ISSN No. 0976-5697

Volume 2, No. 4, July-August 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Survey of Traditional or Character Oriented Symmetric Key Cryptography

Sukalyan Som*
Department of Computer Science
Barrackpore Rastraguru Surendranath College
Kolkata, India
sukalyan.s@gmail.com

Saikat Ghosh
Department of Computer Science
Barrackpore Rastraguru Surendranath College
Kolkata, India
ghoshsaikat6@gmail.com

Abstract: Cryptography is the art and science of achieving security by converting sensitive information to un-interpretable form such that it cannot be interpreted by anyone except the intended recipient. An innumerable set of cryptographic schemes persist in which each of it has its own affirmative and feeble characteristics. This survey describes the various traditional symmetric key cryptosystems for convinced selection of both basic knowledge and finding future scopes. In the introduction basic definitions of cryptography, cryptanalysis and cryptology is stated. In the next section the fundamental terminologies are stated. Third section differentiates between symmetric and asymmetric key cryptography. Hence, explanations on traditional symmetric key cryptography are presented along with the weaknesses of the schemes. To conclude factors for selecting appropriate cryptographic schemes arte explained.

Keywords: Asymmetric key cryptography; Monoalphabetic Cipher; Polyalphabetic Cipher; Symmetric key cryptography; Transposition Cipher.

I. INTRODUCTION

Cryptography defined as a set of techniques and study of mathematics related to aspects of information security such as confidentiality, authenticity, integrity, non-repudiation [1]. While cryptography is the science of securing data, cryptanalysis is the science of analyzing and breaking secure communication. The person who uses cryptanalysis is called cryptanalyst or attacker [2]. Cryptology is a combination of both cryptography and cryptanalysis. Cryptanalytic attacks can be cipher text only, known plaintext, chosen plaintext, chosen cipher text, adaptive chosen plaintext, brute force attack, key guessing attack etc [3].

II. BASIC TERMINOLOGY OR FUNDAMENTAL LITERATURE

Plaintext or Cleartext signifies a message that can be understood by the sender, the recipient and also by anyone else who gets an access to that message.

Cipher refers to the algorithm(s) for transforming an intelligible message to unintelligible form.

When a plain text message is codified using any suitable scheme, the resulting message is known as Ciphertext.

A key is a number (or a set of numbers) that the cipher, as an algorithm, operates on.

The method of disguising plaintext in such a way as to hide its substance is called encryption. Decryption is the process of reverting ciphertext to its original plaintext.

As depicted in Fig 1, to encrypt a message, an encryption cipher, an encryption key and plaintext is needed which creates the ciphertext. To decrypt a message a decryption cipher, a decryption key and the cipher text is needed which reveals the original plaintext.

In cryptography, it is customary to use three characters Alice, Bob and Eve in an information exchange scenario. Alice is the person who needs to send secure data. Bob is the recipient of the data. Eve is the person who somehow disturbs the communication between Alice and Bob by intercepting messages to uncover the data or by sending her own disguised messages.

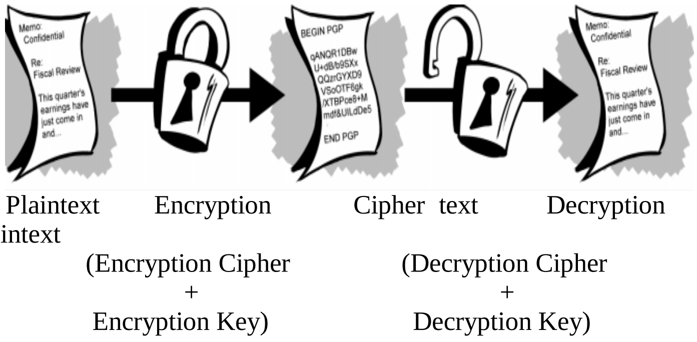

Figure 1. Encryption and decryption

III. SYMMETRIC AND ASYMMETRIC CRYPTOSYSTEMS SELECTING A TEMPLATE

Depending on the key(s) used for an encryption and decryption a cryptosystem can be classified as – Symmetric Cryptosystem and Asymmetric Cryptosystem.

As shown in Figure 2, Symmetric Cryptosystems use the same key, known as Secret key, to both encrypt and decrypt the message. It has a problem to transport the secret key from the sender to the receiver and in tamperproof fashion.

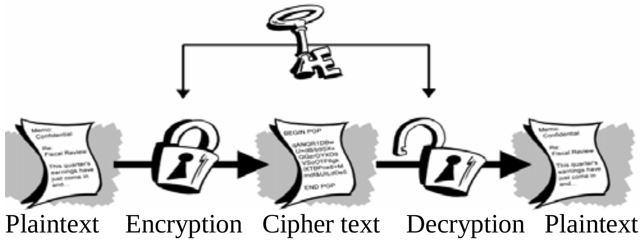

Figure 2. Symmetric Cryptosystem or Secret key Cryptosystem or Private key Cryptosystem

Conventional encryption has benefits. It is very fast. It is especially useful for encryption of data that is not going anywhere. However, conventional encryption alone as a means for transmitting secure data can be quite expensive simply due to the difficulty of secure key distribution. For a sender and recipient to communicate securely using conventional encryption, they must agree upon a key and keep it secret between themselves.

As shown in Figure 3, Asymmetric Cryptosystems use one key, known as Public key, to encrypt a message and a different key, the Private Key, to decrypt it. These are also known as Public Key Cryptosystems.

The primary benefit of public key cryptography is that it allows people who have no pre-existing security arrangement to exchange messages securely. The need for sender and receiver to share secret keys via some secure channel is eliminated; all communications involve only public keys, and no private key is ever transmitted or shared.

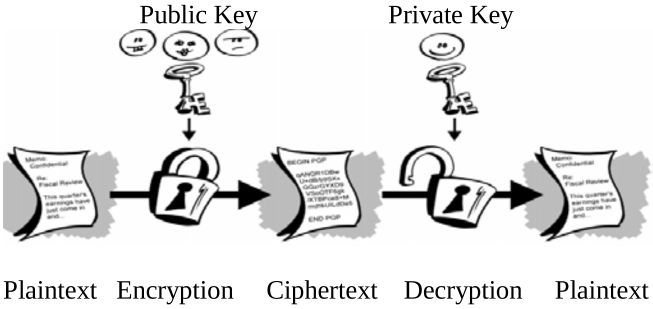

Figure 3. Asymmetric key or Public key Cryptosystem

IV. TRADITIONAL SYMMETRIC KEY CRYPTOGRAPHY

Classical Cryptography is based on information theory appeared in 1949 with the publication of “Communication Theory of Secrecy of Systems” by C. Shannon. In Classical Cryptography both plaintext and key length were same to support secrecy through encryption [4].

Symmetric Cryptosystems can be categorized as Traditional or Character Oriented and Modern or Bit Oriented. This survey concerns with Traditional Symmetric key Cryptography. Although these are now obsolete, the goal is to provide descriptions and analysis from which modern ciphers has evolved. There are two primary ways in which a plaintext can be codified to get the ciphertext using traditional key cryptography – Substitution Ciphers and Transposition Ciphers. When these two approaches are clubbed together, we call them Product Cipher.

A. Substitution Cipher

A Substitution Cipher substitutes one symbol with another. If the symbols in the plaintext are alphabetic characters, one character is replaced with another and if the symbols are digits, one digit is replaced with another. Substitution Ciphers can be categorized as either Monoalphabetic or Polyalphabetic ciphers.

i. Monoalphabetic Substitution Cipher

In Monoalphabetic cipher a symbol in the plaintext is always substituted with another symbol to form the Ciphertext i.e. the relationship between a symbol in plaintext and in Ciphertext is always one-to-one.

a) Shift Cipher Or Additive Cipher

Shift cipher or additive cipher is considered to be the simplest among all Monoalphabetic ciphers. Assuming that both the plaintext and Ciphertext symbols are alphabets, each character is assigned Z_{26}. The result of a mod n is always a non-negative integer less than n and thus the modulo operation creates a set, which in modular arithmetic is referred to as the set of least residues modulo n or Zn. Here, the secret key between Bob and Alice is an integer in Z_{26}. As depicted in Fig 4, the encryption cipher adds the key to the plaintext; the decryption cipher subtracts the key from the Ciphertext. All the operations are done in Z_{26}.

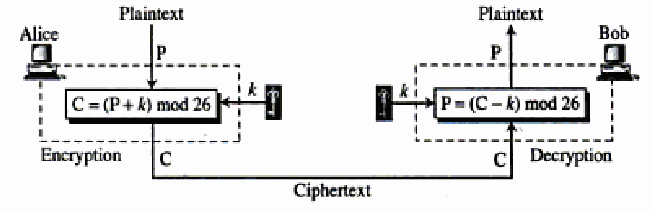

Figure 4. Additive Cipher or Shift Cipher

The shift cipher is interpreted as during encryption characters are shifted down and during decryption characters are shifted up.

A Caesar cipher is an example of shift cipher, is one of the simplest and most widely known encryption techniques. Here, k is 3, A would be replaced by D, B would become E, and so on. It was named after Julius Caesar, who, according to Suetonius, used it with a shift of three to protect messages of military significance [5].

b) Multiplicative Cipher [6]

In Multiplicative Cipher the encryption cipher specifies multiplication of the secret key with the plaintext and the decryption cipher specifies division of Ciphertext by the key as shown in Fig 5.

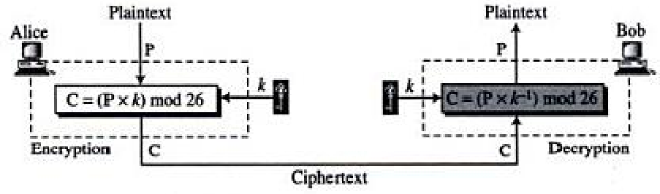

Figure 5. Multiplicative Cipher

In multiplicative cipher, the plaintext and the Ciphertext are integers in Z_{26}; the key is an integer in Z_{26}*.

c) Affine Cipher [7]

Affine cipher can be looked upon as a combination of both Additive cipher and Multiplicative cipher with a pair of keys.

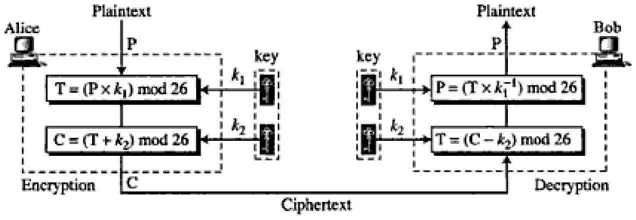

Figure 6. Affine Cipher

As depicted in Fig 6, the first key, k_1 is used with the multiplicative cipher to get a temporary result T and the second; k_2 is then used with the Additive cipher over the temporary result T to get the ciphertext. While decrypting the ciphertext first subtraction is applied with the key k_2 to get the temporary result and hence division is applied with key k_1 to get the plaintext.

d) Homophonic Cipher [8]

The Homophonic Substitution Cipher involves replacing each letter with a variety of substitutes, the number of potential substitutes being proportional to the frequency of the letter. For example, the letter 'a' accounts for roughly 8% of all letters in English, so we assign 8 symbols to represent it. Each time an 'a' appears in the plaintext it is replaced by one of the 8 symbols chosen at random, and so by the end of the encipherment each symbol constitutes roughly 1% of the ciphertext. The letter 'b' accounts for 2% of all letters and so we assign 2 symbols to represent it. Each time 'b' appears in the plaintext either of the two symbols can be chosen, so each symbol will also constitute roughly 1% of the ciphertext. This process continues throughout the alphabet, until we get to 'z', which is so rare that is has only one substitute.

ii. Polyalphabetic Substitution Cipher

In Polyalphabetic substitution cipher, primarily proposed by Leon Battista (1568), each occurrence of a symbol may have a different substitute.

To create a Polyalphabetic cipher, effort should be made to make each ciphertext character dependent on both the corresponding character and the position of the plaintext character in the message. This implies that the secrete key should be a stream of subkeys, in which each subkey depends somehow on the position of the plaintext character that uses that subkey for encipherment. In order words, we need to have a key stream $k = (k_1, k_2, k_3 ...)$ in which k_i is used to encipher the i^{th} character in the plaintext to create the i^{th} character in the ciphertext.

a) Autokey Cipher

In this cipher, the key is a stream of subkeys, in which each subkey is used to encrypt the corresponding character in the plaintext. The first subkey is a predetermined value secretly agreed upon by Alice and Bob. The second subkey is the value of the second Plaintext and so on. The name, Autokey, implies that the subkeys are automatically created during the encryption process.

The first Autokey cipher was invented by Girolamo Cardano, and contained a fatal defect. Like many Autokey ciphers it used the plaintext to encrypt itself; however, since there was no additional key, it is no easier for the intended recipient to read the message than anyone else who knows that the cipher is being used [9].

b) Playfair Cipher

Playfair cipher was, invented by Charles Wheatstone in 1854 but named after Lord Playfair, Wheatstone's friend, used by the British army in World War I and by the Australians in World War II.

The Playfair encryption is a two step process. At first a 5X5 matrix is created and populated, which is used to store a keyword or phrase that becomes the key for encryption and decryption. The following rule is considered during the matrix formation.

a. The keyword is entered in the matrix row-wise – left-to-right and then top-to-bottom.
b. Duplicated letters are dropped.
c. The remaining spaces in the matrix are filled with the rest of the English alphabets that were not a part of the keyword. While doing this, I and J are combined in the same cell.

At the concluding step the plain text is converted into digraphs i.e. groups of 2 letters (e.g. "HelloWorld" becomes "HE LL OW OR LD"), and they are mapped out on the matrix. The two letters of the digraph are considered as the opposite corners of a rectangle in the key table. It should noted that the relative position of the corners of this rectangle. Then the following rules are applied, in order, to each pair of letters in the plaintext:

a. If both letters are the same (or only one letter is left), add an "X" after the first letter. Encrypt the new pair and continue. Some variants of Playfair use "Q" instead of "X", but any uncommon monograph will do.
b. If the letters appear on the same row of your table, replace them with the letters to their immediate right respectively (wrapping around to the left side of the row if a letter in the original pair was on the right side of the row).
c. If the letters appear on the same column of your table, replace them with the letters immediately below respectively (wrapping around to the top side of the column if a letter in the original pair was on the bottom side of the column).
d. If the letters are not on the same row or column, replace them with the letters on the same row respectively but at the other pair of corners of the rectangle defined by the original pair. The order is important – the first letter of the encrypted pair is the one that lies on the same row as the first letter of the plaintext pair.

To decrypt the INVERSE (opposite) of the above mentioned last three rules is done keeping the 1st as-is [10].

c) Vigenere Cipher [11]

A Vigenere cipher, developed by Blasé de Vigenere, a 16th century French mathematician, is different from Autokey and Playfair cipher in the sense that the Vigenere key steam does not depend on the plaintext characters; it depends only on the position of the characters in the plaintext.

The key stream is a repetition of an initial secret key stream of length m, where $1 \leq m \leq 26$. Figure 7 describes the cipher where $(k_1, k_2, k_3... k_m)$ is the initial secret key agreed to by Alice to Bob.

$P = P_1P_2P_3......$ $C = C_1C_2C_3......$ $K = [(k_1, k_2, k_3... k_m), (k_1, k_2, k_3... k_m) ...]$

Encryption: $C_i = (P_i + k_i) \bmod 26$

Decryption: $P_i = (C_i – k_i) \bmod 26$

Figure 7. Vigenere Cipher

d) Hill Cipher [12][13]

The Hill cipher, invented by mathematician Lester S. Hill in 1929, constitute the first general method for successfully linear algebra to polygraphic ciphers, and for applying it in a way that is, in fact, practical.

The plaintext is divided into equal-size blocks, which are encrypted one at a time in such a way that each character in the block contributes to the encryption of other characters in the block.

Here the key is a square matrix of size m X m in which m is the size of the block. Consider this matrix to be K = ((kij)), $\forall i = 1(1)n, \, j = 1(1)n$.

Here we will see the encryption scheme for one block. If we call the m characters in the plaintext block P1, P2, P3 ..., Pm; the corresponding characters in the ciphertext block are C1, C2, C3... Cm then Ci = $\sum_{i=1, j=1}^{m} PiKij$

B. *Transposition Cipher*

A transposition cipher does not substitute one symbol for another; rather it changes the location of the symbols. Transposition ciphers may either be keyless or keyed.

i. Keyless Transposition Cipher

Primitive transposition ciphers were keyless. There are two basic methods for permutation of the symbols. In the first one, the plaintext is written into a table column by column and then ciphered row by row. In the second method the reverse procedure is chosen.

a) Rail Fencing Cipher

Simple most keyless transposition cipher, Rail Fencing cipher, involves writing plain text as sequence of diagonals and then reading it row-by-row to produce ciphertext. To decrypt divide the array of characters is divided with 2 and one odd position and one even position is taken. They are padded in a new array. Consider the plaintext to be “Come Home Tomorrow” which after encryption becomes “Cm oeTmrooeHm oorw”.

As described in Fletcher Pratt's Secret and Urgent, it is "written by ruling a sheet of paper in vertical columns, with a letter at the head of each column. A dot is made for each letter of the message in the proper column, reading from top to bottom of the sheet. The letters at the head of the columns are then cut off, the ruling erased and the message of dots sent along to the recipient, who, knowing the width of the columns and the arrangement of the letters at the top, reconstitutes the diagram and reads what it has to say." [14]

b) Simple Columnar Transposition Cipher

In this approach the plaintext is stored row-by-row in table of a pre-defined size. Then the symbols are read column wise in a random column order.

To improve the simple columnar transposition, more complexity can be added, by performing the above operations multiple rounds.

ii. Keyed Transposition Cipher

The keyed transposition cipher divides the plaintext into groups of predetermined size, known as blocks, and then uses a key to permute the characters in each block separately.

iii. Vernam Cipher

The Vernam cipher, named after Gilbert Sandford Vernam, also known as One-time pad, is implemented using a random set of non-repeating characters as the input ciphertext. The most important part in this mechanism is once an input ciphertext for transposition is used, it is never used again for any other message. The length of input ciphertext is equal to the length of the original plaintext. The algorithm for encryption in this technique is as follows:

a. Each plaintext alphabet is treated as a number in an increasing sequence in Z_{26}.
b. The same thing is done for input ciphertext.
c. Each number corresponding to the plaintext alphabet is added to the corresponding input ciphertext.
d. If the sum thus produced is more than 26 then 26 is subtracted from it.
e. Each number of the sum is translated back to the corresponding alphabet and thus the output ciphertext is produced.

V. CONCLUSION

In this survey the effort has been made to present an overview of the different popular traditional symmetric key ciphers. The selection of right cryptographic technique relies on time, memory and security where memory constraint is highly significant in case of small devices as they have low memory where the time is most important for processing speed [15].

A. Time: How much time will be needed for encrypting and decrypting the data and how much time is need to fulfill the pre-requisites before starting an encryption.
B. Memory: How much memory will be need especially in case of small devices like PDAs, smart cards, RFID tags.
C. Security: Selected encryption scheme should meet the confidentiality, integrity (authentication, non-repudiation) and availability.

VI. REFERENCES

[1] Bement A. L. et. al. (2004), Standards for Security Categorization of Federal Information and Information Systems, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg, MD 20899-8900

[2] Ayushi, (2010), A Symmetric Key Cryptographic Algorithm, International Journal of Computer Applications (0975 - 8887) Volume 1. No. 15.

[3] Atul Kahate, (2008)Cryptography and Network Security, Tata McGraw-Hill Education, pg. 47.

[4] Ijaz Ali Shoukat , Kamalrulnizam Abu Bakar and Mohsin Iftikhar, “A Survey about the Latest Trends and Research Issues of Cryptographic Elements”, p 141, International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 2, May 2011, ISSN 1694 0814.

[5] Wobst, Reinhard (2001). Cryptology Unlocked. Wiley. pp. 19. ISBN 978-0-470-06064-3.

[6] http://www.nku.edu/~christensen/section%206%20multiplicative%20ciphers.pdf

[7] http://www.math.cornell.edu/~kozdron/Teaching/Cornell/135Summer06/Handouts/affine.pdf

[8] http://www.simonsingh.net/The_Black_Chamber/homophoniccipher.htm

[9] Kahn, David, The Codebreakers, revised edition, 1996, p. 144.

[10] Smith, Michael Station X: The Codebreakers of Bletchley Park (1998, Channel 4 Books/Macmillan, London) ISBN 0 7522 2189 2.

[11] Singh, Simon (1999). "Chapter 2: Le Chiffre Indéchiffrable". The Code Book. Anchor Books, Random House. p 63–78. ISBN 0-385-49532-3.

[12] Lester S. Hill, Cryptography in an Algebraic Alphabet, The American Mathematical Monthly Vol.36, June-July 1929, pp.306–312.

[13] Lester S. Hill, Concerning Certain Linear Transformation Apparatus of Cryptography, The American Mathematical Monthly Vol.38, 1931, pp.135–154.

[14] Pratt, Fletcher (1939). Secret and Urgent: The story of codes and ciphers. Aegean Park Press. p 143-144. ISBN 0894122614.

[15] Fontaine. C. and Galand. F. (2007), A Survey of Homomorphic Encryption for Nonspecialists, EURASIP Journal on Information Security Volume 2007, Article ID 13801, 10 pages, doi:10.1155/2007/13801, Hindawi Publishing Corporation.