DOI: http://dx.doi.org/10.26483/ijarcs.v9i1.5206

Volume 9, No. 1, January-February 2018

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

TRANSPOLY HILL CIPHER — AN IMPROVEMENT OVER TRADITIONAL HILL CIPHER

Ashish Pandey
Department of Computer Science and Engineering
KNIT, Sultanpur
Sultanpur, India

Stuti Pandey
Department of Computer Science and Engineering
KNIT, Sultanpur
Sultanpur, India

Abhay Kumar Agarwal
Department of Computer Science and Engineering
KNIT, Sultanpur
Sultanpur, India

Abstract: Every encryption technique is based on any one of two facts- substitution or transposition. Sometimes the combination of both techniques is used. Substitution technique is based on replacement of given plaintext letters by other letters or symbols to get the ciphertext. The basic concept behind this paper is to improve the strength of Hill cipher, 'a multiple-letter substitution cipher' against known plaintext attack. This paper presents a modified approach which is an enhancement over traditional Hill cipher by using symmetric matrix of plaintext letters.

Keywords: substitution cipher; transposition cipher; hill cipher; symmetric matrix; cryptanalysis; diffusion; confusion

I. INTRODUCTION

Substitution and transposition are two basic mechanisms used in encryption techniques. Transposition is permuting the positions of the given plaintext letters to be encrypted. While substitution technique is replacement of given plaintext letters by other letters or symbols to get the ciphertext. Two methods of substitution cipher exist for encryption:

a) First approach involves encrypting multiple letters of plaintext.

b) Second approach involves the use of multiple cipher alphabets [1]. For e.g. Hill Cipher.

Encryption algorithm of Hill Cipher takes 'n' consecutive plaintext letters and replaces for them 'n' ciphertext letters. The replacement is driven by 'n' linear equations and each letter is assigned a numerical value (a=0, b=1,…..z=25). The main concept behind the robustness of Hill cipher is the use of large matrix size [2, 3]. By using large matrix size, it conceals the frequency information which is more common for single-letter encryption. Hill cipher can be easily broken with a known plaintext attack but it is robust against a ciphertext-only attack.

In the presented paper, a modified approach is proposed for Hill cipher by means of a symmetric matrix of plaintext letters. The concept of transposition cipher, diffusion and confusion is also used to improve the strength of Hill cipher against cryptanalysis or known plaintext attack.

The next section presents the concept and definition of traditional Hill cipher. Subsequently proposed method including its algorithm is presented. Lastly, conclusions and future scope is presented.

II. AN OVERVIEW ABOUT HILL CIPHER

Encryption algorithm of Hill cipher takes 'n' consecutive plaintext letters and replaces them with 'n' ciphertext letters. The replacement is driven by 'n' linear equations in which each letter is assigned a numerical value (a=0, b=1….z=25).

For m=3, the system can be described as follows:

$c_1 = (k_{11}p_1 + k_{12}p_2 + k_{13}p_3) \bmod 26$ (1)

$c_2 = (k_{21}p_1 + k_{22}p_2 + k_{23}p_3) \bmod 26$ (2)

$c_3 = (k_{31}p_1 + k_{32}p_2 + k_{33}p_3) \bmod 26$ (3)

Where p_1, p_2 and p_3 are numerical values corresponding to plaintext letters. k_{11}, k_{12}…….k_{32}, k_{33} are elements of key matrix (square matrix) used in encryption. c_1, c_2 and c_3 are numerical values corresponding to ciphertext letters calculated after encryption [4].

These equations can be expressed in terms of matrices as follows:

$$= \quad \bmod 26 \qquad (4)$$

i.e. **C = KP mod 26**

Here 'C' is a ciphertext matrix of order 3×1 and 'P' is plaintext matrix of order 3×1. 'K' is an encryption key matrix of order 3×3. All operations are performed for mod 26.

For example, consider the plaintext "paymoremoney" and use the encryption key **K =** [1]

The first three letters of the plaintext (m = 3) are represented by the vector P = = . Since numerical value of p = 15, a = 0 and y = 24. Then using eq. (4)

$$= \quad \bmod 26$$

$$= \quad \bmod 26 =$$

i.e. we get c_1 = 11, c_2 = 13 and c_3 = 18. Hence, ciphertext letters corresponding to numerical values of c_1, c_2 and c_3 are L,

N and S respectively. Similarly, the ciphertext for the entire plaintext is "LNSHDLEWMTRW".

Decryption process is demonstrated as follows:

$$\mathbf{P = K^{-1}C \bmod 26} \quad (5)$$

Where 'C' and 'P' is column matrix of length 3, representing the ciphertext and plaintext. Here K^{-1} plays the role of decryption key of order 3×3 which is obtained by taking the inverse of encryption key 'K'. For any square matrix, if its inverse exist then it must satisfy the equation $K K^{-1} = K^{-1}K = I$, where I is the identity matrix [5]. Hence, the encrypted text "LNSHDLEWMTRW" requires the inverse of the encryption key matrix 'K' for decryption.

In this case, the inverse is:

$\mathbf{K^{-1}} = \begin{pmatrix} 4 & 9 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix}$. So Decryption of encrypted text "LNSHDLEWMTRW" is as follows:

The first three letters of the ciphertext (m = 3) are represented by the vector $C = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 11 \\ 13 \\ 18 \end{pmatrix}$.

Then using eq. (5)

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 4 & 9 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix} \begin{pmatrix} 11 \\ 13 \\ 18 \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 431 \\ 494 \\ 570 \end{pmatrix} \bmod 26 = \begin{pmatrix} 15 \\ 0 \\ 24 \end{pmatrix}.$$

$p_1 = 15$, $p_2 = 0$ and $p_3 = 24$. Hence, plaintext letters corresponding to numerical values of p_1, p_2 and p_3 are p, a and y respectively. Similarly, the plaintext for the entire ciphertext is "paymoremoney". It is easily seen that if the matrix $\mathbf{K^{-1}}$ is applied to the ciphertext, then the plaintext is recovered.

The main concept behind the robustness of Hill cipher is the use of large matrix size. For an 'n×n' Hill cipher, suppose we have 'n' plaintext-ciphertext pairs, each of length n. we label the pairs

$P_y = \begin{pmatrix} p_{1y} \\ p_{2y} \\ \vdots \\ p_{ny} \end{pmatrix}$ and $C_y = \begin{pmatrix} c_{1y} \\ c_{2y} \\ \vdots \\ c_{ny} \end{pmatrix}$ such that $C_y = KP_y$ for $1 \leq y \leq n$ and for some unknown key matrix K. Now define two 'n×n' matrices $R = (p_{xy})$ and $S = (c_{xy})$ [1, 6]. Then we can form the matrix equation S = KR. If inverse of R exists, then we can calculate $K = SR^{-1}$. If R is non-invertible, then a new version of R can be formed with additional plaintext-ciphertext pairs until an invertible R is obtained.

III. PROPOSED METHOD

As it is known that the Hill cipher can be easily broken with a known plaintext attack but it is robust against a ciphertext-only attack. The proposed method focuses on making Hill cipher robust against known plaintext attack. It is achieved by using the concept of diffusion and confusion along with transposition.

Proposed encryption algorithm is as follows:

Step-1: Take first m-letters (p_1, p_2,….p_m) from plaintext and calculate its equivalent numerical values to build a symmetric matrix of order mXm. We get matrix 'P'.

Step-2: Multiply key matrix 'K' of order mXm with the matrix obtained in step-1.

Step-3: After multiplication perform mod 26 operation on the result obtained in step-2. The resultant matrix after mod 26 operations is the ciphertext matrix 'C'of order mXm.

Step-4: Convert the numerical values in ciphertext matrix 'C' into the letters (alphabets).

Step-5: Finally, apply transposition on ciphertext letters by using shifting operation (circular shift, Shifting may be either one letter or two letters or so on).

Algorithm 1: Encryption algorithm

In proposed method, we diffuse the m-plaintext column vector (column matrix) into a mXm symmetric matrix of plaintext. So the plaintext column vector $P = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$ can be diffused into symmetric plaintext matrix $P = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_3 & p_1 \\ p_3 & p_1 & p_2 \end{pmatrix}$ (for m=3)

Considering again the same plaintext "paymoremoney" and same encryption key K for proposed encryption algorithm, the new ciphertext matrix is calculated as follows:

$\mathbf{P} = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_3 & p_1 \\ p_3 & p_1 & p_2 \end{pmatrix} = \begin{pmatrix} 15 & 0 & 24 \\ 0 & 24 & 15 \\ 24 & 15 & 0 \end{pmatrix}$ and encryption key $\mathbf{K} = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$.

According to algorithm 1

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_3 & p_1 \\ p_3 & p_1 & p_2 \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix} \begin{pmatrix} 15 & 0 & 24 \\ 0 & 24 & 15 \\ 24 & 15 & 0 \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} 375 & 483 & 663 \\ 819 & 747 & 774 \\ 486 & 333 & 78 \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} 11 & 15 & 13 \\ 13 & 19 & 20 \\ 18 & 21 & 0 \end{pmatrix}$$

According to step-4, numerical values in ciphertext matrix are converted into letters. Hence ciphertext letters for plaintext "pay" is "LPNNTUSVA" (written row-wise).

Finally transposition using shifting operation (circular shift) is applied on ciphertext letters. For example: 2-letters shifting is applied on above ciphertext letters. The result is VALPNNTUS.

Same process is applied for remaining letters of plaintext. After diffusion of 3-plaintext letter of column vector in 3X3 plaintext symmetric matrix, we get 9-letters ciphertext matrix. All the letters of ciphertext matrix are completely or almost different to each other. So by using this diffusion, we create confusion for attacker to perform cryptanalysis attack i.e. known plaintext-ciphertext pairs attack.

Proposed decryption algorithm is as follows:

Step-1: First apply transposition on ciphertext letters by using shifting operation in reverse order (circular shift). Shifting of letters should be same as in encryption end but in reverse direction.

Step-2: Take m^2-letters ($c_1, c_2, \ldots . c_m^2$) from ciphertext and calculate its numerical equivalent values for ciphertext matrix 'C' of order mXm.

Step-3: Multiply key inverse matrix K^{-1} of order mXm with matrix obtained in step 2.

Step-4: After multiplication perform mod 26 operation on the result. The resultant matrix after mod 26 operations is the plaintext symmetric matrix 'P' of order mXm.

Step-5: Take 'm' numerical values either from first row or first column (which are same) of symmetric plaintext matrix. Convert the numerical values into the letters. These are the plaintext letters that has been sent.

Algorithm 2: Decryption algorithm

Decryption of the ciphertext "VALPNNTUS" obtained by algorithm 1 is as follows:

According to step-1 of decryption algorithm, apply transposition by shifting operation (circular shift) on ciphertext letters "VALPNNTUS" in reverse direction. The result is "LPNNTUSVA". So ciphertext matrix

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} 11 & 15 & 13 \\ 13 & 19 & 20 \\ 18 & 21 & 0 \end{pmatrix}$$ and decryption key

$$\mathbf{K^{-1}} = \begin{pmatrix} 4 & 9 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix}.$$

According to algorithm 2

$$\begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_3 & p_1 \\ p_3 & p_1 & p_2 \end{pmatrix} = \begin{pmatrix} 4 & 9 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix} \begin{pmatrix} 11 & 15 & 13 \\ 13 & 19 & 20 \\ 18 & 21 & 0 \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_3 & p_1 \\ p_3 & p_1 & p_2 \end{pmatrix} = \begin{pmatrix} 431 & 546 & 232 \\ 494 & 674 & 535 \\ 570 & 717 & 312 \end{pmatrix} \bmod 26$$

.

$$\begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_3 & p_1 \\ p_3 & p_1 & p_2 \end{pmatrix} = \begin{pmatrix} 15 & 0 & 24 \\ 0 & 24 & 15 \\ 24 & 15 & 0 \end{pmatrix}$$ i.e. original plaintext symmetric matrix. Hence plaintext letters for ciphertext "LPNNTUSVA" is "pay".

IV. Conclusions and Future Scope

In the presented paper, plaintext column vector is diffused into a symmetric plaintext matrix of order mXm. By using this diffusion concept, m-plaintext letters produces m^2 different or almost different ciphertext letters. Again on applying circular shift on m^2 ciphertext letters creates confusion for intruders to perform cryptanalysis. Cryptanalysis is based on some knowledge about plaintext-ciphertext pairs. But in proposed method, diffusion of plaintext letters plus transposition of ciphertext letters makes cryptanalysis almost impossible to perform. For an intruder, mapping from m-plaintext letters to m^2 circular shifted ciphertext letters are cumbersome. Hill cipher is robust enough against ciphertext-only attack. So, using proposed method it can also easily thwart cryptanalysis based on statistical analysis.

In future, analysis of other methods can be presented. Alternative encryption and decryption algorithm can also be developed.

V. References

[1] W. Stallings, Cryptography and Network Security, 4th ed., Pearson Education India, pp. 42–45.

[2] Behrouz A. Forouzan and D. Mukhopadhyay, CRYPTOGRAPHY AND NETWORK SECURITY, 2015, Mc Graw Hill India.

[3] S. Bose and P. Vijayakumar, CRYPTOGRAPHY AND NETWORK SECURITY, 1st ed., 1 March 2016, Pearson Education India.

[4] B. Menezes, Network Security and Cryptography, 1st ed., 1 April 2010, Cengage.

[5] Keith M. Martin, Everyday Cryptography: Fundamental Principles and Applications, Oxford University Press India.

[6] W. Mao, Modern Cryptography: Theory & Practice, Pearson Education