Hill Cipher Modifications: A Detailed Review

Narendra B. Parmar, Dr.Kirit R. Bhatt
PG Student, Department of E & C, Sardar Vallabhbhai Patel Institute of Technology Vasad, Gujarat, India
Professor, Department of E & C, Sardar Vallabhbhai Patel Institute of Technology Vasad, Gujarat, India

ABSTRACT: Secured communication of information is important across the globe. Cryptography is one of the methods to attain security of information. The Hill cipher (HC) is one of the famous and known symmetric encryption algorithm based on linear matrix transformation. Hill Cipher has several advantages such as masquerading letter frequencies of the plaintext and high throughput. Despite the ease and speed of the Hill Cipher, the original Hill Cipher is no longer used due to the vulnerability against known plaintext-ciphertext attack. So to enhance security of Hill cipher various methods are proposed by different researchers. This paper aims to present review of various techniques used for Hill Cipher modification to improve performance of Hill Cipher in data security.

KEYWORDS: Symmetric Encryption, Hill Cipher, Cryptanalysis, Known plaintext attack

INTRODUCTION

In today’s world of digital communication sharing of information is increasing significantly via networks. The information being transmitted is vulnerable to various passive and active attacks. Therefore, the information security is one of the most challenging aspects of communication. Cryptography plays an important role in secure communication and it provides an excellent solution to offer the necessary protection against the data intruders. In the cryptography the original information is referred as plaintext and encrypted information is referred as cipher text. The transformation of plaintext into unintelligible data known as cipher text is the process of encryption. Decryption is the process of conversion of cipher text into plain text [13].

The Hill cipher was invented by L.S. Hill in 1929 [1, 2]. It is a famous a classical symmetric cipher based on matrix transformation. Hill cipher is a monoalphabetic polygraphic substitution block cipher. Hill Cipher has resistant towards frequency analysis, high speed and high throughput. Hill Cipher is vulnerable against known-plaintext attack. Hill cipher is a block cipher algorithm where plaintext is divided into equal size blocks. In a Hill cipher, the key is a non-singular matrix of size $b \times b$ where $b$ is the size of the block. Let us consider an arbitrary plaintext string of length $n$, defined over an alphabet of order $M$. We divide that plaintext into $l$ blocks of length $b$, where $b$ is an arbitrarily chosen positive integer and $l = \lfloor n / b \rfloor$. Now if the length $n$ is not a multiple of $M$, the last plaintext block must be padded with $n - b l$ extra characters. Each character in the alphabet is coded with a unique integer in $\{0,1,...,M-1\}$, in other words, all the characters in the alphabet are mapped to the ring $\mathbb{Z}_M$. The $n$ plaintext blocks can be rewritten as $l \times b$ matrix $P$ over $\mathbb{Z}_M$. The secret key matrix $K \ b \times b$ with coefficients in $\mathbb{Z}_M$. Hill encryption can be performed by computing

$$C = E_K (P) = KP \mod M \quad (1)$$

Similarly, decryption is performed by computing

$$P = D_K (C) = K^{-1}C \mod M \quad (2)$$

There might be some problems due to the fact that decryption require $K^{-1}$ over $\mathbb{Z}_M$. But in fact, those matrices $K$ with determinant 0, or with a determinant that has common factors with the modulus $M$, will be singular over $\mathbb{Z}_M$, and so they do not have inverse and therefore they will not be eligible as key matrices in the Hill cipher algorithm [1,2,3,4,14].

CRYPTANALYSIS ATTACKS

Cryptanalysis is a science and art of breaking the Cipher and gets knowledge of original information transmitted. Various types of cryptanalysis attacks are described as below:
a. Ciphertext-only attack: The cryptanalyst has the ciphertext of several messages, all have been encrypted using the same encryption algorithm. The cryptanalyst’s job is to recover the plaintext or to deduce the key used to encrypt the messages, in order to decrypt other messages encrypted with the same keys [13].

b. Known-plaintext attack: The cryptanalyst has access not only to the ciphertext of several messages, but also to the plaintext of those messages. His job is to deduce the key used to encrypt the messages or an algorithm to decrypt any new messages encrypted with the same key [13].

c. Chosen-plaintext attack: The cryptanalyst not only has access to the ciphertext and associated plaintext for several messages, but he also chooses the plaintext that gets encrypted. This is more powerful than a known-plaintext attack, because the cryptanalyst can choose specific plaintext blocks to encrypt, ones that might yield more information about the key [13].

d. Chosen-ciphertext attack: The cryptanalyst can choose different ciphertexts to be decrypted and has access to the decrypted plaintext. His job is to deduce the key [13].

To avoid security problems of original Hill Cipher various modification methods for Hill Cipher is proposed by researchers.


In Hill Cipher while decryption inverse Key matrix is required so if matrix is not invertible then it cannot be used as Key in Hill Cipher. This problem can be solved by using Involutory, Permuted and Reiterative Key Matrix Generation method for Hill Cipher.

Generation of Involutory matrix:

\[ A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix} \]

1. Select \((b - 1) \times (b - 1)\), a non-singular matrix \(A_{22}\) which has \((b - 2)\) number of Eigen values of either +1 or -1 or both.
2. Determine the other Eigen value \(\lambda\) of \(A_{22}\).
3. Set \(a_{11} = -\lambda\).
4. Obtain the consistent solution of all elements of \(A_{21} \& A_{12}\) by using the equation \(A_{21}A_{12} = I - A_{22}\).
5. \(A_{21} = A_{21} \times k\) and \(A_{12} = A_{12} / k\) where \(k\) any value from 1 to b-1
6. Finally formulate the matrix \(A\) by putting values in above matrix.

Formulation of Permuted Matrix:

This scheme used "random" permutations of columns and rows of a matrix to form a "different" key for each block data encryption. \(P_{ki}\) be a permutation matrix generated as

\[
P_{ki} = \begin{cases} 
1 & \text{for } \begin{cases} 
i = k, j = l \\
i = l, j = k \\
i = j, i \neq k, i \neq l 
\end{cases} 
0 & \text{otherwise}
\end{cases}
\]

Thus \(P_{ki}A\) exchanges \(k^{th}\) row with \(l^{th}\) row and \(k^{th}\) column with \(l^{th}\) column. So \(k^{th}\) and \(l^{th}\) diagonal elements of original matrix \(A\) are exchanged.
Generation of Reiterative Matrix:
1. Let $A$ be a diagonal matrix with diagonal elements $d_i = 1, 2, \ldots, b$ where $d_i \neq j$ when $i \neq j$. Then determine the smallest $n_i$ such that $d_i^{n_i} = 1$.
2. Calculate $n = \text{LCM}(n_1, n_2, \ldots, n_b)$
3. Pick up any random invertible square matrix $B$
4. Generate $C = B^{-1}AB$
5. Then $n$ calculated in the step 2 will be smallest integer such that $C^n = I$.

Secunity Analysis: Involutary matrices eliminate necessity of matrix inverses for Hill decryptions. So that same machinery can be used both for encryption and decryption of messages and reduce time required for decryption in Hill cipher scheme. Permuted matrix and reiterative key generation method generates “different” key for each block of data encryption, thereby significantly increases its strength against various cryptanalysis attacks [5].

b. Hill cipher modification based on Affine Transform and one way Hash function [6]
Mohsen Toorani and Abolfazi Falahati proposed cryptosystem that includes a ciphering core as shown in Figure 1, and a one-pass protocol which is shown in Figure 2. Each block of data is encrypted using a random number. For avoiding multiple random number generations, only one random number is generated at the beginning of encryption and the corresponding random number of the following data blocks is recursively generated using a one-way hash function in a hash chain, as it is depicted in Figure 1. The basic random number that is generated prior to the encryption should be securely shared between the participants. This can be done using the introduced one pass protocol that is depicted in Figure 2 where the encryption and decryption procedures should be followed from Figure 1.

Encryption
$a_i = H(a_{i-1})$

if $a_i \neq 0; v_0 = a_i \pmod{p}$

if $a_i = 0; v_0 = 1$

$f = (v_{i+1} \mod n) + 1$

$v_i = k_0 + \tilde{v}_{i-1}a_i \pmod{p}$

$Y = v_0XK + V \pmod{p}$

Decryption
$a_i = H(a_{i-1})$

if $a_i \neq 0; v_0 = a_i \pmod{p}$

if $a_i = 0; v_0 = 1$

$f = (v_{i+1} \mod n) + 1$

$v_i = k_0 + \tilde{v}_{i-1}a_i \pmod{p}$

$X = v_0^{-1}(Y - V)K^{-1} \pmod{p}$

Figure 1: Ciphering core [6]

Encryption
$a_0 \in \mathbb{R} [1, p-1]$

$b \in \mathbb{R} [1, n^2]$

$i = \lceil b/n \rceil$

$j = b - n(i - 1)$

$r = a_0k_j \pmod{p}$

encryption:
$Y = E_{K}(X \rightarrow (Y, a_0, r))$

Decryption

Decryption:

Figure 2: One pass protocol [6]

Encryption process:
1. Sender secretly selects random integers $a_0$ and $b$ where $0 < a_0 < p - 1$ and $1 < b < n^2$ in which $n$ is the rank of the key matrix.
2. Calculate \( r = a_0k_i \mod p \) where \( i = \lceil b / n \rceil \) and \( j = b - n(i - 1) \) in which \( \lceil \cdot \rceil \) denotes the ceiling function.

3. The plaintext message is encoded into some row vectors \( x = [x_1 \; x_2 \; \ldots \; x_n] \). For the \( t \)-th block of data to be encrypted ( \( t = 1, 2, \ldots \) ), \( a_t \) is calculated with a recursive expression as \( a_t = H(a_{t-1}) \) in which \( H(.) \) denotes the one-way hash function. If \( a_t \) is invertible mod \( p \), i.e. \( a_t \neq 0 \mod p \), then \( v_0 = a_t \mod p \). Otherwise, \( v_0 = 1 \).

4. The row vector \( v = [v_1 \; v_2 \; \ldots \; v_n] \) where \( v_i = k_{ij} + \tilde{v}_{i-1}a_i \mod p \) for \( i = 1, \ldots, n \) and \( j = (v_{i-1} \mod n) + 1 \)

\[
\tilde{v}_{i-1} = 2^{\lceil \log_2 v_{i-1} \rceil} + (v_{i-1} \mod 2^{\lceil \log_2 v_{i-1} \rceil})
\]

where \( \chi = \lceil \log_2 v_{i-1} \rceil + 1 \) which denotes the bit-length of \( v_{i-1} \) and \( \lfloor \cdot \rfloor \) indicates the floor.

5. Encrypt text

\[ Y = v_0XK + V \mod p \] (3)

6. The procedure is repeated until all blocks of plaintext become encrypted.

Decryption process:

Receiver computes \( u = k_{ij}^{-1} \mod p \) and \( a_0 = ru \mod p \) in which \( i = \lceil b / n \rceil \) and \( j = b - n(i - 1) \). He uses \( a_0 \) for decrypting the ciphertext \( X = v_0^{-1}(Y - V)K^{-1} \mod p \) (4)

Security Analysis: The proposed cryptosystem provides the security against the known plaintext attack since \( n \) equations cannot be used for solving an unknown \( n \times n \) matrix and 2\( n \) unknown parameters. Choosing a large prime number \( p \) as the modulus extremely enhances the keyspace so the brute-force or the ciphertext-only attack does not break the cipher. The random number after a secure transmission is recursively encoded with the one-way hash function so it differs for each block of plaintext. The chosen-ciphertext and chosen-plaintext attacks are also thwarted since the random number \( a_0 \) that its knowledge is essential for such attacks, is exchanged through a secure protocol [6]. But Liam Keliher and Anthony Z. Delaney develop Chosen Plaintext attack which breaks the secure variant of Hill Cipher [7].

c. Hill Cipher modification based on H-S-X (Hill-Shift-XOR)[8]

Hill cipher algorithm cannot effectively encrypt images that contain large areas of a single color. Bibhudendra Acharya et al. have proposed a novel technique which is a modified version of Hill cipher algorithm for image encryption named H-S-X (Hill-Shift-XOR) which can be applied to any type of images whether they are color or gray. Figure 3. shows block diagram of H-S-X.

![Figure 3: Block diagram of H-S-X [8]](image-url)
Algorithm H-S-X:
Step1: 2-level Hill Cipher is applied to original Image using a random matrix having odd determinant.

Step2: Appropriate $P_1$ and $P_2$ values are selected for the row and column shifting functions where, $P_1$ and $P_2$ are generators for the elements co-prime to congruence modulo $n$ and $m$ respectively where, $n$ is no. of columns in the original image and $m$ is no. of rows in the original image.

2a: $i^{th}$ row pixel values circularly right shifted according to the formula

$$P_1^{i+1} + \text{ceil}(i/b) \cdot K_{r(i \mod b)} \pmod{n}$$

where, $i$ – Corresponding row number, $b$ – Size of key matrix used for shifting, $K_{r(i \mod b)}$ – Sum of the values of the $i^{th}$ row of the key matrix, ceil() – The Ceiling function.

2b: $j^{th}$ row pixel values circularly down shifted according to the formula

$$P_2^{j+1} + \text{ceil}(j/b) \cdot K_{c(j \mod b)} \pmod{m}$$

Where, $j$ – Corresponding column number, $K_{c(j \mod b)}$ – Sum of the values of the $j^{th}$ column of the key matrix.

Step3: Block wise XOR operation is performed onto resultant image using the key matrix or one of its permutations or a masked version of the key.

Security analysis: H-S-X algorithm is more secure to brute force attacks as compared to original Hill cipher algorithm. A Brute Force Attack requires $2^{8 \cdot n^2}$ number of key generations; where $n$ is the order of key matrix. This scheme is resistant against known plaintext attacks due to the shifting steps involved in Step2. It is also resistant to Chosen Plaintext attacks, if the H-S-X steps are repeated. H-S-X scheme is slow compared to original Hill cipher [8].

d. Hill Cipher modification based on Bitwise operation [9]

Ahmed Desokyi and Anju Panicker Madhusoodhanan propose Hill Cipher modification for application of binary data. The bitwise matrix multiplication is used for encryption and decryption process in which when values in the matrices are multiplied, bitwise AND is used and when values are added bitwise XOR is used. The Key matrix generation, encryption and decryption algorithm are shown below:

Key Generation Algorithm:
1. Generate a $b \times b$ random binary matrix
2. Find its inverse
3. If the matrix is not invertible, go to Step 1
4. Store the key and its inverse
5. Repeat Steps 1 to 4 for 8 times for 8 different planes
6. Encryption Algorithm:

Encryption of plaintext is performed using bitwise multiplication of Plaintext and Key matrix by converting plaintext into series of planes as shown in figure 4. The plaintext $[P]$ can be written as:

$$[P] \rightarrow [P_i]; i = 1, 2, \ldots, 8$$

To encrypt 8 planes, we generate 8 different keys randomly. The keys $[K_i], i = 1, 2, \ldots, 8$, are $b \times b$ binary matrices that are invertible. Ciphertext $[C_i] = [P_i][K_i] \pmod{2}$; $i = 1, 2, \ldots, 8$.

Decryption Algorithm:

Decryption of ciphertext is performed using bitwise multiplication of ciphertext and inverse Key matrix by converting ciphertexts into series of planes as shown in figure 5.
**Security Analysis:** The plaintext is converted into a series of planes and each plane is encoded using a different key matrix so known plaintext attack is infeasible. Other known methods of attacks like chosen plaintext and chosen ciphertext attacks also become increasingly difficult as the plaintext and ciphertext are scrambled as a result of using them as planes. Hill Cipher does not preserve the statistics of the plaintext. Therefore it is extremely difficult to break a Hill Cipher code with ciphertext only attack [9].

**e. Hill cipher modification based on circulant matrices [10]**

Adinarayana Reddy K et al. proposes modification to Hill cipher based on circulant matrices, where a prime circulant matrix is shared as a secret key and a non-singular matrix G chosen as a public key such that the determinant of coefficient matrix $G_a$ is zero. The algorithm as follows:

1. Select a $b \times b$ non-singular matrix G in GF (M) as a public key such that $\det (G_a) = 0$.
2. Select a $b \times b$ prime circulant matrix A in GF (M) as a secret key.

**Calculate key**

$$K = AGA^{-1} \mod M$$  \hspace{1cm} (5)

**Encryption:**

$$C_i = KP_i + V_i^T \mod M$$  \hspace{1cm} (6)

Where, $P_i$ is $i^{th}$ plaintext block of size b, $C_i$ is $i^{th}$ cipher text block, $V_i$ is $i^{th}$ row of the prime circulant matrix A

**Decryption:**

$$K^{-1} = AG^{-1}A^{-1} \mod M$$  \hspace{1cm} (7)

$$P_i = K^{-1}(C_i - V_i^T) \mod M$$  \hspace{1cm} (8)

Where, $K^{-1}$ is inverse Key matrix, $V_i$ is $i^{th}$ row of the prime circulant matrix A

**Security Analysis:** For each plaintext block encryption we use a different vector V by rotation. This enhances performance of Hill Cipher against known-plaintext attack and chosen-plaintext attack. It also overcomes the ciphertext-only attack, since the modulus is a prime number. This algorithm is easily implementable. It reduces the key storage requirement from $b^2$ elements of $Z_M$ to just $b$ elements, because matrix is fully specified by its first row. It also reduces the running time required to compute matrix multiplication. Security of the proposed cryptosystem is
based on the difficulty of solving multivariable polynomial equations i.e. $K = AGA^{-1}$ mod $M$. It is difficult to solve if the modulus is large prime number [10].

**f. Hill Cipher modification based on Binary LU Encryption (BLUE) [11]**

Abdelahaliem Abbas Othman proposed the method, Binary LU Encryption (BLUE), matrices are defined over binary field $\{0, 1\}$. Since the modulus is 2, multiplication and addition operations can be replaced by binary “AND” and “XOR” respectively. Using binary matrices with logical operations for encryption and decryption, instead of matrices of integers and modular arithmetic gives high performance and allows for easy hardware implementation of BLUE.

Process shown below:

1. Generate two invertible, $b \times b$ lower and upper triangular matrices, $L$ and $U$, with ones in the main diagonal in order to ensure the existence of the inverse of each matrix.
2. Set $S \leftarrow LU$ and compute $S^{-1}$
3. Set $Q \leftarrow SS$ and $Q^{-1} \leftarrow S^{-1}S$
4. $Q$ is an encryption operator and $Q^{-1}$ is the decryption operator.
5. **Encryption**

   $$C_i = Q(P_i \oplus K_i)$$  \hspace{1cm} (9)

   Where $P_i, K_i$ and $C_i$ are $i^{th}$ message, key and ciphertext respectively

6. **Decryption**

   $$P_i = (Q^{-1}C_i \oplus K_i)$$  \hspace{1cm} (10)

**Security Analysis:** The relation between Plaintext ($P$), Key matrix ($K$), encryption operator ($Q$) and Ciphertext ($C$) is many-to-one so all kinds of cryptanalysis attacks like Ciphertext only, Known plaintext, Chosen plaintext and Chosen cipher text are found to be useless for proposed system [11].

### III. CONCLUSION

In this paper various techniques for Hill Cipher modification are described which provide security against various cryptanalysis attacks like Ciphertext only, Known plaintext, Chosen plaintext and Chosen cipher text. But as we increase security strength of Hill Cipher the computation cost and complexity of algorithm is increases. The Computational Cost of different Algorithms for encryption and decryption of each block of data is given below, where $n$ is rank of key matrix without considering inverse key calculation which is required for decryption:

**Table 1: Computational Cost of Different Algorithms for each block of data encryption and Decryption:**

<table>
<thead>
<tr>
<th>Different Algorithms</th>
<th>Operation</th>
<th>Multiplication</th>
<th>Addition</th>
<th>Inverse</th>
<th>Hash</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Hill Cipher</td>
<td>Encryption</td>
<td>$n^2$</td>
<td>$n^2 + n$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Decryption</td>
<td>$n^2$</td>
<td>$n^2 + n$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Involuntary, Permuted and Reiterative Key Matrix Generation Methods for Hill Cipher System*</td>
<td>Encryption</td>
<td>$n^2$</td>
<td>$n^2 + n$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Decryption</td>
<td>$n^2$</td>
<td>$n^2 + n$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Taroomi’s Secure variant</td>
<td>Encryption</td>
<td>$n^2 + 2n$</td>
<td>$n^2 + n + 1$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Decryption</td>
<td>$n^2 + 2n$</td>
<td>$n^2 + n + 1$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Binary LU Encryption</td>
<td>Encryption</td>
<td>$n^2$ (bitwise AND)</td>
<td>$n^2$ (bitwise XOR)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Decryption</td>
<td>$n^2$ (bitwise AND)</td>
<td>$n^2$ (bitwise XOR)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
*Self Invertible Key matrix is used so no inverse matrix calculation is required for Decryption

REFERENCES

Papers:

Books:

Websites:

BIOGRAPHY

Narendra B. Parmar is a PG student in Department of Electronics & Communication Engineering at Sardar Vallabhbhai Patel Institute of Technology Vasad, Gujarat, India. He received Bachelor of Engineering in Electronics & Communication Engineering (BE EC) degree in 2012 from GTU, Gujarat, India.

Dr.Kirit R. Bhatt is a Professor in Department of Electronics & Communication Engineering at Sardar Vallabhbhai Patel Institute of Technology Vasad, Gujarat, India. He received Bachelor of Engineering in Electronics Engineering degree in 1997 and Master of Engineering in Microprocessor system and application degree in 2001 from MS Uni. Baroda, Gujarat, India. He achieved Ph.D degree in 2013 from MS Uni.