An Efficient Data Hiding Scheme Using Steganography and Cryptography Technique

Mangesh Kulkarni, Prasad Jagtap, Ketan Kulkarni

NMIET, Information Technology, India, Maval

Abstract- In the 21st century, with more use of internet there has been significant increase in various attacks on the network. Here we describe the technique of how can a secrete communication take place between two or more parties secretly. In this paper the technique of cryptography and Steganography are used all together for to achieve message confidentiality. In cryptography, message gets encrypted and in Steganography, it gets embedded within the selected image. For that, we first substitute the original message by using the fourteen square substitution algorithm. After the substitution of text, we then encrypt this text message using RSA algorithm. The encrypted message is then hidden within image by Least Significant Bit (LSB) method. This image works as a carrier file and is sent to the receiver. At the receivers end, same operations are performed to decrypt the original message in reverse order. It is found that here we are using the double ciphering techniques which makes the system very robust and secures it from known hacking attacks. It makes very difficult for the intruders to hack the image and then decrypt the message in a feasible amount of time thus securing it from many known network attacks.

Index Terms- Steganography, Cryptography, Fourteen square algorithm, encryption, decryption.

I. INTRODUCTION

Network security is collection of tools designed to protect the data from the hacker and intruders. Since the start of the internet one of the most important factors of Information and Technology is the security of its biggest asset, the information. Cryptography was created as a technique for security purpose and maintaining the confidentiality in communication. Uptill now there has been a constant evolution in invention of different types of techniques to encrypt and decrypt of data and keeping message secrete. Unfortunately, cryptography alone does not serve to keep the contents of message secrete. Many times the cryptanalysis does lot of work for hackers and there is failure to maintain the confidentiality of message. It is also pretty tough to keep the existence of the message secrete. To improve current system, we have used Steganography with cryptography for increasing robustness. Steganography literally means covered writing. It is the practice of hiding a message in such a manner that its very existence is concealed. This is achieved by hiding information inside information. The word Steganography is derived from the Greece words Stegos means cover and Grafia means writing. Steganography can be achieved by using image, audio, video carriers. The most flexible and efficient way to use it is with help of images, hence in our software, we are using it in form of image. Some of the most common methods of image Steganography proposed by different researchers are discussed in following section.

II. EXISTING SYSTEM

A Least Significant Bit (called LSB) method is used for hiding the information. The 8th bit of carrier files every byte is substituted by 1 bit of secret information. This approach may give the disadvantage that it causes noticeable distortion when number of embedding in each pixel exceeds three. To improve the drawback of distortion many adaptive methods for steganography have been proposed. Another approach in LSB method for hiding information within encrypted image data is checking the entropy and correlation value of stego image and original image are same after embedding. To implement this number of horizontal and vertical blocks will be generated at the sender side and then it is mixed with encrypted image before sending it to receiver. The receiver will need transformation table to reconstruct the information received from sender. But here only secret information is sent instead of whole transformation table. Each bytes LSB contains the hidden data within the encrypted image data randomly based on secret key before sending it. The values of entropy and correlation must be same before and after insertion process. This gives the advantage of being undetected and security enhancement of encrypted image.

By using Discrete Wave length Transformation a new steganography technique is used which embeds our secret message within the frequency domain. It consists of two operations one is horizontal and the other is vertical. In horizontal first the pixels are scanned from left to right and then perform addition and subtraction on the neighboring pixels and the sum is stored on top and the difference at the bottom. The advantage of this is the security is maintained and the message is useless without key matrix and decoding rules. Also as far as possible the technique can be used to reduce extra data in stego image data compress the size of key matrix.

To inherit noise places a new method of digital steganogra-phy Spread Spectrum Image steganography (SSIS). It includes error control coding, image restoration and those similar

to spread spectrum communication are merged within this system. The hidden message is embedded within noise and then added to digital cover image. The color imagery and audio signals method may be the extensions of this technique.

A new technique of hamming code whose goal is to improve the embedding rate which give two different schemes. The first one is derived from blind watermarking scheme. The second scheme includes relative entropy between cove data and stego data tends to be zero. For digital watermarking the main care to be taken off the subsequent processing should not destroy the embedded information. This property makes digital watermarking technology attractive for steganography information embedding is followed by lossy consumption. It gives advantage such that the performance of scheme is compared with respect to security embedding distortion and rate of embedding.

There are lots of techniques present for substitution algo-rithms. The history of using square substitution method all started back from as basic as two squares.

In next method a text steganography using two square reverse ciphers is proposed. In this, there are two steps for encryption. In the first step, get the first step cipher by using table1 and table2. After getting the first step cipher we divide that into 2-2 characters. After that swap the 2-2 characters positions and get the final cipher. The cover image is divided into bytes. In each byte do not embed; but in few selected bytes based on the bit pattern of the cipher text. In the selected bytes the 7th bit position that is LSB minus one are to be embedded. The 8th bit means the LSB bit position. Advantage here is it provides two level security- cryptography and steganography. As compared to LSB method this algorithm is much better in terms of intrusion prevention. But disadvantage is Table1 and Table 2 is divided in to alphabets where q is missing also digits are not included. In future this approach an be extendable to send secret image in cover image.

This approach can be extendable to audio and video carrier files. A method for steganography using six square substitution ciphers which includes only alphabets.However the research further continued and six squares were designed. This made the algorithm secure but however it still did not had any provision for substitutions of special symbols and numbers. Just the squares were increased for proper ordered matrix and greater substitution flexibility.

Next in this research field to evolve was the recent 12 square matrix. All most all cons were tried to eliminate in this method. The 12 square was able to encrypt not only all the alphabets but also the special symbols and the numbers. Precisely the 12 square uses six of its squares arranged in 5 by 5 matrices and other six squares in 6 by 7 matrices. First 6 squares consist of alphabets and the other 6 squares consist of the numbers and symbols.

Though the 12 square methods was a secure one it still had some cons in it. Namely important ones were like the letter q was missing in the table. It was done so for the sake of proper matrix setting. Also some special characters and digits were missing in the table. Because of this whenever the letter q was detected it remained unchanged in the enci-phering process. This existing system is only able to perform one time encryption which make security of single level. In 12 square method the disadvantage is that the message was only substituted the 12 squares table and it could get easily hacked by the attackers. This system suffered lots of drawbacks so in order to overcome all the cons of previous approaches we have proposed a new optimal system that is described in next section.

This was the latest existing system working to current date but it has the following drawbacks as chalked below:

Consists of only 12 squares

Letter q is missing in the table

Some special characters are missing

Uppercase and lowercase are not differentiated No different mechanism for key generation

No robust mechanism present as there is only one level of encryption

Only one level of message encryption is done

III. PROPOSED SYSTEM

The Proposed system provides high level of security by dual ciphering method. In this system we are substituting the original message by using fourteen square substitution algorithms and after that we are applying the RSA encryption algorithm on substituted text and that encrypted cipher text is embedded in image. Hence the message is dual enciphered thus taking the security measures of information to next level.

A. FOURTEEN SQUARE SUBSTITUTION CIPHER

In the earlier paper the twelve square cipher substitution is represented which includes only lower case alphabets, numbers and some special characters. This is extended in to fourteen square substitution ciphers to include numerals, upper and lower alphabets and all special characters including space bar which are present on keyboard. At first, the fourteen-square substitutes the alphabets, digits and special characters so it is less susceptible to frequency analysis attacks. It uses eight 9 by 6 matrices each arranged in a square, as shown in figure-1. Each of the 9 by 6 matrices contains the letters of the alphabets (upper case and lower case) and another six 6 by 7 matrices arranged in a squares for digits and special characters, as shown in figure-2. All the special characters from your desktop/laptop keyboard are included in this figure. So the following describes how the figure-1 is prepared: in square-1, we have taken fifty two alphabets and two special characters, out of which twenty six are capital letters and twenty six, are small letters. In each row we have arranged nine alphabets and each column contains six alphabets. Square- 2 is made from square- 1 by taking the first row of the square-1 to sixth row place and other rows one position up. Similarly square-3 is created from square 2 by taking the first row of square 2 to sixth row place and other rows are position up.The same procedure is for square- 4 which created from square-3 by taking first row of square-3 to sixth row of square-4 and other rows are position up. In square-5, we have converted rows into column and interchanged first and last alphabets. The same step follows in square-6 to square-8 by taking first row of previous square to sixth row and other rows one position up.The same method we have followed in figure- 2. It is shown in figure-1 and figure-2.

Square 1	Square 2	Square 3	Square 4
s t u v w x y z A	B C D E F G H I J	K L M N O P Q R S	T U V W X Y Z @?
B C D E F G H I J	K L M N O P Q R S	T U V W X Y Z @?	a b c d e f g h i
K L M N O P Q R S	T U V W X Y Z @?	a b c d e f g h i	j k l m n o p q r
T U V W X Y Z @?	a b c d e f g h i	j k l m n o p q r	s t u v w x y z A
a b c d e f g h i	j k l m n o p q r	s t u v w x y z A	B C D E F G H I J
j k l m n o p q r	s t u v w x y z A	B C D E F G H I J	K L M N O P Q R S
Square 5	**Square 6**	**Square 7**	**Square 8**
r y E K Q W a g m	t z F L R X b h n	u A G M S Y c i o	v B H N T Z d j p
t z F L R X b h n	u A G M S Y c i o	v B H N T Z d j p	w C I O U @ e k q
u A G M S Y c i o	v B H N T Z d j p	w C I O U @ e k q	x D J P V ? f l s
v B H N T Z d j p	w C I O U @ e k q	x D J P V ? f l s	r y E K Q W a g m
w C I O U @ e k q	x D J P V ? f l s	r y E K Q W a g m	t z F L R X b h n
x D J P V ? f l s	r y E K Q W a g m	t z F L R X b h n	u A G M S Y c i o

Fig. 1. Plain Text and Cipher Text Table(Alphabets)

Square 9	Square 10	Square 11
0 1 2 3 4 5 6	7 8 9 ` ~ !	# $ % ^ & * (
7 8 9 ` ~ !	# $ % ^ & * (	) _ - + = { [
# $ % ^ & * (	) _ - + = { [	}] ; : " ' \
) _ - + = { [	}] ; : " ' \	| <, >. € /
}] ; : " ' \	| <, >. € /	0 1 2 3 4 5 6
| <, >. € /	0 1 2 3 4 5 6	7 8 9 ` ~ !
Square 12	**Square 13**	**Square 14**
0 6 ! & + ; <	1 7 * = : ,	1 7 * = : ,
1 7 * = : ,	2 8 # ({ " >	2 8 # ({ " >
2 8 # ({ " >	0 6 ! & + ; <	3 9 $) [' .
3 9 $) [' .	3 9 $) [' .	4 ` % _ } \ €
4 ` % _ } \ €	4 ` % _ } \ €	5 ~ ^ -] | /
5 ~ ^ -] | /	5 ~ ^ -] | /	0 6 ! & + ; <

Fig. 2. Plain Text and Cipher Text Table(Digits and Special Characters)

The plain text is read from left to right. If the character is alphabet it will refer to the figure-1, otherwise if it is a number or a special character it will refer to figure-2. While scanning the plain text the first alphabets plain text is in square-1 and its cipher is in same row and column location in square-5. The second alphabet, its plain text is in square- 2 and its cipher text is in same row and column location of square6. The third alphabet, its plain text is in square-3 and cipher text is in same row and column location of square-7. The fourth alphabet, its plain text is in square 4 and its cipher text is in same row and column location of square-8 similarly fifth alphabet corresponds to square-1 and square-5, 6th alphabet corresponds to square-2 and square-6, 7th alphabet corresponds to square-3 and square-7 and so on. Secret message is combination of alphabets, digits and special characters. While scanning for the secrete message, for the special characters and digits it will refer to the figure-2. The special character (including digit), its plain text is in square-9 and cipher text is in same row and column location of square-12. For another special character(including digits) it follow the same procedure as mentioned for figure-1 characters. How it will work can be understood by below little example.

If plain text is: Hello

Its cipher text would be: bUJJ?

After this successful substitution of the plain text, again this cipher text will be encrypted by the one of the most secure encryption algorithm. For this we are using the RSA algorithm form encryption and decryption process and lastly the encrypted message will be embedded inside the image. The substituted image is then re-encrypted using the RSA in following way:

B. RSA ALGORITHM

(1) Given an integer key size 1¡=k¡=4000, generate random prime large Integers p, q with k bits; let n = pq and generate random large Integer d such that p,q ¡ d ¡ n and relatively prime to (p-1)(q-1). Let e = inverse of d mod (p-1)(q-1). Save p, q, d in hexadecimal format.

(2) Given two prime large Integers set p, q = these (check their primality) and generate d, e as in (1). Save p, q, d in hex format.

(3) (3)Given p, q, d, check primality of p, q; check p, q ¡ d ¡ pq and GCD(d, (p-1) (q-1)) = 1 and if ok, generate e = inverse of d mod (p-1)(q-1). Save p, q, d.

(4) read p, q, d and compute e = inverse of d mod (p-1)(q-1).

(5) Divide input data into x blocks

(6) Add padding bits into x blocks

(7) Convert blocks into integer and calculate exponential value

(8) Calculate modulus n of above calculated value

(9) Repeat process for all blocks to generate encrypted data

(10) For decryption, Divide input data into x blocks

(11) Add padding bits into x blocks

(12) Convert blocks into integer and calculate exponential value using key d

(13) Calculate modulus n of above calculated value

(14) Repeat process for all blocks to generate decrypted data

IV. EMBEDDING PROCESS

After the substitution and conversion (into RSA), the code is then finally ready to get embed in image, also known as carrier image.For this process the carrier file is transformed into binary form. Each pixel becomes 1 byte. The cipher text of the secret

message is converted into bytes. Now calculate the number of bytes, suppose it is n. divide it by 2, say it is x. the x called as the index variable. The value x=0, corresponds to 6th and 7th bit location, x=1 corresponds to 7th and 8th bit locations of any pixel (byte) of the digital image. If present value of x=0 then hide the two bits of cipher text in 6th and 7th bit locations of the present pixel (byte), and next value of x is 1 for the next pixel. If present value of x=1 then hide the two bits of the cipher text in 7th and 8th bit location of the present pixel, and next value of x is 0 for next pixel.

Carrier File Byte	Operation	Location	Index Variable, x
Byte A	Embed(11)	6^{th} and 7^{th}	1
Byte B	Embed(00)	7^{th} and 8^{th}	0
Byte C	Embed(10)	6^{th} and 7^{th}	1
Byte D	Embed(11)	7^{th} and 8^{th}	0
Byte E	Embed(01)	6^{th} and 7^{th}	1
Byte F	Embed(11)	7^{th} and 8^{th}	0
Byte G	Embed(10)	6^{th} and 7^{th}	1
Byte H	Embed(10)	7^{th} and 8^{th}	1
Byte I	Embed(10)	6^{th} and 7^{th}	0
Byte J	Embed(10)	7^{th} and 8^{th}	1
Byte K	Embed(10)	6^{th} and 7^{th}	0
Byte L	Embed(10)	7^{th} and 8^{th}	1
	So On		

Fig. 3. BYTE Selection using Index Variable

EXAMPLE

Consider the cipher text to be sent is: 11001011 01111010 10101010.

This data is three byte. So n=3 and x=1. Suppose the different bytes of digital image are A,B,C etc. from figure-3 we can see that in byte A of carrier file we embedded the data bits 11 in 6th and 7th bit location, and the next value of x becomes 0. So we embed the next data bits 00 in 7th and 8th location and so on. In figure-3 every image there will be some bytes representing the image features which should not be get altered. In JPEG image having size one Mega Bytes, there will be a maximum of hundred bytes carrying the image characteristics. If we modify these characteristics then image will be get disturbed, so these characteristics should not be alter. For different image format it is different.

V. CONCLUSION AND FUTURE WORK

It is observed that algorithm works fine. It provides three level securities one at substitution level, second at cryptog-raphy level, and third at Steganography level. If at all the intruder suspects data it is very difficult for him to steal data. The degradation in image is not noticeable. The size of the image I not increase after embedding process. The presented system have no support for cross platforms like android, Mac and linux. We believe to extend our work in all the platforms to increase the flexibility in our system. Further, we will also try to bring our system on the smart phones. Many other techniques can be used in increasing security and further decreasing total time complexity of total system.

REFERENCES

[1] Gandharba Swain, Saroj Kumar Lenka, ”Steganography using the Twelve Square Substitution Cipher and Index Variable”,IEEE transactions on Image Processing, 2011, pp. 84-88.

[2] Mohammad Ali Bani Younes and Aman Jantan, ”A New Steganography Approach for Image Encryption Exchange by using the LSB insertion”, International Journal of Computer Science and Network Security, Vol 8, No 6,2008, pp. 247-254.

[3] Lisa M. Marvel and Charles G. Boncelet, ”Spread Spectrum Image Steganography”, IEEE Transactions on Image Processing, Vol. 8, No.8, 1999, pp.1075-1083.

[4] Joachim J. Eggers, R.Bauml and Bernd Girod, ”A Communications Approach to image steganography”, Proc. of SPIE Volume 4675, SanJose, Ca, 2002, pp. 1-12.

[5] Gandharba Swain, Saroj Kumar Lenka, ”A Technique for Secure Com-munication using Message Dependent Steganography”, Special issue of IJCCT, Vol. 2,No. 12,2010.

AUTHORS

First Author – Mangesh Kulkarni, NMIET, Information Technology, India, Maval, Email: kulkarnimangesh03@gmail.com
Second Author – Prasad Jagtap, NMIET, Information Technology, India, Maval, Email: prasadjagtap1407@gmail.com
Third Author – Ketan Kulkarni, NMIET, Information Technology, India, Maval, Email: kulkarniketan92@gmail.com