GENERALIZED AFFINE TRANSFORMATION BASED ON CIRCULANT MATRICES

Adi Narayana Reddy K[a], Vishnuvardhan B[b], Durga Prasad K[c]

[a] Department of CSE, Hyderabad Institute of Technology and Management, Hyderabad, India.
aadi.iitkgp@gmail.com
[b] Department of IT, JNTU Jagityala, Karimanagar.
mailvishnu@yahoo.co.in
[c] Department of IT, Padmasri Dr BVRIT, Narsapur, Medak.
dpcse007@gmail.com

Abstract

The secure transmission of any form of data over a communication medium is prime important across the globe or in research arena. Cryptography is a branch of cryptology and it provides security for data transmission between any communicating parties. The Hill cipher is one of the symmetric key substitution algorithms. Hill Cipher is vulnerable to known plaintext attack. This paper presents an enhancement to the Hill cipher by utilizing the circulant matrices. The proposed technique shares a prime circulant matrix as a secret key and we choose a non-singular matrix as a public key in such way that the determinant of the coefficient matrix is zero. Computational cost shows that the proposed technique is efficient and it thwarts all the security attacks.

Keywords:

Circulant Matrix, Determinant, Hill Cipher, Primitive Root, Substitution Cipher.

1. Introduction

Today, information is one of the most valuable assets. Information transmission across the network is of prime importance in the present age. Cryptography is the branch of cryptology and it provides security to the transmitted data between the communicating parties. There are various algorithms to provide security for the information. Traditional symmetric ciphers use substitution in which each character is replaced by other character. Lester S. Hill invented the Hill cipher in 1929. Hill cipher is a classical substitution technique that has been developed based on linear transformation. It has both advantages and disadvantages. The main advantages are disguising letter frequencies of the plaintext; high speed, high throughput, and the simplicity because of using matrix multiplication and inversion for enciphering and deciphering. The disadvantages are, it is vulnerable to known plaintext attack and the inverse of the shared key matrix may not exist always. To overcome the drawbacks of Hill cipher algorithm many modifications are presented. In our paper we present a modification to the Hill cipher by the utilization of special matrices called circulant matrices. A circulant matrix is a matrix where each row is rotated one element to the right relative to the preceding row vector. In literature circulant matrices are used in many of the cryptographic algorithms. Advanced Encryption Standard (AES) uses circulant matrices to provide diffusion at bit level in mix columns step. Circulant matrices can be used to improve the efficiency of Lattice-based cryptographic functions. Cryptographic hash function Whirlpool uses circulant matrices.

The paper is systematized accordingly: Section 2 presents an over view of Hill cipher modifications. Section 3 presents a proposed Hill cipher modification. Section 4 explains security analysis. Conclusion of the proposal is in the section 5.

2. Literature Review on Hill Cipher Modifications

Classical substitution is one of the mechanisms to provide security for data. Hill cipher is a polyalphabetic substitution cryptographic system based on simple linear transformation. This is a block cipher and the plaintext is divided into blocks and each block is encrypted by using same shared key matrix. The encryption of the plaintext block P is defined as $C = PK$ (mod m) in which C is the ciphertext block, K is an n x n key matrix where $k_{ij} \in Z_m$ in which Z_m is ring of integers modulo m where m is a natural number that is greater than one. The value of modulus m was 26 in the original Hill cipher but its value can be optionally chosen. The participants securely shares key matrix K. The decryption of ciphertext C is defined as $P = CK^{-1}$ (mod m). All operations are performed over Z_m. The decryption of Hill cipher algorithm is possible only if the key matrix K is invertible or equivalently, the GCD (det K (mod m), m) = 1. Many of square matrices, generally, are not invertible over Z_m. According to Overbay [20], the key space of the Hill cipher is GL (N, Z_m), the group of n x n matrices that are invertible over Z_m. As it is proved that, when $m = \prod_i p_i^{e_i}$ is a composite modulus, we have:

$$| GL(n, Z_m) | = \prod_i \left(p_i^{(e_i - 1)n^2} \prod_{k=0}^{n-1} (p_i^n - p_i^k) \right)$$

and the proposition of n x n invertible matrices will be:

$$f(n, m) = \prod_i \prod_{j=1}^{n} (1 - \frac{i}{p_i^j})$$

The probability of randomly selected any square matrix to be invertible is about 1 for any large prime modulus, while it is almost 0 for a composite modulus with many different prime divisors, so the risk of determinant having common factors with the modulus can be reduced by taking a prime number as modulus. Thus, the key space of a prime modulus is larger than composite modulus.

Several researches have been done to improve the security of Hill cipher. Yi-Shiung Yeh [25] presented a new polygraph substitution algorithm based on different bases. Their algorithm uses two co-prime base numbers that are securely shared between the participants. Although their algorithm thwarts the known-plaintext attack, requires many mathematical manipulations. It is time consuming and is not efficient for dealing bulk data. Saeednia [22] tried to make Hill cipher secure by using dynamic key matrix obtained by random permutations of columns and rows of the master key matrix and transfers an encrypted plaintext and encrypted permutation vector to the receiving side. The numbers of dynamic keys are generated n! Where n refers the size of the key matrix. Each plaintext is encrypted by a new key matrix that prevents the known-plaintext attack on the plaintext but it is vulnerable to known-plaintext attack on permutation vector, the same vulnerability of original Hill cipher. Chefranov [4, 5] proposed a modification to [22] that works similar to Hill cipher permutation method, but it does not transfer permutation vector, instead both sides use a pseudo-random permutation generator, and only the number of the necessary permutation is transferred to the receiver. The number of dynamic keys is the same as [22]. Ismail [13] tried to improve the security of Hill cipher by introduction of an initial vector that multiplies each row of the current key matrix to produce the corresponding key of each block but it has several inherent security problems. Lin Ch [16, 17] claimed that taking random numbers and using one-way hash function thwarts the known-plaintext attack to the Hill cipher but their scheme is vulnerable to chosen-ciphertext attack. Mohsen Toorani [18, 19] proposed a symmetric cryptosystem based on affine transformation. It uses one random number and generates other random numbers recursively using HMAC in chain. Ahmed Y Mahmoud [2] proposed a modification to Hill cipher based on Eigen values HCM-EE. The HCM-EE

generates dynamic encryption key matrix by exponentiation with the help of Eigen values but it is time consuming. Circulant matrices are playing major role in diffusion. We propose an enhanced Hill cipher based on circulant matrices and it reduces the space and time complexity. The comparisons of time complexities were presented in appendix.

3. Proposed Cryptosystem

The proposed algorithm is developed on the concept of circulant matrices. Once the key is calculated and is used for encryption of every plaintext block along with random number. We will start with the following basic concepts

Basic Concepts

Circulant Matrix: A circulant matrix is a matrix where each row rotates one element to the right relative to the preceding row vector. Thus a circulant matrix can be written as

$$\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \cdots & c_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}$$ and is denoted by $\text{circ}(c_0, c_1, c_2, \cdots c_{n-1})$

Prime circulant matrix:

An n x n circulant matrix is prime circulant if gcd of row vector is 1. For example the 4 x 4 circulant matrix with row vector (a, b, c, d) is prime circulant if gcd (a, b, c, d) = 1.

Coefficient matrix:

Let G be a matrix and the coefficient matrix of G is denoted as G_c and is defined as circ (circ (row 1), circ (row 2) . . . circ (row n)) where row 1, row 2. . . row n are row vectors of matrix G and circ (row i) is the circulant matrix of row i. For example if G is a 2 x 2 matrix then G_c is 4 x 4 matrix

$$G = \begin{bmatrix} g_1 & g_2 \\ g_3 & g_4 \end{bmatrix}$$

$$G_c = \begin{bmatrix} g_1 & g_2 & g_3 & g_4 \\ g_2 & g_1 & g_4 & g_3 \\ g_3 & g_4 & g_1 & g_2 \\ g_4 & g_3 & g_2 & g_1 \end{bmatrix}$$

Proposed Affine Hill Cipher

The proposed cryptosystem is developed based on circulant matrices. A prime circulant matrix is shared secretly by the participants and a non-singular matrix G is chosen as a global key (i.e. public key) such that the determinant of the coefficient matrix G_c is zero. In order to give randomization and to strengthen it against chosen-ciphertext and chosen-plaintext attacks, each plaintext block is encrypted by a random number. The proposed cryptosystem uses a large prime number p and a primitive root g such that $1 < g < p$ in which p and g are publicly available to anyone. The algorithm follows as:

1. Alice and Bob shares a prime circulant matrix secret key A

2. Choose a publicly available non-singular matrix G such that the determinant of the coefficient matrix G_c is zero over mod p.
3. Alice generates a random number r where $1 \le r \le n$ in which n is size of the row.
4. To encrypt t^{th} block (t = 1, 2, 3 …), Alice computes random numbers r_t recursively by using the expression $r_t = g^{r_{t-1}}$ where $r_0 = a_r$, in which a_r is the r^{th} element of the secret key matrix A.
5. Alice computes $V = [v_1, v_2, \cdots, v_n]$ where $v_i = r_1 * a_i \pmod p$ for i= 1, 2, 3,…, n and $v_i = r_t * v_i \pmod p$ where for t = 2,3, ...
6. Sender computes $K = AGA^{-1} \pmod p$
7. The t^{th} block of plaintext M_t can be encrypted as $C_t = r_t M_t K + V_t \pmod p$ for (t = 1, 2, 3 …) where C is ciphertext.
8. Sender transmits (C, r) to the receiver.
9. Receiver computes $K^{-1} = AG^{-1}A^{-1} \pmod p$, vector V using random number r and also computes r_t recursively.
10. The t^{th} block of ciphertext C_t can be decrypted as $M_t = r_t^{-1}(C_t - V_t)K^{-1} \pmod p$ where $r_t^{-1} \pmod p$ is the inverse of the recursively generated random number r_t.

The proposed cryptosystem neutralizes all the security drawbacks of Hill cipher. It thwarts the known-plaintext attack since n equations cannot be used to solve an unknown n x n matrix and n unknown parameters. Choosing a large prime as modulus increases the key space and it over come ciphertext-only attack. The random numbers are generated by using exponentiation. The chosen-ciphertext and chosen-plaintext attacks are also thwarted since the knowledge of random number r_0 is essential for such attacks.

Computational Cost

In this section, the time complexity of the proposed algorithm is evaluated. Let T_{Enc} and T_{Dec} denote the running time for encryption and decryption of each block of plaintext respectively. In our proposed algorithm we compute key once and that key will be used for all the blocks of plaintext. We will not consider the time complexity of key generation and we have:

$$T_{Enc} \cong (n^2 + n)T_{Mul} + n^2 T_{Add}$$

$$T_{Dec} \cong (n^2 + n)T_{Mul} + n^2 T_{Add}$$

In which, T_{Add} and T_{Mul} are the time needed for the for the calculation of scalar modular addition and multiplication respectively. The total processing time for enciphering /deciphering whole blocks of plaintext /ciphertext can be estimated simply multiplying the running time of each block of data by total number of blocks. The total number of blocks depends on length of input data. If the length of data is L alphabets and is not multiple of n, it should be padded until it becomes multiple of n so the number of blocks is $\left\lceil \frac{L}{n} \right\rceil$. The running time for encrypting the whole plaintexts is:

$$T_{Total\_Enc} \cong \left\lceil \frac{L}{n} \right\rceil ((n^2 + n)T_{Mul} + n^2 T_{Add})$$

The running time for decrypting the whole ciphertexts is:

$$T_{Total\_Dec} \cong \left\lceil \frac{L}{n} \right\rceil ((n^2 + n)T_{Mul} + n^2 T_{Add})$$

Table 1 (in Appendix) gives comparison between the required number of operations for encrypting/decrypting each block of plaintext/ciphertext in the proposed algorithm and those of other algorithms.

4. Security Analysis

The proposed algorithm provides security in two levels. In the first level, the matrix K, vector V and random number r_t are used for encryption of the plaintext block. In this vector V and r_t are computed for each block of plaintext and K is computed once and used for encryption of all the blocks. It thwarts the known plaintext attack since n equations cannot be used to solve an unknown n x n matrix K and n + 1 unknown parameters of row vector V and random number r_t. The random numbers are generated recursively as $r_t = g^{r_{t-1}}$ for t = 1,2,3… where $r_0 = a_r$ and r is a shared random number. It also thwarts chosen-plaintext and chosen-ciphertext attacks, since the knowledge of random number is must required for these attacks. According to chosen-ciphertext attack/ chosen-plaintext attack, Eve has access to corresponding plaintext of the chosen ciphertexts. For example Eve has access to a cryptographic module that automatically performs decryption. Now Eve as a set of equations $C_t = r_t M_t K + V_t \bmod p$, t = 1, 2, 3… n+1 where C_t and M_t are known parameters. To solve the equations the knowledge of r_t is must. Choosing a large prime number p as a modulus increases the key space so that the brute-force attack or equivalently, ciphertext-only attack does not have any benefit for the attacker. In the second level, the security is based on the difficulty of solving multivariable polynomial equations i.e. solving of the equation $K = AGA^{-1} (\text{mod } p)$. It is difficult to solve if the modulus is a large prime number. After simplification this equation becomes

$$G_c X = K$$

Where the elements of X are the combination of elements of A and A^{-1} and G_c is the coefficient matrix of G and its determinant is zero. For example A is an 2 x 2 prime circulant matrix and G is 2 x 2 global non-singular matrix, then the above equation becomes

$$\begin{bmatrix} g_1 & g_2 & g_3 & g_4 \\ g_2 & g_1 & g_4 & g_3 \\ g_3 & g_4 & g_1 & g_2 \\ g_4 & g_3 & g_2 & g_1 \end{bmatrix} \begin{bmatrix} ac \\ ad \\ bc \\ bd \end{bmatrix} = \begin{bmatrix} k_{11} \\ k_{12} \\ k_{21} \\ k_{22} \end{bmatrix}$$

This produces less than n^2 equations since the determinant of coefficient matrix G_c is zero with $2n^2$ unknowns. In which n^2 unknowns are combination of elements of A and A^{-1}, remaining n^2 unknowns are elements of the matrix K. This makes the algorithm more secure. It shows that the computational cost of the proposed algorithm is more than the Hill cipher but it thwarts the security vulnerabilities.

5. Conclusion

The proposed cryptosystem is based on circulant matrices. The structure of the cryptosystem is similar to Affine Hill cipher and each block of plaintext is encrypted using a different random numbers. Circulant matrices reduce the storage requirement from n^2 to n. The proposed cryptosystem provides security in two levels, so it makes difficultly for adversary to break the system. It thwarts known-plaintext attack, and it also thwarts chosen- ciphertext attack and chosen-plaintext attack since the knowledge of such attack is essential. It thwarts ciphertext-only attack, if the modulus is a large prime number.

References

1. Adi Narayana Reddy K, Vishnuvardhan B, Madhuviswanath V, Krishna AVN. “A Modofied Hill Cipher based on Circulant Matrices”, Procedia Technology (ELSEVIER) Second International Conference on Computer, Communication, Control and Information Technology. Volume 4 2012 page 114-118

2. Ahmed, Y. Mahmoud, Alexander, G. Chefranov. "Hill Cipher Modification Based on Eigenvalues HCM-EE," Proc. Of the Second International Conference on Security of Information and Networks (SIN2009), October 6-10, 2009, North Cyprus, Turkey.
3. Bauer, C. and Millward, K. 2007. Cracking Matrix Encryption Row by Row, Cryptologia, 31(1), 76-83
4. Chefranov A. G, "Secure Hill Cipher Modification SHC-M" Proc. Of the First International Conference on Security of Information and Network (SIN2007) 7-10 May 2007, Gazimagusa (TRNC) North Cyprus, Elci. A., Ors, B., and Preneel, B (Eds) Trafford Publishing, Canada, 2008: pp 34-37, 2007
5. Chefranov, A. G. 2008. Secure Hill Cipher Modification SHCM, Proc. of the First International Conference on Security of Information and Networks (SIN2007) 7-10 May 2007, Gazimagusa (TRNC) North Cyprus, Elçi, A., Ors, B., and Preneel, B. (Eds.) Trafford Publishing, Canada. 34-37
6. D. Kalman and J. E. White, "Polynomial Equations and Circulant Matrices", Amer. Math. Monthly 108 (2001), 821-840.
7. Daniele Macciancio, Oded Regev, "Lattice-based Cryptography" July 22 2008.
8. Dan Kalman and James E. White, "Polynomial Equations and Circulant Matrices," The Mathematical Association of America [Monthly 108 November 2001.
9. Galvin, W. P. 1984. Matrices with Custom-Built Eigenspaces, this MONTHLY, 91, 308-309.
10. Gupta, J. Singh, R. Chaudhary, Cryptanalysis of an Extension of the Hill Cipher," Cryptologia, Vol.31, pp.246-253, 2007
11. Hill L S, "Cryptography in an Algebraic Alphabet", American Mathematical Monthly 1929; 36: 306-312.
12. Hill L S, "Concerning certain Linear Transformation Apparatus of Cryptography," American Mathematical Monthly 1931; 38: 135-154
13. Ismail IA. Amin M, Diab H. "How to Repair the Hill Cipher", Journal of Zhejiang University – Science A 2006; 7: 2022 - 2030
14. Koblitz N. A course in Number theory and Cryptography. Springer-Verlag: New York, 1987; 64-74
15. L. Law, A. Menezes, M. Qu, J. Solinas, and S. Vanstone, "An efficient Protocol for Authenticated Key Agreement," Journal of Designs, Codes and Cryptography, Vol.28, pp.119-134, 2003
16. Lin C. H, Lee C. Y, and Lee C. Y, "Comments on Saeednia's improved scheme for Hill cipher," Journal of the Chinese institute of engineers, Vol.27, No. 5, pp. 743-746, 2004.
17. Li C, Zhang D and Chen G, "Cryptanalysis of an image encryption based on the Hill cipher," Journal of Zhejiang University – Science A, Vol.9, No.8, pp. 1118-1123, 2008.
18. Mohsen Toorani, Abolfazl Falahati, "A Secure Variant of the Hill Cipher," Proceedings of the 14th IEEE Symposium on Computers and Communications (ISCC'09), pp.313-316, July 2009
19. Mohsen Toorani, Abolfazl Falahati , "A Secure Cryptosystem based on Affine Transformation," Journal of Security and Communication Networks, Vol. 4, No. 2, pp. 207-215, Feb. 2011
20. Overbey, J. Traves, W. and Wojdylo, J. 2005. On the Key Space of the Hill Cipher, Cryptologia, 29(1), 59-72
21. Pohlig SC, Hellman ME. An Improved Algorithm for Computing Logarithms in GF (p) and Its Cryptographic Significance. *IEEE Transactions on Information Theory* 1978; **24**: 106–111
22. Saeednia's S, "How to Make the Hill cipher Secure," Cryptologia Journal 2000; 24: 353-360.
23. Schneier , B., Kelsey, J., Whiting, D., Wagner, D., Hall, C., Ferguson, N. "The Twofish Algorithm: A 128-bit Block Cipher". John Wiley and Sons (1999)

24. Y. Rangel-Romeror, R. Vega-Garcia, A. Menchaca-Mendez, D. Acoltzi-Cervantes, L. Martinez-Ramos, M. Mecate-Zambrano, F. Montalvo-Lezama, J. Barron-Vidales, N. Cortez-Duarte, F. Rodriguez-Henriquez. Comments on "How to repair Hill cipher", Journal of Zhejiang University SCIENCE A, 2008 9(2):211-214

25. Y.S. Yeh, T.C. Wu, C.C. Chang, and W.C. Yang, "A New Cryptosystem Using Matrix Transformation," Proceedings of the 25th IEEE International Carnahan Conference on Security Technology, pp.131-138, Oct. 1991.

26. Yu, X. Y. Zhang, J. et al. 2006. A New Measurement Method of Image Encryption. Journal of Physics, 48(1), 408-411

Appendix:

The Computational Cost of Different Algorithms for encryption/decryption of each block of data

Different Algorithms	Operation	T_{Add}	T_{Mul}	T_{Inv}	T_{Hash}
Hill Cipher	Encryption	n^2	n^2-n	-	-
	Decryption	n^2	n^2-n	1	-
Affine Hill Cipher	Encryption	n^2	n^2	-	-
	Decryption	n^2	n^2	1	-
Saeednia's Scheme	Encryption	n^3	n^3	1	-
	Decryption	n^3	n^3	1	-
Lin C. H Scheme	Encryption	n^2+n+3	n^2+4	-	n+1
	Decryption	n^2+n+3	n^2+4	1	n+1
Tarooni's Scheme	Encryption	n^2+2n	n^2+n+1	-	1
	Decryption	n^2+2n	n^2+n+1	1	1
The Proposed Scheme	Encryption	n^2+n	n^2	-	-
	Decryption	n^2+n	n^2	-	-