A Novel Cryptosystem Using Matrix Transformation

Bibhudendra Acharya, *Member, IEEE*, S. K. Patra, *Member, IEEE* and G. Panda, *Senior Member, IEEE*

***Abstract*--An improvement of the Hill cipher is proposed in this paper. The drawback of the Hill cipher 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. The proposed variant of the Hill cipher that overcomes these disadvantages. To overcome the drawbacks, the proposed cryptosystem uses randomly generated self-invertible matrix as an encryption key for each block encryption. Moreover, this method eliminates the computational complexity involved in finding inverse of the matrix while decryption.**

***Index Terms*--cryptosystem, encryption, decryption, Hill Cipher, Self-invertible matrix.**

I. INTRODUCTION

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, electronic commerce 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 deciphers 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.

The Hill cipher algorithm is one of the symmetric key algorithms that have several advantages in data encryption 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. In the proposed Cryptosystem 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. Hence, this cryptosystem eliminates the computational complexity involved in finding inverse of the matrix while decryption. To overcome the weak security of the Hill algorithm, the proposed technique generates the encryption key to form a different key for each block encryption.

The paper is organized as follows. Following the introduction, the basic concept of Hill cipher is outlined in section II. Section III discusses about the modular arithmetic. In section IV, proposed cryptosystem is presented. Finally, section V describes the concluding remarks.

B. Acharya, S. K. Patra, and G. Panda are with the Department of Electronics and Communication Engineering, National Institute of Technology Rourkela, India(e-mail: bibhudendra@gmail.com, {gpanda, skpatra}@nitrkl.ac.in).

II. Hill Cipher

Hill ciphers are an application of linear algebra to cryptology. It was developed by the mathematician Lester Hill. The Hill cipher algorithm takes m successive plaintext letters and substitutes m ciphertext letters for them. The substitution is determined by m linear equations in which each 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 element. Then, a $m \times m$ matrix A is used as a key of the system such that A is invertible modulo 26 [5]. Let a_{ij} be the entry of A. For the plaintext block $x = (x_1, x_2, ..., x_m)$ (the numerical equivalents of m letters) and a key matrix A, the corresponding ciphertext block $y = (y_1, y_2, ..., y_m)$ can be computed as

Encryption:

$$(y_1, y_2, ..., y_m) = (x_1, x_2, ..., x_m)\, A \quad (\text{mod } 26), \quad \dots (1)$$

Where

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

The ciphertext is obtained from the plaintext by means of a linear transformation.

Decryption:

The reverse process, deciphering, is computed by

$$(x_1, x_2, ..., x_m) = (y_1, y_2, ..., y_m)\; A^{-1} \quad (\text{mod } 26), \quad \dots (2)$$

Where

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{bmatrix}^{-1} (\text{mod} 26)$$

Since the block length is m, there are 26^m different m letters blocks possible, each of them can be regarded as a letter in a 26^m-letter alphabet. Hill's method amounts to a monoalphabetic substitution on this alphabet [6].

III. Modular Arithmetic

The arithmetic operation presented here are addition, subtraction, unary operation, multiplication and division. Based on this the self invertible matrix for Hill cipher algorithm is generated. The congruence modulo operator has the following properties [7]:

1. $a \equiv b \mod p$ if $n \mid (a - b)$
2. $(a \mod p) = (b \mod p) \Rightarrow a \equiv b \mod p$
3. $a \equiv b \mod p \Rightarrow b \equiv a \mod p$
4. $a \equiv b \mod p$ and $b \equiv a \mod p \Rightarrow a \equiv c \mod p$

Let $Z_p = [0, 1, ..., p - a]$ the set residues modulo p. If modular arithmetic is performed within this set Z_p, the following equations present the arithmetic operations:

1. Addition:
$(a + b) \mod p = [(a \mod p) + (b \mod p)] \mod p$
2. Negation:
$-a \mod p = p - (a \mod p)$
3. Subtraction:
$(a - b) \mod p = [(a \mod p) - (b \mod p)] \mod p$
4. Multiplication:
$(a * b) \mod p = [(a \mod p) * (b \mod p)] \mod p$
5. Division:
$(a / b) \mod p = c$ when $a = (b * c) \mod p$

The following Table I exhibits the properties of modular arithmetic.

TABLE I
Properties of Modular Arithmetic

Commutative Law: $(\omega + x) \mod p = (x + \omega) \mod p$ $(\omega * x) \mod p = (x * \omega) \mod p$
Associative Law: $[(\omega + x) + y] \mod p = [\omega + (x + y)] \mod p$
Distribution Law: $[\omega * (x + y)] \mod p = [\{(\omega * x) \mod p\} * \{(\omega * y) \mod p\}] \mod p$
Identities: $(0 + a) \mod p = a \mod p$ $(1 * a) \mod p = a \mod p$
Inverses: For each $x \in Z_p$, there exists y such that $(x + y) \mod p = 0$ then $y = -x$ For each $x \in Z_p$ there exists y such that $(x * y) \mod p = 1$

IV. Proposed Cryptosystem

As Hill cipher decryption requires inverse of the matrix, we suggest the use of self-invertible matrix generation method while encryption with the Hill Cipher. In the self-invertible matrix generation method, the matrix used for the encryption is itself self-invertible. So, in the proposed cryptosystem at the time of decryption, we need not to find inverse of the matrix Moreover in the proposed cryptosystem, algorithm generates the different key matrix for each block encryption instead of keeping the key matrix constant. This increases the secrecy of data. In order to generate different key matrix each time, the encryption algorithm randomly generates the seed number and from this key matrix is generated. A method of generating a random self-invertible even matrix is

The algorithm generates $n \times n$ matrix where n is even and utilized for generating a self-invertible matrix.

Let s be the seed for generating the random number,
t be the multiplier generating the random number,
p the modulus (necessarily to be a prime number), and
k a scalar constant.

Form a random matrix of $\frac{n}{2} \times \frac{n}{2}$ A_{11} with elements as

$a_{11} = s$

$a_{12} = st$

.
.
.

$a_{21} = st^{\frac{n}{2}}$

.
.
.

$a_{2n/2} = st^{n-1}$

Thus $a_{ij} = st^m$

where $m = (i-1)\frac{n}{2} + j - 1$, $1 \le i \le \frac{n}{2}$ and $1 \le j \le \frac{n}{2}$

Then form $A_{22} - A_{11}$

Set $a_{i+n/2,\, j+n/2} = -a_{ij}$ for $i, j = 1$ to $n/2$

Then form A_{12} as $A_{12} = k(I - A_{11})$

One can also take $k(I + A_{11})$ that means

$a_{i,j+\frac{n}{2}} = k(1 - a_{ij})$ *for* $i = j = -ka_{ij}$ for

$i \ne j$ with $i,\ j = 1$ *to* $n/2$

Form A_{21} as $A_{21} = \frac{1}{k}(I + A_{11})$

Thus $a_{i+n/2}, j = \frac{1}{k}(1 + a_{ij})$ for $i = j$ $= \frac{a_{ij}}{k}$ *for* $i \ne j$

$i \ne j$ with $i,\ j = 1$ *to* $n/2$

Finally formulate A

Example: For 6×6 random matrix (modulo 13)
s = seed value = 5, t = multiplier 7, k =3, then

$$A_{11} = \begin{bmatrix} 5 & 9 & 11 \\ 12 & 6 & 3 \\ 8 & 4 & 2 \end{bmatrix},\quad A_{22} = \begin{bmatrix} 8 & 4 & 2 \\ 1 & 7 & 10 \\ 5 & 9 & 11 \end{bmatrix},$$

$$A_{12} = 3(I - A_{11}) = \begin{bmatrix} 1 & 12 & 6 \\ 3 & 11 & 4 \\ 2 & 1 & 10 \end{bmatrix},$$

$$A_{21} = \frac{1}{3}[I + A_{11}] = \begin{bmatrix} 2 & 3 & 8 \\ 4 & 11 & 1 \\ 7 & 10 & 1 \end{bmatrix}$$

$$\text{So } A = \begin{bmatrix} 5 & 9 & 11 & 1 & 12 & 6 \\ 12 & 6 & 3 & 3 & 11 & 4 \\ 8 & 4 & 2 & 2 & 1 & 10 \\ 2 & 3 & 8 & 8 & 4 & 2 \\ 4 & 11 & 1 & 1 & 7 & 10 \\ 7 & 10 & 1 & 5 & 9 & 11 \end{bmatrix}$$

The flowcharts for the encryption & decryption methods are represented in Fig. 1 & 2.

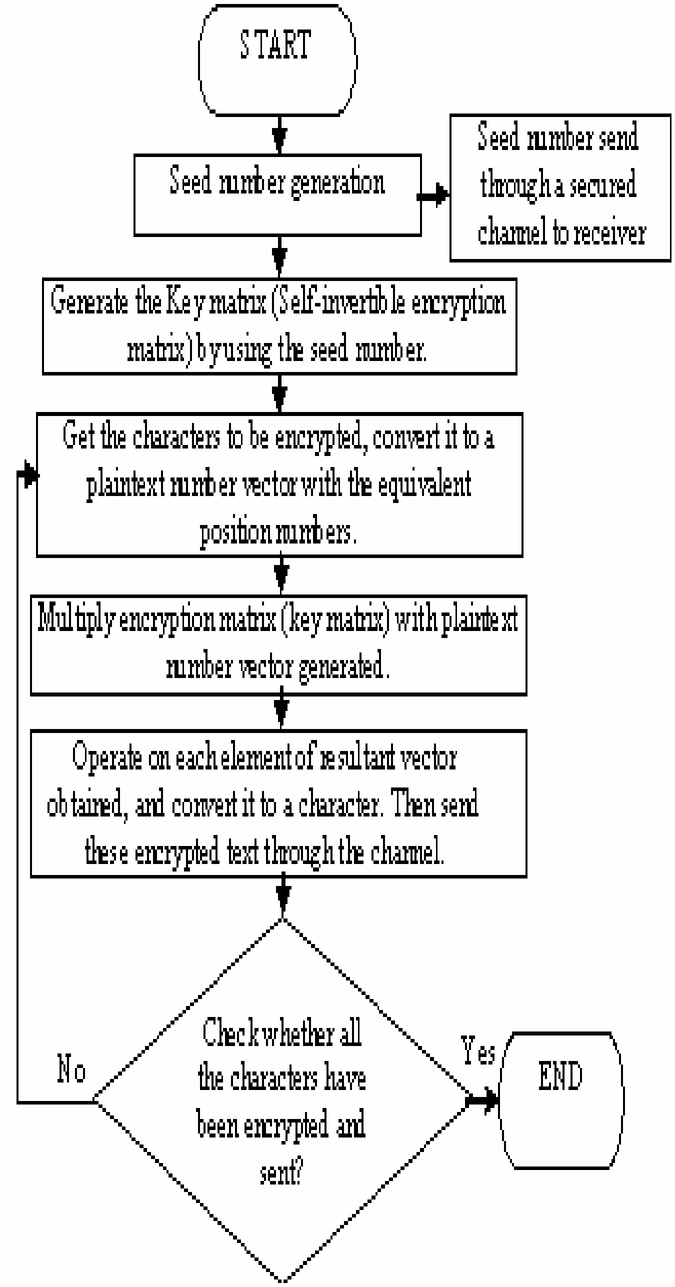

START
Seed number generation
Seed number send through a secured channel to receiver
Generate the Key matrix (Self-invertible encryption matrix) by using the seed number.
Get the characters to be encrypted, convert it to a plaintext number vector with the equivalent position numbers.
Multiply encryption matrix (key matrix) with plaintext number vector generated.
Operate on each element of resultant vector obtained, and convert it to a character. Then send these encrypted text through the channel.
Check whether all the characters have been encrypted and sent?
No
Yes
END

Fig 1. Flow chart for Encryption

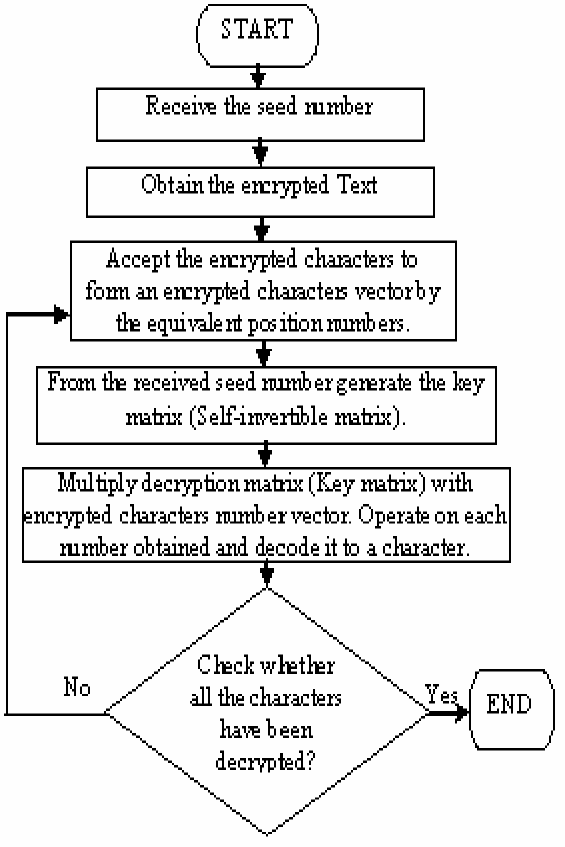

START
Receive the seed number
Obtain the encrypted Text
Accept the encrypted characters to form an encrypted characters vector by the equivalent position numbers.
From the received seed number generate the key matrix (Self-invertible matrix).
Multiply decryption matrix (Key matrix) with encrypted characters number vector. Operate on each number obtained and decode it to a character.
Check whether all the characters have been decrypted?
No
Yes
END

Fig 2. Flow chart for Decryption

V. Conclusion

This paper presents a symmetric cipher that is actually a variation of the Hill cipher. The proposed cryptosystem 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 cryptosystem uses a different key for each block encryption, thereby significantly increases its resistance to various attacks. This cryptosystem encompass less computational complexity, as inverse of the matrix is not required while decrypting in Hill Cipher. Also the proposed method for generating self-invertible matrix can also be used in other algorithms where matrix inversion is required.

VI. 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.

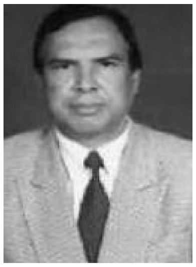
[6] Barr T.H., Invitation to cryptography (Prentice Hall, 2002)

[7] W. Stallings, *Cryptography and network security* (4th edition, Prentice Hall, 2005).

VII. Biographies

Bibhudendra Acharya received B.E. degree in Electronics & Telecommunication Engineering from Dr. B. A. Marathawada University, Aurangabad in 2002 and M.Tech. in Telematics and signal processing from National Institute of Technology, Rourkela in 2004. He has 2 years of teaching and 3 years of Industry experience. Currently, he is a Ph.D. student in Department of Electronics & Communication Engineering in NIT, Rourkela. He is a member of IEEE, IE (India), and ISTE (India).

Sarat Kumar Patra received B.E. degree in Electronics & Telecommunication Engineering from University College of Engineering, Burla in 1986 and masters degree in Electronics systems and Communication from Regional college of Engineering, Rourkela in 1992. He has received doctoral degree in 1998 from Edinburgh University, U.K. Currently, he is a professor in Department of Electronics & Communication Engineering at National Institute of Technology, Rourkela. He has 21 years of teaching and research experience. In the years from 1986 to 1989, he was in the Defense research and development Organization, India. He has more than 20 publications in National/ International Journals and conferences. His research areas of interest are Signal Processing, Soft Computing, Mobile Communication and Wireless Security. He is a member of IEEE, CSI, IE (India), ISTE (India), and IETE (India).



Ganapati Panda received Ph.D. degree in digital signal processing from IIT, Kharagpur, India, and post doctorate from University of Edinburgh, UK, in 1982 and 1984—1986, respectively. He has published more than 160 papers in referred research journals and conferences. He carries out research work in the field of DSP, soft-computing and digital communication. He is a fellow of INAE and National Academic of Science, India. Presently he is working as a professor in the Department of Electronics & Instrumentation Engineering at NIT, Rourkela, India.