital Media:Priciples, Algorithms and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[7] C.Gayathri, V.Kalpana "Study on Image Steganography Techniques ",Computer Science & Engineering, School of Computing, SASTRAUNIVERSITY, Journal of Engineering and Technology (IJET).

[8] Vijay kumar sharma, Vishal shrivastava "A steganography Algorithm for hiding image in image by improved lsb substitution by minimize detection" journal of theoretical and applied information technology 15th february 2012. vol. 36 no.1.

[9] Neil F Johnson, Sushil Jajodia, "Exploring Stenography: Seeing the Unseen", IEEE Computer, Feb 1998.

[10] D.C. Wu and W.H. Tsai. "A steganographic method for images by pixel value differencing", Pattern Recognition Letters 24: 1613-1626, 2003.

[11] Vidyasagar M. Potdar, Elizabeth Chang, "Grey Level Modification Stegnography for Secret Communication", 2nd IEEE International Conference on Industrial Informatics INDIN 2004 June 24th, 26th June, Berlin, Germany, Submitted Tuesday, May 25, 2004.

[12] Yadav, Rajkumar, Rishi, Rahul, Batra, Sudhir, "A new Steganography Method for Gray Level Images using Parity Checker",International Journal of Computer Applications(0975-8887) Volume 11-No. 11, December 2010.

[13] AviKak, Public-Key Cryptography and the RSA Algorithm Lecture Notes on "Computer and Network Security",June 20, 2013 4:40pm c 2013AvinashKak, Purdue University.

[14] T.Lakshmi, P.GangaBhavani,"A Narrative Data Embedding Method Using Adaptive Pixel Pair Matching",International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 1, Issue 7, September 2013.