

Secured Data Communication using Novel Modification to Hill Cipher Algorithm with Self Repetitive Matrix

Rahul. R. Ravan¹, Atul R. Nigavekar²

¹Student, Electronics, Kolhapur Institute of Technology, Kolhapur, Maharashtra, India
myrahul2004@yahoo.co.in

²Associate Professor, Electronics, Kolhapur Institute of Technology, Kolhapur, Maharashtra, India
atulrn@indiatimes.com

Abstract: *The Hill cipher algorithm is one of the symmetric key algorithms that have several advantages in data encryption. However, a drawback of this algorithm is that the inverse of the matrix used for encrypting the plaintext does not always exist. So, if the matrix is not invertible, the encrypted text cannot be decrypted. This paper presents a variant of the Hill cipher that overcomes this disadvantage. The proposed technique adjusts the encryption key to form a different key for each block encryption. The proposed variant yields higher security compared to the original one. Also in this paper, a method of generating self-invertible matrix for Hill Cipher algorithm has been proposed. In the self-invertible matrix generation method, the matrix used for the encryption is itself self-invertible. So, at the time of decryption, we need not to find inverse of the matrix. Moreover, this method eliminates the computational complexity involved in finding inverse of the matrix while decryption.*

Keywords: Hill Cipher, Encryption, Decryption, Self-invertible matrix, modified Hill Cipher.

1. Introduction

The desire to transmit messages securely is not new. For centuries, people have had a need to keep their communications private. Today, in the Information Age, as the Internet and other forms of electronic communication become more prevalent, electronic security is becoming increasingly important. Cryptography, the science of encryption, plays a central role in mobile phone communications, pay-TV, e-commerce, sending private emails, transmitting financial information, security of ATM cards, computer passwords and touches on many aspects of our daily lives [1]. Cryptography is the art or science encompassing the principles and methods of transforming an intelligible message (plaintext) into one that is unintelligible (ciphertext) and then retransforming that message back to its original form. In modern times, cryptography is considered to be a branch of both mathematics and computer science, and is affiliated closely with information theory, computer security, and engineering [2].

Cryptography systems can be broadly classified into: symmetric and asymmetric. Symmetric cryptosystems use the same key (the secret key) to encrypt and decrypt a message, and asymmetric cryptosystems use one key (the public key) to encrypt a message and a different key (the private key) to decrypt it. Asymmetric cryptosystems are also called public key cryptosystems.

Symmetric encryption is referred to as conventional encryption or single key encryption. Conventional encryption can be further divided into categories of classical techniques and modern techniques. The hallmark of conventional encryption is that the cipher or key to the algorithm is shared, i.e., known by the parties involved in the secured communication. Substitution Cipher is one of the basic components of classical ciphers. A substitution cipher is a method of encryption by which units of plaintext are substituted with ciphertext according to a regular system; the units may be single letters (the most common), pairs of

letters, triplets of letters, mixtures of the above, and so forth. The receiver decipheres the text by performing an inverse substitution [3]. The units of the plaintext are retained in the same sequence in the ciphertext, but the units themselves are altered. There are a number of different types of substitution cipher. If the cipher operates on single letters, it is termed a simple substitution cipher; a cipher that operates on larger groups of letters is termed polygraphic. A monoalphabetic cipher uses fixed substitution over the entire message, whereas a polyalphabetic cipher uses a number of substitutions at different times in the message— such as with homophones, where a unit from the plaintext is mapped to one of several possibilities in the ciphertext. Hill cipher is a type of monoalphabetic polygraphic substitution cipher.

Hill cipher is a block cipher that has several advantages such as disguising letter frequencies of the plaintext, its simplicity because of using matrix multiplication and inversion for enciphering and deciphering, its high speed, and high throughput [4]. But the drawback of this algorithm is that the inverse of the matrix used for encrypting the plaintext does not always exist. So, if the matrix is not invertible, the encrypted text cannot be decrypted. Moreover, Hill cipher can be easily broken with a known plaintext attack revealing weak security. This paper presents a variant of the Hill cipher that overcomes these disadvantages. This Modified Hill Cipher Algorithm, initially checks the matrix used for encrypting the plaintext, whether that is invertible or not. If the encryption matrix is not invertible, then the algorithm modifies the matrix such a way that its inverse exist. To overcome the weak security of the Hill algorithm, the proposed technique adjusts the encryption key to form a different key for each block encryption. Also in this paper, a method of generating self-invertible matrix for Hill Cipher algorithm has been proposed. In the self -invertible matrix generation method, the matrix used for the encryption is itself self-invertible. So, at the time of decryption, we need not to find inverse of the matrix. Moreover, this method eliminates the computational complexity involved in finding inverse of the matrix while decryption.

The paper is organized as follows. Following the introduction, the basic concept of Hill Cipher is outlined in section 2. Section 3 discusses about the modular arithmetic. In section 4, proposed Modified Hill Cipher Algorithm is presented. Finally, section 5 describes the concluding remarks.

2. Hill Cipher

Hill Cipher are an application of linear algebra to cryptography. It was developed by the mathematician Lester Hill. The Hill Cipher algorithm take m successive plain text letters and substitute m ciphertexts letter from them. The substitute is determined by m linear equation in which in which character is assigned a numerical value (a=0, b=1, ..., z=25). Let m be a positive integer, the idea is to take m linear combinations of the m alphabetic characters in one plaintext element and produce m alphabetic characters in one ciphertext elements. Then mxm matrix is used as a key of the system such that A is invertible module 26 [5]. Let a_{ij} be the entry of A. For the plain text block $x=(x_1, x_2, \dots, x_m)$ (the numerical equivalent of m letters) and a key matrix A, the corresponding ciphertext block $y=(y_1, y_2, \dots, y_m)$ can be computed as: Encryption: $(y_1, y_2, \dots, y_m) = (x_1, x_2, \dots, x_m)A \pmod{26}, \dots(1)$

Where,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}$$

The cipher text is obtained from a plaintext by means of a linear transformation. Decryption:

The reverse process deciphering is computed by $(y_1, y_2, \dots, y_m) = (x_1, x_2, \dots, x_m)A^{-1} \pmod{26}, \dots(2)$

Where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}$$

4. Proposed Modified Hill Cipher Algorithm

This algorithm generates the different key matrix for each block encryption instead of keeping the key matrix constant. This increases the secrecy of data. Also algorithm checks the matrix used for encrypting the plaintext, whether that is invertible or not. If the encryption matrix is not invertible, then the algorithm modifies the matrix such a way that it's inverse exist. The new matrix we obtain after modification of key matrix is called as Encryption matrix and with the help of this matrix encryption operation is performed. In order to generate different key matrix each time, the encryption algorithm randomly generates the matrix which is also used as a key.

For example:

$$\text{Key matrix, } K = \begin{pmatrix} K_{11} & K_{12} & K_{13} & K_{14} & K_{15} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} \\ K_{31} & K_{32} & K_{33} & K_{34} & K_{35} \\ K_{41} & K_{42} & K_{43} & K_{44} & K_{45} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} \end{pmatrix}$$

Where, $K = \text{seed number}$
 $K_{11} = (\text{seed number})$
 $K_{22} = (K_{11} * 2) \pmod{n}$
 \dots
 $K_{55} = (K_{44} * 5) \pmod{n}$

Where m is successive numbers of plaintext letters taken at time for encryption and n is length of the lookup table (total characters used for encryption and decryption) or we can set this n value as per requirement. Then with the help of random matrix, encryption matrix E is generated. Steps for encryption matrix generation are as follows:

1. Check whether the matrix K is invertible or not.
2. If inverse of matrix K does not exist, then adjust the diagonal elements (Increment the values of diagonal elements, one element at a time) so that the inverse of the resultant matrix (matrix obtained after changing diagonal elements) is invertible. This matrix becomes the Encryption matrix E.

In this algorithm if takes m successive plain text characters and substitutes for then m ciphertext characters. The substitution is determined by m linear equations in which each character is assigned a numerical value (we can take the characters ASCII equivalent numbers or we can assign a look up table. This case can be expressed in term of column vectors and matrices:

$$\begin{matrix} C1 = \\ C2 = \\ C3 = \\ C4 = \\ C5 = \end{matrix} \begin{pmatrix} E_{11} & E_{12} & E_{13} & E_{14} & E_{15} \\ E_{21} & E_{22} & E_{23} & E_{24} & E_{25} \\ E_{31} & E_{32} & E_{33} & E_{34} & E_{35} \\ E_{41} & E_{42} & E_{43} & E_{44} & E_{45} \\ E_{51} & E_{52} & E_{53} & E_{54} & E_{55} \end{pmatrix} \begin{pmatrix} P1 \\ P2 \\ P3 \\ P4 \\ P5 \end{pmatrix} \pmod{n}$$

Or $C=EP \pmod{n}$, where C and P are column vectors of length 3, representing th ciphertext and plain text respectively, and E is a 5.5 encryption matrix. All operations are performed mod n.

For decryption, from the random matrix once again similar way E matrix is generated. Decryption required using the modulo inverse of the matrix E. The inverse E^{-1} of matrix E is defined by the equation

$$E.E^{-1} = E^{-1}.E = I \quad \dots (4)$$

Where I is the matrix that is all zeros except for ones along the main diagonal from upper left to lower right. Hence decryption matrix D is generated by doing modulo inverse of encryption matrix. Multiply decryption matrix D with received ciphertext number vector C and then do modulo

operation. Then operate on the output resultant vector, substitute its equivalent characters and which is the plaintext.

We can explain this as

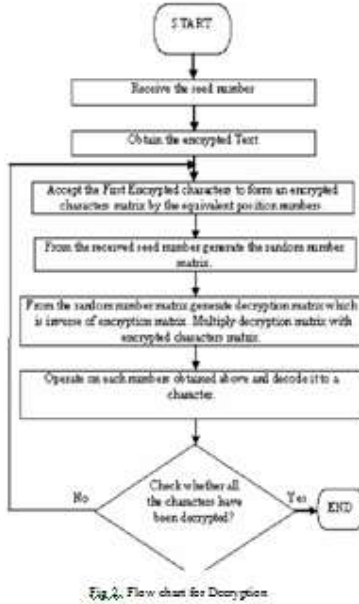
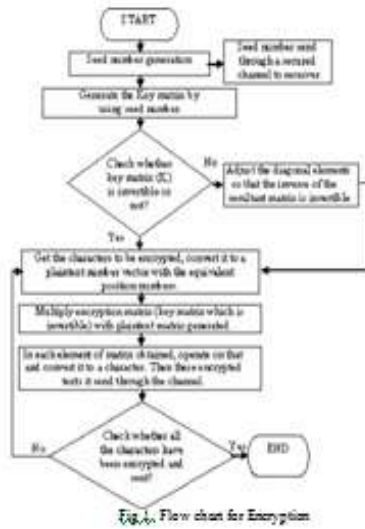
$$p=D.C=E^{-1} C \quad \dots(5)$$

In general algorithm can be expressed as follows:

$$C=EP \text{ mod } n \quad \dots(6)$$

$$P=E^{-1} C \text{ mod } n=E^{-1} EP=P \quad \dots(7)$$

The flow chart for the encryption and decryption methods are represented in figure 1 & 2



Example: (For Modulo 97)

Let $m = 5, n = 97$ and Seed number $S = 141$

Then, $K_{11} = 141$

$$K_{22} = (K_{11} * 2) \text{ mod } n$$

..

..

$$K_{55} = (K_{44} * 5) \text{ mod } n$$

Hence Key Matrix:

$$K = \begin{pmatrix} 141 & 0 & 0 & 0 & 0 \\ 0 & 90 & 0 & 0 & 0 \\ 0 & 0 & 78 & 0 & 0 \\ 0 & 0 & 0 & 24 & 0 \\ 0 & 0 & 0 & 0 & 24 \end{pmatrix}$$

Consider the plaintext to be encrypted is “event”. Letters of the plaintext are represented by their equivalent number vector (30 47 30 39 45)

Then with the help of key matrix, encryption matrix is generated. Encryption matrix we get as

$$K = \begin{pmatrix} 62 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 22 & 0 & 0 \\ 0 & 0 & 0 & 35 & 0 \\ 0 & 0 & 0 & 0 & 35 \end{pmatrix}$$

Then ciphertext for the plaintext is [17 88 78 7 23]

Decryption is done by doing inverse method of above and the, cipher text is converted to the original as “event”

Thus replacing the vector numbers (30 47 30 39 45) by their ASCII values we get the word “event”.

$$\begin{pmatrix} 30 \\ 47 \\ 30 \\ 39 \\ 45 \end{pmatrix} = \begin{pmatrix} 36 & 0 & 0 & 0 & 0 \\ 0 & 81 & 0 & 0 & 0 \\ 0 & 0 & 75 & 0 & 0 \\ 0 & 0 & 0 & 61 & 0 \\ 0 & 0 & 0 & 0 & 61 \end{pmatrix} \cdot \begin{pmatrix} 17 \\ 88 \\ 78 \\ 7 \\ 23 \end{pmatrix} \pmod{97}$$

5. Self-invertible Matrix Generation Method

As Hill cipher decryption requires inverse of the matrix, we suggest the use of self-invertible matrix generation method while encryption in the 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, we need not to find inverse of the matrix. A general method of generating self-invertible matrix is

Let $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ be an $n \times n$ Self-invertible matrix partitioned to $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$

A_{11} is a 1×1 matrix = $[a_{11}]$,
 A_{12} is a $1 \times (n-1)$ matrix = $[a_{12} \ a_{13} \ \dots \ a_{1n}]$

A_{21} is a $(n-1) \times 1$ matrix = $\begin{pmatrix} a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$ and
 A_{22} is a $(n-1) \times (n-1)$ matrix = $\begin{pmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$

So, $A_{12} A_{21} = I - A_{11}^2 = -1 - a_{11}^2 \dots (8)$
 and $A_{12} (a_{11}I + A_{22}) = 0 \dots (9)$
 Also, $a_{11} = -1$ (of the Eigen values of A_{22} other than 1)
 Since $A_{21} A_{12}$ is a singular matrix having the rank 1
 and $A_{21} A_{12} = I - A_{22}^2 \dots (10)$
 So, A_{22}^2 must have rank of $(n-2)$ with Eigen values +1 of $(n-2)$ multiplicity.
 Therefore, A_{22} must have Eigen values ± 1 .
 It can also be proved that the consistent solution obtained for element A_{21} & A_{12} by equating (10) term by term will also satisfy the equation (8).

A

- which has $(n-2)$ number of Eigen values of either +1 or -1 or both.
- ii. Determine the other Eigen value λ of A_{22} .
 - iii. Set $a_{11} = -\lambda$.
 - iv. Obtain the consistent solution of all elements of A_{21} & A_{12} by using (10).
 - v. Formulate the matrix.

Example: (For Modulo 13)

Let $A = \begin{pmatrix} 9 & 6 & 10 \\ 12 & 10 & 2 \\ 5 & 3 & 4 \end{pmatrix}$ which has Eigen values $\lambda = \pm 1, 10$

So, $A_{11} = [3]$,
 and one of the consistent solutions of $A_{12} = [11 \ 9 \ 4]$,
 and $A_{21} = \begin{pmatrix} 10 \\ 4 \\ 5 \end{pmatrix}$.

So, $A = \begin{pmatrix} 3 & 11 & 9 & 4 \\ 10 & 9 & 6 & 10 \\ 2 & 12 & 10 & 2 \\ 5 & 5 & 3 & 4 \end{pmatrix}$

Another consistent solution of $A_{12} = [1 \ 2 \ 11]$ and $A_{21} = \begin{pmatrix} 6 \\ 6 & 9 & 6 & 10 \\ 9 & 12 & 10 & 2 \\ 3 & 5 & 3 & 4 \end{pmatrix}$.

6. Conclusion

This paper presents a symmetric cipher that is actually a variation of the Hill cipher. The proposed algorithm is called Modified Hill Cipher Algorithm. This algorithm eliminates the drawback of using a random key matrix in Hill cipher algorithm for encryption, where we may not be able to decrypt the encrypted message, if the matrix is not invertible. As this algorithm uses a different key for each block encryption thereby significantly increases its resistance to various attacks. Also this paper suggests efficient methods for generating self-invertible matrix for Hill Cipher algorithm. These methods encompass less computational complexity as inverse of the matrix is not required while decrypting in Hill Cipher. These proposed methods for generating self-invertible matrix can also be used in other algorithms where matrix inversion is required.

References

- [1] G.R. Blakley, Twenty years of cryptography in the open literature, Security and Privacy 1999, *Proceedings of the IEEE Symposium*, 9-12 May 1999.
- [2] H. Imai, G. Hanaoka, J. Shikata, A. Otsuka, A.C. Nascimento, Cryptography with information theoretic security”, Information Theory Workshop, 2002, *Proceedings of the IEEE*, 20-25 Oct 2002.
- [3] A. J. Menezes, P.C. Van Oorschot, S.A. Van Stone, *Handbook of applied cryptography* (CRC press, 1996).
- [4] J. Overbey, W. Traves, J. Wojdylo, On the keyspace of the Hill cipher. *Cryptologia*, 29(1), 2005, 59-72.
- [5] K. Petersen, Notes on number theory and cryptography, 2000. [Http://www.math.unc.edu/Faculty/petersen/Coding/cr2.pdf](http://www.math.unc.edu/Faculty/petersen/Coding/cr2.pdf).
- [6] Barr T.H., Invitation to cryptography (Prentice Hall, 2002)
- [7] W. Stallings, *Cryptography and network security* (4th edition, Prentice Hall, 2005)