IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 2, Ver. IX (Mar-Apr. 2014), PP 49-53
www.iosrjournals.org

Survey: Recent Modifications in Vigenere Cipher

Ranju S Kartha[#], Varghese Paul[*]
[#] *School of Computer Science Mahatma Gandhi University Kottayam, Kerala*
[*] *Department of Information Technology, CUSAT* Cochin, Kerala

Abstract: *In order to secure the information there are different polyalphabetic substitution ciphers are available. Out of these Vigenere cipher is considered to be most efficient and simplest one. Due to its repeating nature of the key it is also vulnerable to attacks. To overcome this, there are many researches going on to modify the conventional Vigenere Cipher. In this paper we have presented a review of recent modifications in Vigenere cipher and its cryptanalysis.*
Keywords*: Polyalphabetic Cipher, Vigenere Cipher, Kasiski Method, Index of Coincidence IC, Linear Feedback Shift Register LFSR.*

I. INTRODUCTION

In today's world the information security issues are more challenging and complex due to the introduction of internet and distributed systems. Applications ranging from secure commerce, payments to private communications and protecting passwords are some of the important aspects of security. Cryptography plays a crucial role in providing security to data transmitted over the network. Cryptography referred exclusively to encryption, which is the process of transforming an intelligible message into one that is unintelligible and decryption is the reverse that is moving from the unintelligible cipher text back to plaintext. A cipher is a pair of algorithms that create the encryption and the reversing decryption. The detailed operation of a cipher is controlled both by the algorithm and in each instance by a key [8]. The security of encrypted data depends on the strength of cryptographic algorithm and the secrecy of the key. Cryptanalysis is the science of breaking ciphers, and cryptanalysts try to defeat the security of cryptographic systems.

The two basic building blocks of all classical ciphers are substitution and transposition. In substitution technique letters of plaintext are replaced by numbers and symbols. If plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with cipher text bit patterns. Monoalphabetic substitution ciphers replace each letter in the plaintext with another letter to form the ciphertext. One of the main problems with monoalphabetic substitution ciphers is that they are so vulnerable to frequency analysis. A polyalphabetic cipher is any cipher based on substitution, using multiple substitution alphabets. Each occurrence of a character may have different substitute. So it has the advantage of hiding the letter frequency of the underlying language. Instead of there being a one-to-one correspondence, there is a one-to-many relationship between each letter and its substitutes. The best known and simplest of such algorithm is referred to as Vigenere cipher.

II. VIGENERE CIPHER

The late 1500s, Blaise de Vigenere proposed a polyalphabetic system Vigenere cipher that is difficult to decipher. In this method the alphabetic text is encrypted using a series of different Caesar ciphers based on the letters of a keyword. In a Caesar cipher, the cipher-text is formed by shifting each letter in the plaintext, by a fixed position to the right in the alphabet. This shift is performed modulo 26. For example, in a Caesar cipher of shift 3, A would become D, B would become E and so on. The Vigenère cipher consists of several simple substitution ciphers in sequence with different shift values. For encryption and decryption, a table of alphabets can be used, termed a tabula recta, Vigenere square, or Vigenere table.

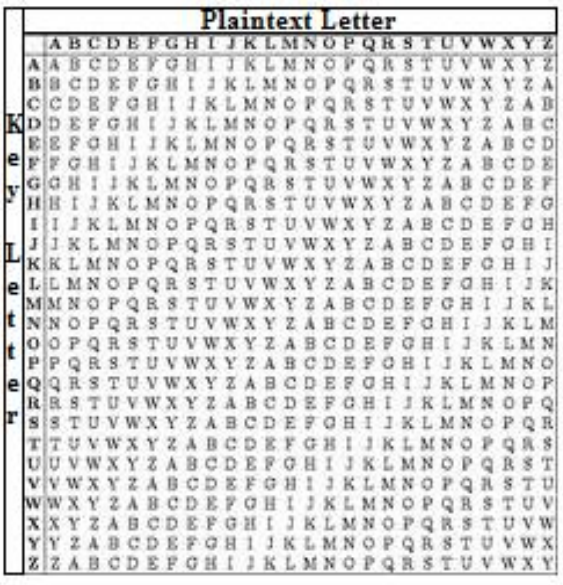

Fig. 1 Vigenere Table

This table comprises of alphabets written out 26 times in different rows, each alphabet is cyclically left shifted compared to the previous alphabet, resulting 26 possible combinations of Caesar ciphers [8].

For example, suppose the plaintext to be encrypted is CRYTOGRAPHY and keyword is CIPHER, for encryption repeats the keyword until it matches the length of the plaintext. Encryption is performed by going to the row in the table corresponding to the key, and finds the column heading the corresponding letter of the plaintext character; the letter at the intersection of corresponding row – column of the Vigenere Square produce the ciphertext character. The rest of the plaintext is encrypted in the same way.

Plaintext : CRYPTOGRAPHY

Keyword : CIPHERCIPHER

Ciphertext : EZNWXFIZPWLP

And decryption is performed by going to the row corresponding to the key; find the position of the ciphertext letter in this row, and then the corresponding column's label will be the plaintext. The Vigenere cipher was called unbreakable because the number of possible solutions grows with the length of the text by a power of 26.

III. CRYPTANALYSIS OF VIGENERE CIPHER

Cryptanalysis of classical ciphers is made possible because of the redundancy in the linguistic structure of natural languages [9]. The most frequent letters in the ciphertext are simply shifted versions of the most frequent letters in the plaintext. So the cryptanalyst can easily decrypt a ciphertext by performing frequency analysis on the letters in the ciphertext. The Vigenere cipher masks the frequency with which a character appears in a language, which makes the use of frequency analysis more difficult.

The primary weakness of the Vigenère cipher is the repeating nature of its key. If a cryptanalyst correctly find the key's length, then the cipher text can be easily broken [9]. Cryptanalysis of the Vigenere cipher has 2 main steps, first one is to identify the period of the cipher (the length of the key), then find the specific key. Once the length of the key is known, it is easy to derive the letters of the key. We can divide the ciphertext into that many simple substitution cryptograms. Then using frequency analysis we can solve the resulting shift ciphers. There are two methods to find the key length. They are the Kasiski method - to find the repeated text sequence of at least three in the ciphertext and the Index of Coincidence - to predict the number of alphabets used for substitution. Both methods cannot break Vigenere ciphers if the ciphertext is short or does not include repetitions.

The Kasiski method is based on the following observation: repeated portions of plaintext encrypted with the same portion of the keyword result in identical ciphertext segments [1].This method follows the rule: if a message is encrypted with *m* alphabets (key length is *m* for Vigenere cipher), and if a particular word or letters group appears *d* times in the plaintext, then it should be encrypted approximately *d/m* times from the same alphabet [11]. The distance between the repeated patterns in cipher text should be a multiple of the key length or

say the number of alphabets used. So in this algorithm first we can identity repeated patterns of three or more letters, then for each pattern, calculate the distance between the positions of starting point of successive instances of the pattern. Next determine the greatest common divisor of all distances, and then the key length should be one factor of the GCD.

($d \equiv 0 \pmod{m}$, m is the key length)

After the length of the key has been found, we can divide the ciphertext into m different pieces and applies the method used to cryptanalyze the additive cipher, including frequency attack. Each ciphertext piece can be decrypted and put together to create the whole plaintext. In other words, the whole cipher text does not preserve the single-letter frequency of the plaintext, but each piece does [2].

William Friedman developed statistical methods for determining whether a cipher is monoalphabetic or polyalphabetic and for determining the length of the keyword if the cipher is polyalphabetic. Friedman's test for determining whether a cipher is monoalphabetic or polyalphabetic is based Index of Coincidence. The Index of Coincidence measures the probability that two randomly selected letters of the string are identical. Mathematically we can compute the Index of Coincidence IC for a given letter-frequency distribution as

$$IC = \frac{\sum_{i=1}^{c} n_i(n_i - 1)}{N(N-1)/c},$$

where N is the length of the text and n_1 through n_c are the frequencies (as integers) of the c letters of the alphabet (c = 26 for monobasic English). This formula gives the ratio of the total number of coincidences observed to the total number of coincidences that one would expect from the null model.

For an arbitrary string of English text Index of Coincidence is 0.065 and for random string it is 0.038. If the ciphertext were generated by a monoalphabetic cipher, the frequencies of letters for the ciphertext alphabet should be nearly the same as for English – but in a different order. If the cipher were generated by a polyalphabetic cipher, the frequencies of the letters would become more nearly uniform – more nearly the same for each letter. If the Index of Coincidence of ciphertext is closer to 0.065, the more likely we have a monoalphabetic cipher. If it closer to 0.038, the more likely that we have a polyalphabetic cipher.

Using this Index of Coincidence we can estimate the keyword length for the Vigenere cipher. The following formula gives an estimate of the keyword length:

$$m \approx \frac{0.0265n}{0.065 - IC + n(IC - 0.0385)},$$

where n is the number of letters in the ciphertext message, IC is the Index of Coincidence and m is the keyword length. For verifying the keyword length m, divide the ciphertext into m substrings and compute the Index of Coincidence for each substring. If all IC values of the substrings are around 0.065, then m is the correct keyword length. Otherwise (IC value ≈0.038) m is not the correct keyword length.

IV. MODIFICATIONS IN VIGENERE CIPHER

The primary weakness of the Vigenère cipher is the repeating nature of its key, so if the period of key is greater than plaintext we can increase the security of Vigenere cipher. To increase the period of key the length of Vigenere key needs to be enlarged. If we choose a long key, it is harder to remember and also its transmission through a secure channel is expensive.

A novel approach is presented by combing the LFSR key with Vigenere cipher key [10]. Paper introduces the cipher system which exploits the advantages of Vigenere cipher with LFSR based stream ciphers and mitigates the weaknesses of these individual ciphers. This system generates large key from short keyword using LFSR concept. Here plaintext is encrypted with large and pseudorandom letter generated key which flatten the letter frequencies of cipher text. So it will more difficult to identify the length of the key. Since period of this Cipher is large enough and also the key stream is pseudorandom letters leading to decrease the effectiveness of Kasiski and Index of Coincidence (IC) attacks.

Another method for modifying the Vigenere Cipher algorithm, by automatically changing the cipher key after each encryption step [3]. In this method the key varies as it is used in the encryption process and the successive keys that will be dependent on the initial key value. The second step key will be as a result of a function that operated on the first step (initial key) and so forth. As the key varies after each successive encryption that eliminating the repeating nature of the key as in conventional Vigenere Cipher. The ciphertext will have different encryption key pattern so the cryptosystem will be more difficult against the frequency attack.

Currently there are several good cryptographic algorithms like AES and DES. And these algorithms utilize the two factors needed for a cryptographically secure algorithm, confusion and diffusion. Confusion means making the correlation between the cipher text and plain text as complex as possible. Diffusion is used to mask the statistical properties of the data by spreading it throughout the cipher text. Stream ciphers, like Vigenère and Caesar require only one round. Vigenère cipher is very good at the confusion aspect of the cryptography, but they lack the benefit of diffusion. By adding a few bits of random padding to each byte before the message are encrypted using Vigenère, one can add the benefit of diffusion [3]. The amount of random bits is determined by a one way function, $F(x)$, consisting of a prime p, a generator g less than p and a positive constant less than eight c. The message length should be less than p to prevent the possible detection of cycles.

$$F(x) = (g^{x} + c) \bmod p,$$

where x represents the n^{th} character of a message. Each byte has been padded and it is concatenated to the last bit of the previously padded bytes. The key needed to perform encryption and decryption using this methodology is as follows:

Key: (p, g, c, Vigenère key).

In this crypt-system the natural repetitiveness of a language is obscured and diffused by the random padded bits. The characteristics of the enciphered padded text will be different from the enciphered text generated using conventional Vigenère cipher. So the Kasiski method for finding the key length is no longer effective when the message is padded with random bits. The main drawback of this method is that the size of the encrypted message will be increased by around 56%. For areas with low bandwidth or limited storage capacity this cipher cannot be used. However for most communication channels where encryption is required, a moderate increase in message size will not have a significant impact. The other major drawback is that a good random number generator is required to create an effective cipher. If one bit is lost; the remainder of the message will be useless [4].

For modification of Vigenère cipher, numbers, punctuations, mathematical symbols may be used for key in place of characters to make it more difficult for brute force attack. If random numbers are used for key that increases the difficulty to decipher the message [5]. In conventional Vigenere Cipher the plaintext is considered as a sequence of alphabets without any space between them. It may create a problem for receiver to read the message by inserting spaces between words. The converted sentence may or may not form a meaningful one. Even though it is meaning full the sentence, it may not be the exact plain text, because, the receiver needs to guess the exact place to insert space in decrypted plaintext. Hence the user will be under pressure to choose the place for inserting space. In order to conquer this difficulty an enhanced polyalphabetic cipher with extended vigenere table was developed [6]. In extended Vigenere cipher a new symbol is added into the row and column of the Vigenere Tableau. The new symbol can be used to represent or to locate the blank space in the plaintext. Hence the user can easily encrypt or decrypt the message or plaintext without any ambiguity.

TheVigenere table was enhanced by Dennie Van Tassel in his paper and was constructed by using 36X36 matrix comprising of 26 alphabets and numbers from 0 to 9. It was named as modern Vigenere cipher. Recently, Dr. Udaya has enhanced the Modern Vigenere table [7] by constructing 68X68 matrix, consisting of alphabets (1 to 26), numbers (0 to 9) and all the symbols present on the keyboard (32). It helps to encrypt and decrypt the combination of all kinds of text and the symbols on key board [7]. To fulfill the need in the field of secrecy modified the Vigenere table into a 256X256 matrix and named it as a comprehensive Vigenere table with 128 ASCII and 128 Extended ASCII characters [8]. If we assume that the length of the key is 256 characters and without repetition of any characters, then there will be 256! Hence brute force attack is not possible.

V. CONCLUSION

Encryption is the foolproof protection against deliberate abuse of information. Cryptography with its benefits can protect communications and stored information from unauthorized access. Lots of researchers are going on for modifying the ciphers which are vulnerable for various attacks. As a result cryptography grows without any boundary, and which in turn cause an increase in activities of the cryptanalyst to find new loopholes. Hence cryptography offers immense potential for research activities. This paper is a review of recent modifications in Vigenere cipher and its cryptanalysis.

ACKNOWLEDGMENT

We like to thank the Director of School of Computer Science, Mahatma Gandhi University Kottayam for providing all the facilities to complete the task.

REFERENCES

[1] Alfred J. Menezes, Paul C. van Oorschot, Scott A., "*Handbook of Applied Cryptography*", CRC Press, 1996.
[2] Behrouz A Forouzan, "*Cryptography and Network Security*", Tata McGraw-Hill Publishing Company Limited, 2007.
[3] Quist-Aphetsi Kester, "*A cryptosystem based on Vigenère cipher with varying key*", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 10, December 2012.
[4] Phillip I Wilson and Mario Garcia, "*A Modified Version of the Vigenère Algorithm*", IJCSNS International Journal of Computer Science and Network Security, Volume 6 No.3B, March 2006.
[5] C. R. S. Bhardwaj, "*Modification of Vigenère Cipher by Random Numbers, Punctuations & Mathematical Symbols*", IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661 Volume 4, Issue 2, Sep.-Oct. 2012.
[6] Ravindra Babu Kallam, Dr.S. Udaya Kumar, Md Abdul Rasool, Dr. A. Vinaya Babu and Puskur Pavan, "*An Enhanced Polyalphabetic Cipher using Extended Vigenere Table*", International Journal of Advanced Research in Computer Science, Mar-Apr 2011.
[7] Kallam Ravindra Babu, Dr. S. Udaya Kumar, Dr. A. Vinaya Babu and Dr. M. Thirupathi Reddy, "*An Enhanced Cryptographic Substitution Method for Information Security*", International Journal of Mathematical Archive- 2(10), 2011.
[8] Prof. Ravindra Babu Kallam, Dr.S.Udaya Kumar, Dr. A.Vinaya babu and V.Shravan kumar, "*A Contemporary Polyalphabetic Cipher using Comprehensive Vigenere*", World of Computer Science and Information Technology Journal (WCSIT) ISSN: 2221-0741 Vol. 1, No. 4,167-171, 2011.
[9] Eskicioglu and Litwin, "*Cryptography*", IEEE Potentials, February/March.
[10] Abdul Razzaq and Yasir Mahmood, Farooq Ahmed, "*Strong Key Machanism Generated by LFSR based Vigenère Cipher*", The 13th International Arab Conference on Information Technology, ACIT''2012, Dec.10-13.
[11] http://www.scribd.com/doc/36928010/Classical-Cryptography