INVOLUTORY MATRIX IN VISUAL CRYPTOGRAPHY

P.Shanmugam[1] & C.Loganathan[2]
[1] Department of Mathematics, Kongu Engineering College, Perundurai – 638052, India
[2] Principal, Maharaja Arts and Science College, Coimbatore – 641407

ABSTRACT
In Hill cipher encryption, the complexity of finding the inverse of the matrix during decryption is eliminated by adopting self invertible matrices. Authentication of hall tickets for candidates during examination is a complicated issue. By evaluating simultaneously encryption and decryption with details of text, photos and signatures of candidates it is easy to prevent malpractices during examination held at various centres of the institution.

Keywords: *Hill cipher, Encryption and decryption of text and image, self invertible matrix.*

1. INTRODUCTION

Cryptography refers to the scientific system that encompasses the principles and methods of transforming data. Using this method, intelligible messages get transformed into unintelligible messages and then this process is reversed to get the message in its original form. In modern times, cryptography is considered as a branch of mathematics and computer science. It is affiliated closely with information theory, computer security, and engineering [3].

In cryptography, in the same file two types of data – one in text format and another in jpeg format – are created. While self invertible matrix of modulo 26 is applied for the text data, self invertible matrix of modulo 256 is used for the image format. The quality of encryption and decryption has been calculated using various measures.

There are two ways in which secret messages can be protected from being stolen during transmission. One method, namely encryption is the process of encoding secretive information in such a way that only the right person using a right key can decode the original information. Another way, i .e. steganography involves a technique which hides secretive information into a carrier so that it becomes less attractive. In this paper, both the techniques are used to insert the text.

This paper deals with self-invertible matrix used in Hill cipher algorithm. It helps to overcome the drawback of using a random key matrix in Hill cipher algorithm for encryption. If the matrix is not invertible it will not be able to decrypt the encrypted message. By avoiding the process of finding inverse of the matrix during decryption, the computational complexity can also be reduced. By using self invertible key matrix of higher order for encryption, high security and better image encryption are achieved. In addition, grayscale as well as colour images are encrypted.

Information hiding can be obtained into four phases are: preliminary, embedded, transmission and extraction. In the preliminary stage an encryption technique is applied while the embedded stage uses an algorithm to hide information. A security issue must be used in each step. Information hiding can be used for different applications, namely military, e-commerce, confidential communication, copyright protection, copy control, authentication digital elections etc. In these domains, hiding information is better than ciphering because in the hiding process there is no way of knowing that a message is hidden behind an image.

Sections 2 and 3 deal with the basic concept of Hill cipher and the generation of self invertible matrix. Section 4 presents the procedure for hiding the information in an image. In section 5, the results of quality of encryption and decryption are discussed while section 6 describes the concluding remarks.

2. HILL CIPHER

Hill cipher, developed by the mathematician Lester Hill in 1929, lies in the manipulation of matrix. For encryption, algorithm takes m successive plain text letters and substitutes m cipher letters. In Hill cipher, each character is assigned a numerical value like $\alpha = 0, \beta = 1, \ldots, \zeta = 25$ [5]. The substitution of cipher text letters leads to 'm' linear equation. This can be expressed $\alpha\sigma X = K\Pi$, $\omega\eta\varepsilon\rho\varepsilon$ C and P are column vectors, representing the plain text and cipher text respectively, and K is the encryption key matrix. All these operations are performed with modulo 26. Decryption requires using the inverse of the matrix K.

The inverse matrix K^{-1} of a matrix K is defined by the equation $K K^{-1} = K^{-1} K = I$, where I is the Identity matrix.

But the inverse of the matrix does not always exist, and when it does, it satisfies the preceding equation. K^{-1} is applied to the ciphertext, and then the plaintext is recovered. In general term, this is written as follows:

For encryption: $C = E_K(P) = K_P$

For decryption: $P = D_K(C) = K^{-1}C = K^{-1}K_P = P$

3. GENERATING SELF-INVERTIBLE MATRIX

As Hill cipher decryption requires inverse of the matrix, there arises a problem whether the inverse of the key matrix does exist or not during decryption. If the matrix is not invertible, then the encrypted text cannot be decrypted. In order to overcome this problem, the self-invertible matrix is used in Hill cipher. In the self-invertible matrix generation method, the matrix used for the encryption is itself self-invertible. So, at the time of decryption, it is not necessary to find the inverse of the key matrix. Moreover, this method eliminates the computational complexity involved in finding the inverse of the matrix during decryption. A is called self-invertible matrix if $A = A^{-1}$. The analysis presented here for generation of self-invertible matrix is valid for matrix of + ve integers that are the residues of modulo arithmetic on a prime number [1].

3.1 Generation of Self Invertible 4 □ 4 Matrix

Let $A = \left[\begin{array}{cc:cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right]$ be self-invertible matrix, partitioned as $A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right]$,

where $A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $A_{12} = \begin{bmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{bmatrix}$, $A_{21} = \begin{bmatrix} a_{31} & a_{32} \\ a_{42} & a_{42} \end{bmatrix}$, $A_{22} = \begin{bmatrix} a_{33} & a_{34} \\ a_{43} & a_{42} \end{bmatrix}$

Then, $A_{12}A_{21} = I - A_{11}^2$, $A_{11}A_{12} + A_{12}A_{22} = 0$,

$A_{21}A_{11} + A_{22}A_{21} = 0$ and $A_{21}A_{12} = I - A_{11}^2$

In order to obtain solution for all the four matrix equations, $A_{12}\,A_{21}$ can be factorized as,

$$A_{12}A_{21} = (I - A_{11})(I + A_{11})$$

So, if $A_{12} = (I - A_{11})k$ or $(I + A_{11})k$ then

$A_{21} = (I + A_{11})\frac{1}{k}$ or $A_{21} = (I - A_{11})\frac{1}{k}$, where k is a scalar constant.

Also $A_{11}A_{12} + A_{12}A_{22} = A_{11}(I - A_{11})k + A_{22}(I - A_{11})k$ or $k(A_{11} + A_{22})(I - A_{11})$

Hence $A_{11} + A_{22} = 0$ or $A_{11} = I$

As $A_{11} = I$ is a trivial solution, then, $A_{11} + A_{22} = 0$.

By solving 3rd and 4th matrix equations, the same solution is obtained.

Algorithm:

1. Select any arbitrary $\frac{n}{2} \times \frac{n}{2}$ matrix A_{22}
2. A_{11} is obtained from $A_{11} = -A_{22}$
3. Take $A_{12} = (I - A_{11})k$
4. Then, $A_{21} = (I + A_{11})\frac{1}{k}$
5. Form the matrix completely as,

Example: (For Modulo 26)

Take $A_{22} = \begin{bmatrix} 1 & 3 \\ 8 & 4 \end{bmatrix}$ then, $A_{11} = -A_{22} = \begin{bmatrix} 25 & 23 \\ 18 & 22 \end{bmatrix}$

Taking $A_{12} = I - A_{11}$ with k = 1, then $A_{12} = \begin{bmatrix} 2 & 3 \\ 8 & 5 \end{bmatrix}$ and $A_{21} = \begin{bmatrix} 0 & 23 \\ 18 & 23 \end{bmatrix}$

So the key matrix $A = \begin{bmatrix} 25 & 23 & 2 & 3 \\ 18 & 23 & 8 & 5 \\ 0 & 23 & 1 & 3 \\ 18 & 23 & 8 & 4 \end{bmatrix}$

3.2 Hill Cipher Encryption of an Image

Hill cipher can be adopted to encrypt grayscale and color images, For grayscale images, the modulus will be 256 (the number of levels is considered as the number of alphabets). In the case of color images, the color image is first decomposed into (R-G-B) components. Secondly, each component (R-G-B) is encrypted separately by the algorithm. Finally, the encrypted components are concatenated together to get the encrypted color image.

4. INSERTING ENCRYPTED TEXT IN A ENCRYPTED IMAGE

To hide the text inside the image, the character value of the text and the pixel value of the image are converted into streams of 8-bit binary. Two pixel pairs of the image in two adjacent rows are used to hide one character of the text. The four least significant bits of the selected pixel value in two adjacent rows are replaced respectively by the four least significant bits and four upper significant bits of one character of the text. The modified 8-bit binary numbers are converted to decimal number which gives the pixel value of the image, after hiding the text. To hide each character of secret message, two pixels are needed. So the number of characters that can be hidden in $(n \times n)$ image is given by the equation: Number of characters $\leq (n \cdot n) \div 2 - n$. In this equation, n pixels are subtracted because the secret text is not set in the first row of the cover image [2]. So start setting data from the second row of the cover image. The first row of the covered image is used to store specific data, like the position of last pixel in the covered image that contains secret data.

4.1 Reconstruction the Secret Text File

Reconstruction of the secret text message is performed by reversing the process used to insert the secret message in the container image.

5. QUALITY OF ENCRYPTION MEASURING FACTORS

One of the most important factors in examining the encrypted image is the visual inspection where the higher the disappearance of the main features is, the better the encryption algorithm will be. But, depending on the visual inspection alone is not enough in judging the complete hiding of the data image content. So the other measuring techniques are considered to evaluate the degree of encryption quantitatively. With the implementation of an encryption algorithm to an image, a change takes place in pixel values as compared to the values before encryption. Such a change may be irregular. This means that the higher the change in pixel values, the more effective the image encryption will be and hence the quality of encryption. So, the quality of encryption may be expressed in terms of the total deviation (changes) in pixel values between the original image and the encrypted one [4]. In addition to the visual inspection, three measuring quality factors will be considered to evaluate and compare among encryption algorithms. These factors are the maximum deviation, the correlation coefficient and irregular deviation.

5.1 The Maximum Deviation Factor

The maximum deviation measures the quality of encryption in terms of how it maximizes the deviation between the original and the encrypted images. The steps in this measure are as follows:

(1) Count the number of pixels of each grayscale value in the range from 0 to 255 and present the results graphically (in the form of curves) for both original and encrypted images (i.e. get their histogram distributions).
(2) Compute the absolute difference or deviation between the two curves and present it graphically.
(3) Count the area under the absolute difference curve, which is the sum of deviations (D) with this representing the encryption quality. D is given by the trapezoidal rule:

$$\mathrm{MDF} = \frac{h_0 + h_{255}}{2} + \sum_{i=1}^{254} h_i,$$

where h_i is the amplitude of the absolute difference curve at value i.

The higher the value of D, the better the encrypted image that deviated from the original image.

5.2 The Correlation Coefficient Factor

Correlation is a measure of the relationship between two variables. If the two variables are the image and its encryption, then they are in perfect correlation (i.e. the correlation coefficient equals one) if they are highly dependent (identical). In that case, the encrypted image is the same as the original image and the encryption process failed in hiding the details of the original image. If the correlation coefficient equals zero, then the original image and its encryption are totally different, i.e. the encrypted image has no distinct features and it is highly independent of the original image. So, the success of the encryption process means smaller values of the correlation coefficient (C.C), which is measured by the following equation:

$$CCF = \frac{cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^{N} (x_i - E(x))(y_i - E(y))}{\sqrt{\sum_{i=1}^{N} (x_i - E(x))^2} \sqrt{\sum_{i=1}^{N} (y_i - E(y))^2}}$$

where $E(x) = \frac{1}{N} \sum_{i=1}^{N} x_i$, x and y are grayscale pixel values of the original and encrypted images.

5.3 The Irregular Deviation Factor

The Irregular Deviation is a quality measuring factor which is based on how much the deviation caused by encryption (on the encrypted image) is irregular. It gives attention to each individual pixel value and the deviation is caused at every location of the input image. This method can be summarized in the following steps:

(1) Calculate the D matrix which represents the absolute values of the difference between each pixel values before and after encryption. So, D can be represented as $D = | I - J |$, where I is the input image, and J is the encrypted image.

(2) Construct the histogram distribution H of the absolute deviation between the input image and the encrypted image. So, H = histogram (D).

(3) Get the average value of how many pixels are deviated at every deviation value (i.e. the number of pixels at the Get the average value of how many pixels are deviated at every deviation value (i.e. the number of pixels at the histogram if the statistical distribution of the deviation matrix is of uniform distribution). This average value (DC) can be calculated as:

$$DC = \frac{1}{256} \sum_{i=0}^{255} h_i ,$$

where h_i is the amplitude of the absolute difference histogram at the value i.

(4) Subtract this average from the deviation histogram, and then take the absolute value of the result:

$$AC(i) = \square | H(i) - \square \square DC |$$

(5) Count the area under the absolute AC value curve, which is the sum of variations of the deviation histogram from the uniformly distributed histogram:

$$IDF = \sum_{i=0}^{255} AC(i).$$

The lower the MDF value, the better the encryption algorithm.

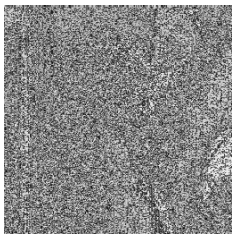
5.4 Results

Several kinds images have been evaluated. The results of the three measuring factors are given in Table 1 where MDF is the maximum deviation measure, CCF is the correlation coefficient measure, and IDF is the irregular deviation measure. For encryption quality the greater MDF is, the better; while for CCF the closer to zero the better, while for IDF the smaller, the better. For better text insertion method, the decrypted image should look like the original image. For this decryption quality, the smaller MDF the better; while for CCF the closer to one the better, while for IDF the larger, the better.

Table 1

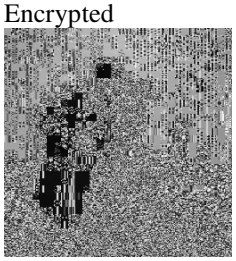
Chipher	MDF		CCF		IDF	
	Encryption	Decryption	Encryption	Decryption	Encryption	Decryption
Lena	22888	36	0.060676	0.999821	68186	130524
Cameraman	38297	36	0.304824	0.999933	73336	130526
Vegetables	17648	32	0.255587	0.999528	71004	130524
Toll	13465	18	0.749398	0.999864	94132	130536
Tools	29025	18	0.569205	0.999524	66390	130536


Lena Original

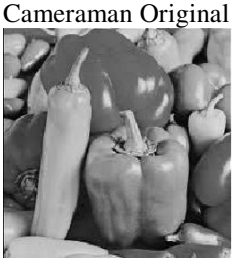

Encrypted

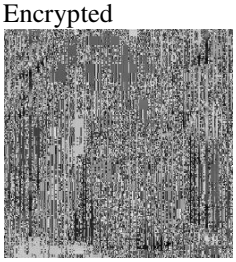

Text inserted and decrypted


Cameraman Original


Encrypted

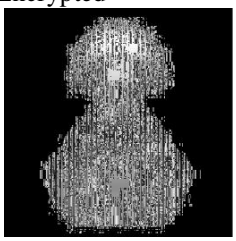

Text inserted and decrypted


Vegetables Original


Encrypted

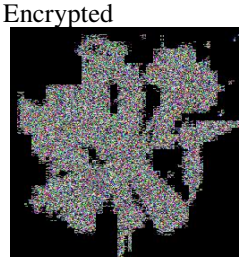

Text inserted and decrypted


Toy Original


Encrypted


Text inserted and decrypted


Tools Original


Encrypted


Text inserted and decrypted

6. CONCLUSION

In this paper, a novel method has been developed from Hill cipher by using self invertible matrix for encrypting text and image and decrypting. The encrypted text is hidden within the encrypted image which is decrypted to make it look like the original image. By increasing the order of the self invertible matrix, better encryption quality is obtained and the security level of the hidden text is ensured to the maximum.

7. REFERENCES

[1]. Bibhudendra Acharya, Girija Sankar Rath, Sarat Kumar Patra, Saroj Kumar Panigrahy. "Novel Methods of Generating Self-Invertible Matrix for Hill Cipher Algorithm", International Journal of Security, Vol 1, Issue 1, 2007, 14-21.

[2]. Habes. A, " Information Hiding in BMP image Implementation, Analysis and Evaluation", Information Transmission in Computer Networks Vol 6, Issue 1, 2006, 1 -10

[3]. Imai H., Hanaoka G., Shikata J., Otsuka A., Nascimento A.C., "Cyptography with Information Theoretic Security", Information Theory Workshop, 2002, Proceedings of the IEEE, 20-25 Oct 2002

[4]. Ismail I.A., AMIN Mohammed, DIAB Hossam, "How to repair the Hill cipher" Ismail et 2022 al. / J Zhejiang Univ SCIENCE A 2006 7(12):2022-2030

[5]. Stallings. W, "Cryptography and Network Security", 4th edition, Prentice Hall, 2005