Generation Of Keymatrix For Hill Cipher Encryption Using Quadratic Form

Dr. K. Mani, A. Barakath Begam

Abstract—Hill Cipher is a polygraphic encryption which uses matrices to transform blocks of plaintext letters into blocks of ciphertext. In Hill cipher, if the encryption key matrix called key matrix is not properly chosen, it is impossible to find the correct decryption matrix. Generally, the key matrix is chosen randomly, but it sometimes fails to form the correct key matrix. Further, obtaining a correct key matrix is not possible in a single run. As there is no deterministic procedure available in generating the key matrix, this paper presents the generation of same using quadratic form.

Index Terms— Hill Cipher, Keymatrix, Quadratic form and Equivalent quadratic form

1 INTRODUCTION

As the demand for effective information security is increasing day by day, security is the major concern to protect such data from adversary. Cryptography is one of the techniques to protect such data. It is a method of storing and transmitting data in a particular form so that only the intended recipient can read and process it. It provides various security services viz., confidentiality, integrity, authentication and non-repudiation. Confidentiality service is required for preventing disclosure of information to unauthorized parties. The authorized party will be unable to determine the keys that have been associated with encryption. The service data integrity provides assurance that the data has not been modified an unauthorized manner after it was created, transmitted or stored. Authentication is a service to recognize a user identity. Non-repudiation service prevents either sender or receiver from denying a transmitted message. Cryptographic algorithms broadly classified into two types viz., classical cryptography and modern cryptography. Classical cipher is a type of cipher that was used historically which is divided into transposition cipher and substitution cipher. In a substitution cipher, letters (or group of letters) are systematically replaced throughout the message for other letters (or group of letters) most probably the replaced letter is not present in the original plaintext. Caesar cipher is one of the popular substitution ciphers. In a transposition cipher, the letter themselves are kept unchanged, but their order within the plaintext is scrambled based on some well defined scheme. Modern cryptography uses various concepts of mathematics such as number theory, computational complexity theory and probability theory. It is divided into two types viz., symmetric-key cryptography and asymmetric-key cryptography. In symmetric-key cryptography, same key is used for both the encryption and decryption. Thus, before encryption is performed, the key must be known to both sender and receiver well in advance and the key management is the major issue. In asymmetric-key cryptography, two keys are involved viz., private-key and public-key. Public-key of the receiver is used for encryption and private-key of the receiver is used for decryption process. Symmetric-key is divided into two types viz., stream cipher and block ciphers. In stream cipher, a character or a letter or a bit is encrypted at a time whereas in block cipher group of letters called blocks are encrypted at a time. Vernam One-Time pad is an example of stream cipher; AES and DES are examples of block cipher encryption. The concept of public-key cryptography was invented by Whitfield Diffie and Martin Hellman and independently by Ralph Merkle. It is divided into three types viz., based on integers factorization, sum of subset problem and discrete logarithms. Hill Cipher is one of the popular classical encryption techniques which was invented by Lester S. Hill in 1929 [1] and technically it is a polygraphic substitution cipher. It can work on diagraphs, trigraphs or theoretically in any sized blocks. It uses linear algebra mathematics and modular arithmetic. In Hill Cipher, any block size may be selected but the block size is determined on the basis of the order of the key matrix, but it might be difficult to define good keys for enciphering large blocks. It uses matrices to transform blocks of plaintext letters into blocks of ciphertext.

The rest of the paper is organized as follows. The various works related to Hill Cipher encryption are presented in section 2. Mathematical concepts related to Quadratic Form (QF), equivalent QF etc., are discussed in section 3. Section 4 describes the proposed methodology in generating the key matrix of Hill Cipher encryption using QF. The proposed methodology is explained with an example section 5. The process of encryption and decryption of Hill Cipher using the generated key matrix and decryption matrix is presented in section 6. Finally, chapter 7 ends with conclusion.
2 Related Work

K. Mani and M. Viswambari [2] proposed a deterministic method to generate a keymatrix from Magic Rectangle (MR). For that, they framed some rules if the matrix taken from MR does not form a keymatrix. Bibhudendra et al., [3] proposed involutary, permuted and reiterative keymatrix generation method for Hill Cipher encryption. The keymatrix inversion problem is solved by involuntary matrix generation method. The Hill Cipher system’s security is enhanced considerably using permuted and reiterative keymatrix. Rushdi A. Hamamreh and Mousa F. [4] suggested a new technique in Hill Cipher algorithm to overcome its major problem non-invertible keymatrix. Further, they indicated that there will be no restriction on key generation or failure of choosing keymatrix which results in very difficult for the attacker to get the key but easier to choose and generate the key. Bibhudendra et al., [5] proposed an advanced Hill Cipher encryption to encrypt an image using a technique different from the conventional Hill Cipher. They proved that the proposed scheme is a faster encryption scheme which overcomes the problem of encrypting the images in homogeneous background. Further, they proved that the proposed scheme is resistant against known-plaintext attack.

L. Sreenivasulu Reddy [6] provided a new model of Hill Cipher using the non-quadratic residues to improve the security on Hill Cipher during encryption and they proved that the proposed algorithm is less vulnerable to known-plaintext attack. Andysah Putera et al., [7] proposed dynamic keymatrix generation method for Hill Cipher using genetic algorithm. In that they indicated that the result achieved by specifying some combinations of numbers which are used as a encryption key for Hill Cipher, it avoids the unnecessary numbers and the keymatrix is unique which should have the determinant value 1. Mani K. and Mahendran R. [8] proposed a deterministic procedure for keymatrix generation in Hill Cipher using classical encryption techniques by using the sequential advancement and permuted procedure methods.

3 MATHEMATICAL PRELIMINARIES

The following definitions and theorems are useful in understanding the concept of QF to generate the encryption matrix of Hill Cipher.

Definition 3.1: (Polynomial Expression) An expression is called a polynomial of \( n \) variables \( x_1, x_2, ..., x_n \) if the sum of terms each of which is the product of an integer and positive integral powers of selected variables.

Definition 3.2: (Homogeneous Polynomial) A polynomial of \( n \) variables is called homogeneous, if the sums of the exponents of the variables in each term are same. This common sum is called degree of the polynomial.[9][10]

Definition 3.3: (Quadratic Form) Given \( X = (x_1, x_2, ..., x_n)^t \)

and

\[
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}
\]

the function

\[
Q(X) = X^tAX = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j
\]

is called quadratic form. The matrix \( A \) can always be assumed symmetric because each element of every pair of coefficients \( a_{ij} \) and \( a_{ji} \) \( (i \neq j) \) can be replaced by \( \frac{(a_{ij}+a_{ji})}{2} \) without changing \( Q(X) \).

A quadratic form in \( n \) variables \( x_1, x_2, ..., x_n \) is denoted by \( Q(x_1, x_2, ..., x_n) \). Evidently, a QF \( Q(x_1, x_2, ..., x_n) \) has terms only of the form \( ax_i x_j \) for \( i \neq j \) and of the form \( bx_i x_i = bx_i^2 \). Thus, QFs with three variables are

\[
Q(x_1, x_2, x_3) = B_{11} x_1^2 + B_{22} x_2^2 + B_{33} x_3^2 + (B_{12} + B_{21}) x_1 x_2 + (B_{23} + B_{32}) x_2 x_3 + (B_{13} + B_{31}) x_1 x_3
\]

Eqn. (3.1) it is expressed as a matrix product using eqn. (3.2)

\[
Q(x_1, x_2, x_3) = x^t B x = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} B_{11} x_1 + B_{12} x_2 + B_{13} x_3 \\ B_{21} x_1 + B_{22} x_2 + B_{23} x_3 \\ B_{31} x_1 + B_{32} x_2 + B_{33} x_3 \end{bmatrix}
\]

\[
= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} B_{ij} = B \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
\]

...(3.2)

where \( x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \end{bmatrix} \) and \( B = [B_{ij}] \). The matrix product will always be a \( 1 \times 1 \) matrix whose only component is \( Q(x_1, x_2, x_3) \). Further, \( Q(x_1, x_2, x_3) = x^t B x \). The components \( B_{ij} \) and \( B_{ji} \) for \( i \neq j \) may be any numbers as long as the sum \( B_{ij} + B_{ji} \) has the value of the coefficient of \( x_i x_j \). Thus, the matrix \( B \) is not unique. If \( B_{ij} \) is chosen equal to \( B_{ji} \) when \( i \neq j \), then the matrix \( B \) will be symmetric and unique. Also, if the coefficients of \( Q(x_1, x_2, x_3) \) are integers and the coefficient of \( x_i x_j \) is even for \( i \neq j \), the corresponding symmetric matrix \( A \) such that \( Q(x_1, x_2, x_3) = x^t A x \) will have integer components.

Definition 3.4: (Positive Definite) A \( Q(X) \) is positive definite if the values of the principal minor determinants of \( A \) are positive (non-negative). In this case \( A \) is said to be positive definite [8]

Definition 3.5: If the QF \( [Q(x_1, x_2, x_3)] = x^t A x \), where \( x \) is the \( n \times 1 \)
The proposed methodology consists of two phases viz., (i) generation of keymatrix for Hill Cipher using equivalent QF, (ii) performing encryption and decryption.

### 4.1 Generation of Keymatrix Using QF

Once the QF with the condition specified in the above are taken, then the equivalent QF is formed denoted as $B_{ij}$ with $i \neq j$. It is a symmetric matrix and also unique. The coefficients of $Q(x_1, x_2, \ldots, x_n)$ are integer and it have the integer components. Once the equivalent quadratic matrix is formed and it is checked for modular inverse. If it has the modular inverse, i.e., $gcd([B_{ij}], 26) = 1$ it is considered as the keymatrix and then it is used for encryption. The modular inverse of encryption keymatrix called decryption matrix $K^{-1}$ which is then used for decryption.

### 5 PROPOSED METHODOLOGY – AN EXAMPLE

In order to understand the relevance of work, let the QF taken is $Q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + 2x_1x_2 - 2x_1x_3$ and it is expressed a matrix product as

$$Q(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} = x^tAx$$

Thus, $Q_1(y_1, y_2, y_3)$ is the encryption of $Q(x_1, x_2, x_3)$ with corresponding matrix $B = M^tAM$ so that $B_{11} = 1$ and if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \text{ then } x = My. \text{ To form the matrix } M \text{ eqn. (5.1) is used.}$$

$$A_{11}Q(x_1, x_2, x_3) = V^2 + Q'(x_2, x_3)$$

where $V = A_{11}x_1 + A_{12}x_2 + A_{13}x_3$ and $Q'(x_2, x_3) = A_{11}A_{22} - A_{12}^2, A_{11}A_{23} - A_{12}A_{13}$ and Further,

$$A_{11}A_{23} - A_{12}A_{13}, A_{11}A_{33} - A_{13}^2$$
2Q(x_1, x_2, x_3) = (2x_1 + x_2 - x_3)^2 + Q'(x_2, x_3)

...(5.2)

Q'(x_2, x_3) = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = x_2^2 + 2x_2x_3 + 3x_3^2

where \( Q'(x_2, x_3) = x_2^2 + 2x_2x_3 + 3x_3^2 \) has the minimum positive value of 1 at (1,0). When \( x_1 = -1 \), then \( V = |2x_1 + x_2 - x_3| = |2x_1 + 1 - 0| \) has minimum value -1. Thus, the values -1, 1 and 0 forms in the first column for matrix \( M \) and the rest of the column values of \( M \) are filled with integers so that \( \det(M) = 1 \). Thus,

\[
M = \begin{bmatrix} -1 & 0 & 2 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

and the matrix \( B \) is calculated using the formula \( B = M^TA \).

Now,

\[
B = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ -1 & -2 & 6 \end{bmatrix}
\]

and

\[
Q_1(y_1, y_2, y_3) = y_1^2 + y_2^2 + 6y_3^2 - 2y_1y_2 - 4y_2y_3 = (y_1 - y_2)^2 + y_2^2 + 5y_3^2 - 4y_2y_3.
\]

The minimum positive value 1 is obtained at (1,0) for

\[
Q_1'(y_2, y_3) = y_2^2 + 5y_3^2 - 4y_2y_3
\]

...(5.3)

The QF \( Q(x_1, x_2, x_3) \) is formed with the corresponding matrix \( C = M_1BM_1 \) so that if \( y = [y_1, y_2, y_3] \) and \( z = [z_1, z_2, z_3] \), then \( y = M_1z \) and \( C = I_3 \). Since, \( Q_1'(y_2, y_3) \) has the minimum value of 1 at (1,0), \( M_1 \) has the form as

\[
M_1 = \begin{bmatrix} 1 & v & w \\ 0 & 1 & s \\ 0 & 0 & t \end{bmatrix} \text{ where } \begin{bmatrix} 1 \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}
\]

To find the value of \( s \) and \( t \), using the formula \( y = M_1z \) from matrix \( B \).

\[
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & v & w \\ 0 & 1 & s \\ 0 & 0 & t \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}
\]

Thus,

\[
y_1 = z_1 + vz_2 + wz_3
\]

\[
y_2 = 0z_1 + z_2 + sz_3
\]

\[
y_3 = 0z_1 + 0z_2 + tz_3
\]

Substitute the values of \( y_2 \) and \( y_3 \) in eqn.(3), we get

\[
Q_1'(y_2, y_3) = z_2^2 + 2sz_2z_3 + 5tz_3^2 - 4z_2 - 4sz_3 - 4tz_2z_3 + tsz_3^2
\]

The coefficients of \( z_2z_3 \) are 2s - 4t. To find \( z_2z_3 \) term, set

\[2s - 4t = 0\]

...(5.4)

and \[\begin{bmatrix} 1 & s \\ 0 & t \end{bmatrix} = 1\]

i.e., \( t - 0s = 1 \)

...(5.5)

Solving eqn (4) and (5) we have \( s = 2 \) and \( t = 1 \)

Hence, \( M_1 = \begin{bmatrix} 1 & v & w \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \)

To find the value of \( v \) and \( w \), using the formula \( y = M_1z \) from matrix \( B \).

\[
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & v & w \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}
\]

\[y_1 = z_1 + vz_2 + wz_3\]

...(5.6)

\[y_2 = 0z_1 + z_2 + 2z_3\]

...(5.7)

\[y_3 = 0z_1 + 0z_2 + z_3\]

...(5.8)

From matrix \( B \), we know

\[V = y_1 - y_3\]

...(5.9)

Substitute eqns. (6) and (7) in eqn. (9), we get

\[V = vz_2 + wz_3 - z_3\]

...(5.10)

Now, The coefficient of \( z_2 \) is \( v \) and \( z_3 \) is \( w \) obtained from eqn. (9).

To find the values of \( z_2 \) and \( z_3 \), set the coefficients equal to 0, i.e., \( v = 0 \) and \( w - 1 = 0 \) and values of \( v \) is 0 and \( w \) is 1 and finally the matrix \( M_1 \) is formed as

\[
M_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}
\]

After obtaining equivalent QF, i.e., \( M_1 \) the equivalent QF matrix is called as keymatrix \( K \) and it is then used for encryption in Hill Cipher. This is because it satisfies the said criteria. Thus,

\[
K = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}
\]
6 ENCRYPTION/DECRYPTION USING HILL CIPHER – AN EXAMPLE

Let the plaintext $M$ be taken for encryption is “welcometoindia”, and $K$ obtained with order $3$ using section 5 is considered. Since, the order of $K$ is $3$, block size is also $3$ and hence $M$ is divided into various blocks as $p_1 = “wel”$; $p_2 = “com”$; $p_3 = “eto”$; $p_4 = “ind”$; $p_5 = “iax”$.

To encrypt the plaintext using Hill Cipher, we have $C_i = KP_i \mod 26$, $i = 1,...,5$ and to decrypt the ciphertext $P_i = K^{-1}C_i \mod 26$ , $i = 1,...,5$. Now,

$$G_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 22 \\ 4 \\ 11 \end{bmatrix} \mod 26 = \begin{bmatrix} 11 \\ 26 \\ 11 \end{bmatrix} = \begin{bmatrix} 11 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} L \\ A \\ L \end{bmatrix}$$

Similar computation can also be performed in computing other characters of the plaintext. Thus, the plaintext “welcometoindia” is encrypted as “LALQMMQVOFTDLUX”.

To perform decryption, first $K^{-1}$ is found using Gauss-Jordan method and it is computed as

$$K^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1 & R_2 \\ R_2 & R_3 \\ R_3 & R_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_1 \leftrightarrow R_1 \\ R_2 \leftrightarrow R_2 \\ R_3 \leftrightarrow R_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} K^{-1} \mod 26 \\ 0 & 1 & 24 \\ 0 & 0 & 1 \end{bmatrix}$$

If $KK^{-1} \mod 26 = I$, then the value of $K^{-1}$ is taken for decryption process.

$$P_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 24 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 11 \\ 264 \\ 11 \end{bmatrix} \mod 26 = \begin{bmatrix} 22 \\ e \\ l \end{bmatrix}$$

Similar computation can also be performed in computing other characters of the ciphertext. Thus, the ciphertext “LALQMMQVOFTDLUX” is decrypted as “welcometoindia”.

7 CONCLUSION

As there is no deterministic procedure available in generating the keymatrix for Hill Cipher encryption, a novel deterministic QF based keymatrix is generated in this paper. It is noted that for every QF there is an equivalent QF available if the QF matrix is positive definite. A $3 \times 3$ QF matrix is used in this paper to generate the keymatrix, an $n \times n$ keymatrix may be generated using an $n$ variables QF. The idea is unique and non-existent.

REFERENCES