International Journal of Computer Science and Software Engineering (IJCSSE), Volume 6, Issue 2, February 2017
ISSN (Online): 2409-4285 www.IJCSSE.org Page: 41-46

Analysis of Polyalphabetic Transposition Cipher Techniques used for Encryption and Decryption

Shaikh Abdul Hannan[1] and Ali Mir Arif Mir Asif[2]

[1] Assistant Professor, Department of CS & IT, Albaha University, Albaha, Saudi Arabia.

[2] Assistant Professor, I.M.S.I.T., Aurangabad, India.

[1]abdulhannan05@gmail.com, [2]mirarifali05@gmail.com

ABSTRACT

Cryptography is considered to be a disciple of science of achieving security by converting sensitive information to an un-interpretable form such that it cannot be interpreted by anyone except the transmitter and intended recipient. An innumerable set of cryptographic schemes persist in which each of it has its own affirmative and feeble characteristics. In this paper we present a perspective on Polyalphabetic Transposition Cipher Techniques which are currently used for encryption and decryption purpose. This paper mainly focuses on practically use of the Polyalphabetic Transposition Cipher Techniques for encryption and decryption purpose. A brief discussion about types of polyalphabetic transposition cipher techniques.

Keywords: *Polyalphabetic, Encryption, Vigenere cipher, Vernam cipher, One—time pad, Decryption.*

1. INTRODUCTION

Cryptography is an art and science. It is a playing major role in information and security division. The main aim of the cryptography is protecting the data from unauthorized users or hackers. "Cryptography is subject contains two parts one is encryption and another one decryption. Encryption is a process converting the plain text to cipher text using some keys. Decryption is a process of converting the cipher text to plain text using the keys" [1]. Encryption is an effective way to achieve the security of data. The word of encryption came in mind of King Julius Caesar because he did not believe on his messenger so he thought to encrypt the data or message by replacing every alphabet of data by 3rd next alphabet [2]. The process of Encryption hides the data in a way that an attacker cannot hack the data. The main purpose of encryption is to hide the data from unauthorized parties from viewing, altering the data [3]. Encryption techniques occur or used by using the shifting techniques, mathematical operations and shifting techniques. The Simple data is known as Plain text and Data after encryption is known as Cipher text. Substitution and transposition techniques are mainly used for it.

In encryption methods, two methods are used for encryption purpose-

a. Substitution techniques-Change the one letter by another using secret key.

b. Transposition techniques-Replace the place of letters of plaintext.

In substitution techniques monoalphabetic and polyalphabetic techniques are used. In monoalphabetic, a single cipher alphabet is used per message. This technique was easy to break because they show the frequency data of plaintext alphabet. So polyalphabetic techniques came into knowledge in which different monoalphabetic substitution as one proceeds through original message [4].

2. TYPES OF POLYALPHABETIC CIPHER

A transposition cipher can easily be recognized by an analysis of character frequencies. Iterating transposition ciphers can greatly increase security, but as with substitution ciphers, almost all such ciphers can be broken. Although many modern cryptosystems incorporate transposition ciphers, the operation on large blocks has the disadvantage of requiring a lot of memory [5].

Polyalphabetic ciphers were invented circa 1467 by the Florentine architect Alberti, who devised a cipher disk with a larger outer and smaller inner wheel, respectively indexed by plaintext and ciphertext characters. Letter alignments defined a simple substitution, modified by rotating the disk after enciphering a few words. The first printed book on cryptography, Polygraphia, written in 1508 by the German monk Trithemius and published in 1518, contains the first tableau – a square table on 24 characters listing all shift substitutions for a fixed ordering of plaintext alphabet characters. Tableau rows were used sequentially to substitute one plaintext

character each for 24 letters, where-after the same tableau or one based on a different alphabet ordering was used. In 1553 Belaso (from Lombardy) suggested using an easily changed key (and key-phrases as memory aids) to define the fixed alphabetic (shift) substitutions in a polyalphabetic substitution. Polyalphabetic ciphers have the advantage over simple substitution ciphers that symbol frequencies are not preserved. However, polyalphabetic ciphers are not significantly more difficult to cryptanalyze, the approach being similar to the simple substitution cipher. In fact, once the block length is determined, the ciphertext letters can be divided into groups (where group consists of those ciphertext letters derived using permutation), and a frequency analysis can be done on each group [6].

A simple substitution cipher involves a single mapping of the plaintext alphabet onto ciphertext characters. A more complex alternative is to use different substitution mappings (called multiple alphabets) on various portions of the plaintext. This results in so-called polyalphabetic substitution. In the simplest case, the different alphabets are used sequentially and then repeated, so the position of each plaintext character in the source string determines which mapping is applied to it. Under different alphabets, the same plaintext character is thus encrypted to different ciphertext characters, precluding simple frequency analysis as per mono-alphabetic substitution.

There are three types of polyalphabetic cipher, these are:-

A. Vigenere cipher
B. Vernam cipher
C. One-time pad cipher

All these three techniques have two features in common- 1) Set of related monoalphabetic substitution rules are used and 2) A key is used for the transformation of plaintext into cipher text [4].

A. Vigenere Cipher

The Vigenere cipher is a method of encrypting alphabetic text by using a series of different Caesar ciphers based on the letters of a keyword. It is a simple form of polyalphabetic substitution [7][8].

The Vigenere cipher has been reinvented many times. The method was originally described by Giovan Battista Bellaso in his 1553 book La cifra del. Sig. Giovan Battista Bellaso; however, the scheme was later misattributed to Blaise de Vigenere in the 19th century, and is now widely known as the "Vigenere cipher".

Though the cipher is easy to understand and implement, for three centuries it resisted all attempts to break it; this earned it the description le chiffre indéchiffrable (French for 'the indecipherable cipher'). Many people have tried to implement encryption schemes that are essentially Vigenere ciphers [9]. Friedrich Kasiski was the first to publish a general method of deciphering a Vigenere cipher.

The first well documented description of a polyalphabetic cipher was formulated by Leon Battista Alberti around 1467 and used a metal cipher disc to switch between cipher alphabets. Alberti's system only switched alphabets after several words, and switches were indicated by writing the letter of the corresponding alphabet in the ciphertext. Later, in 1508, Johannes Trithemius, in his work Poligraphia, invented the tabula recta, a critical component of the Vigenere cipher. The Trithemius cipher, however, only provided a progressive, rigid and predictable system for switching between cipher alphabets.

What is now known as the Vigenere cipher was originally described by Giovan Battista Bellaso in his 1553 book La cifra del. Sig. Giovan Battista Bellaso. He built upon the tabula recta of Trithemius, but added a repeating "countersign" (a key) to switch cipher alphabets every letter. Whereas Alberti and Trithemius used a fixed pattern of substitutions, Bellaso's scheme meant the pattern of substitutions could be easily changed simply by selecting a new key. Keys were typically single words or short phrases, known to both parties in advance, or transmitted "out of band" along with the message. Bellaso's method thus required strong security for only the key. As it is relatively easy to secure a short key phrase, say by a previous private conversation, Bellaso's system was considerably more secure.

Blaise de Vigenere published his description of a similar but stronger autokey cipher before the court of Henry III of France, in 1586. Later, in the 19th century, the invention of Bellaso's cipher was misattributed to Vigenere. David Kahn in his book The Codebreakers lamented the misattribution by saying that history had "ignored this important contribution and instead named a regressive and elementary cipher for him [Vigenere] though he had nothing to do with it" [10].

The Vigenere cipher gained a reputation for being exceptionally strong. Noted author and mathematician Charles Lutwidge Dodgson (Lewis Carroll) called the Vigenere cipher unbreakable in his 1868 piece "The Alphabet Cipher" in a children's magazine. In 1917, Scientific American described the Vigenere cipher as "impossible of translation" [11]. This reputation was not deserved. Charles Babbage is known to have broken a variant of the cipher as early as 1854; however, he didn't publish his work [12]. Kasiski entirely broke the cipher and published the technique in the 19th century. Even

before this, though, some skilled cryptanalysts could occasionally break the cipher in the 16th century [10].

The Vigenere cipher is simple enough to be a field cipher if it is used in conjunction with cipher disks [13]. The Confederate States of America, for example, used a brass cipher disk to implement the Vigenere cipher during the American Civil War. The Confederacy's messages were far from secret and the Union regularly cracked their messages. Throughout the war, the Confederate leadership primarily relied upon three key phrases, "Manchester Bluff", "Complete Victory" and, as the war came to a close, "Come Retribution" [14].

Gilbert Vernam tried to repair the broken cipher (creating the Vernam-Vigenere cipher in 1918), but, no matter what he did, the cipher was still vulnerable to cryptanalysis. Vernam's work, however, eventually led to the one-time pad, a provably unbreakable cipher.

In a Caesar cipher, each letter of the alphabet is shifted along some number of places; for example, in a Caesar cipher of shift 3, A would become D, B would become E, Y would become B and so on. The Vigenere cipher consists of several Caesar ciphers in sequence with different shift values.

To encrypt, a table of alphabets can be used, termed a tabula recta, Vigenere square, or Vigenere table. It consists of the alphabet written out 26 times in different rows, each alphabet shifted cyclically to the left compared to the previous alphabet, corresponding to the 26 possible Caesar ciphers. At different points in the encryption process, the cipher uses a different alphabet from one of the rows. The alphabet used at each point depends on a repeating keyword.

For example, suppose that the plaintext to be encrypted is:

ATTACKATDAWN

The person sending the message chooses a keyword and repeats it until it matches the length of the plaintext, for example, the keyword "LEMON":

LEMONLEMONLE

Each row starts with a key letter. The remainder of the row holds the letters A to Z (in shifted order). Although there are 26 key rows shown, you will only use as many keys (different alphabets) as there are unique letters in the key string, here just 5 keys, {L, E, M, O, N}. For successive letters of the message, we are going to take successive letters of the key string, and encipher each message letter using its corresponding key row. Choose the next letter of the key, go along that row to find the column heading that matches the message character; the letter at the intersection of [key-row, msg-col] is the enciphered letter.

Fig. 1. The Vigenere square or Vigenere table also known as the tabula recta can be used for encryption and descyption

For example, the first letter of the plaintext, A, is paired with L, the first letter of the key. So use row L and column A of the Vigenere square, namely L. Similarly, for the second letter of the plaintext, the second letter of the key is used; the letter at row E and column T is X. The rest of the plaintext is enciphered in a similar fashion:

Plaintext:	ATTACKATDAWN
Key:	LEMONLEMONLE
Ciphertext:	LXFOPVEFRNHR

Decryption is performed by going to the row in the table corresponding to the key, finding the position of the ciphertext letter in this row, and then using the column's label as the plaintext. For example, in row L (from LEMON), the ciphertext L appears in column A, which is the first plaintext letter. Next we go to row E (from LEMON), locate the ciphertext X which is found in column T, thus T is the second plaintext letter.

B. Vernam Cipher

In modern terminology, a Vernam cipher is a symmetrical stream cipher in which the plaintext is combined with a random or pseudorandom stream of data (the "keystream") of the same length, to generate the ciphertext, using the Boolean "exclusive or" (XOR) function. This is symbolized by $\oplus$ [15] and is

represented by the following "truth table", where + represents "true" and − represents "false".

Table 1. Other Names For This Function Are Not Equal (Neq), Modulo 2 Addition (Without Carry) And Modulo 2 Substraction (Without Borrow)

INPUT		OUTPUT
A	***B***	***A⊕B***
−	−	−
−	+	+
+	−	+
+	+	−

The cipher is reciprocal in that the identical keystream is used both to encipher plaintext to ciphertext and to decipher ciphertext to yield the original plaintext:
Plaintext ⊕ Key = Ciphertext
and:
Ciphertext ⊕ Key = Plaintext
If the key stream is truly random and used only once, this is effectively a one-time pad. Substituting pseudorandom data generated by a cryptographically secure pseudo-random number generator is a common and effective construction for a stream cipher. RC4 is an example of a Vernam cipher that is widely used on the Internet.
If, however, the keystream is used for two messages, known to cryptanalysts as a depth, the effect of the keystream can be eliminated, leaving the two plaintexts XORed together. The result is equivalent to a Running key cipher and the two plaintexts may be separated by linguistic crypt analytical techniques.
Ciphertext1⊕ Ciphertext2 = Plaintext1⊕Plaintext2
An operator's mistake of this sort famously allowed the Cryptanalysis of the Lorenz cipher by the British at Bletchley Park during World War II. They diagnosed how the keystream was generated, worked out how to break the cipher, and read vast quantities of high-level messages to and from German high command without ever seeing an actual Lorenz machine [16].

C. One-Time Pad

In cryptography, a one-time pad (OTP) is an encryption technique that cannot be cracked if used correctly. In this technique, a plaintext is paired with random, secret key (or pad). Then, each bit or character of the plaintext is encrypted by combining it with the corresponding bit or character from the pad using modular addition. If the key is truly random, and at least as long as the plaintext, and never reused in whole or in part, and kept completely secret, then the resulting ciphertext will be impossible to decrypt or break [17] [18]. It has also been proven that any cipher with the perfect secrecy property must use keys with effectively the same requirements as OTP keys [19]. However, practical problems have prevented one-time pads from being widely used.

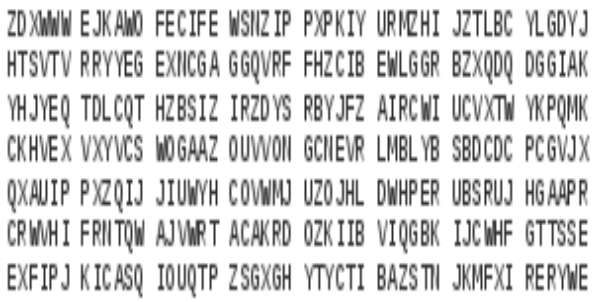

Fig. 2. Expert from a one-time pad

First described by Frank Miller in 1882 [20] [21], the one-time pad was re-invented in 1917 and patented a couple of years later. It is derived from the Vernam cipher, named after Gilbert Vernam, one of its inventors. Vernam's system was a cipher that combined a message with a key read from a punched tape. In its original form, Vernam's system was vulnerable because the key tape was a loop, which was reused whenever the loop made a full cycle. One-time use came later, when Joseph Mauborgne recognized that if the key tape were totally random, then cryptanalysis would be impossible [22].
The "pad" part of the name comes from early implementations where the key material was distributed as a pad of paper, so that the top sheet could be easily torn off and destroyed after use. For ease of concealment, the pad was sometimes reduced to such a small size that a powerful magnifying glass was required to use it. The KGB used pads of such size that they could fit in the palm of one's hand [23], or in a walnut shell [24]. To increase security, one-time pads were sometimes printed onto sheets of highly flammable nitrocellulose, so that they could be quickly burned after use.
There is some ambiguity to the term because some authors use the terms "Vernam cipher" and "one-time pad" synonymously, while others refer to any additive stream cipher as a "Vernam cipher", including those based on a cryptographically secure pseudorandom number generator (CSPRNG) [25].
Frank Miller in 1882 was the first to describe the one-time pad system for securing telegraphy [26] [21].

The next one-time pad system was electrical. In 1917, Gilbert Vernam (of AT&T Corporation) invented and later patented in 1919 (U.S. Patent 1,310,719) a cipher based on teleprinter technology. Each character in a message was electrically combined with a character on a paper tape key. Joseph Mauborgne (then a captain in the U.S. Army and later chief of the Signal Corps) recognized that the character sequence on the key tape could be completely random and that, if so, cryptanalysis would be more difficult. Together they invented the first one-time tape system [25].

The next development was the paper pad system. Diplomats had long used codes and ciphers for confidentiality and to minimize telegraph costs. For the codes, words and phrases were converted to groups of numbers (typically 4 or 5 digits) using a dictionary-like codebook. For added security, secret numbers could be combined with (usually modular addition) each code group before transmission, with the secret numbers being changed periodically (this was called super encryption).

In the early 1920s, three German cryptographers (Werner Kunze, Rudolf Schauffler and Erich Langlotz), who were involved in breaking such systems, realized that they could never be broken if a separate randomly chosen additive number was used for every code group. They had duplicate paper pads printed with lines of random number groups. Each page had a serial number and eight lines. Each line had six 5-digit numbers. A page would be used as a work sheet to encode a message and then destroyed. The serial number of the page would be sent with the encoded message. The recipient would reverse the procedure and then destroy his copy of the page. The German foreign office put this system into operation by 1923 [25].

A separate notion was the use of a one-time pad of letters to encode plaintext directly as in the example below. Leo Marks describes inventing such a system for the British Special Operations Executive during World War II, though he suspected at the time that it was already known in the highly compartmentalized world of cryptography, as for instance at Bletchley Park [27].

The final discovery was by Claude Shannon in the 1940s who recognized and proved the theoretical significance of the one-time pad system. Shannon delivered his results in a classified report in 1945, and published them openly in 1949 [19]. At the same time, Vladimir Kotelnikov had independently proven absolute security of the one-time pad; his results were delivered in 1941 in a report that apparently remains classified [28].

Suppose Alice wishes to send the message "HELLO" to Bob. Assume two pads of paper containing identical random sequences of letters were somehow previously produced and securely issued to both. Alice chooses the appropriate unused page from the pad. The way to do this is normally arranged for in advance, as for instance 'use the 12th sheet on 1 May', or 'use the next available sheet for the next message'. The material on the selected sheet is the key for this message. Each letter from the pad will be combined in a predetermined way with one letter of the message. It is common, but not required, to assign each letter a numerical value, e.g., "A" is 0, "B" is 1, and so on. In this example, the technique is to combine the key and the message using modular addition. The numerical values of corresponding message and key letters are added together, modulo 26. If key material begins with "XMCKL" and the message is "HELLO", then the coding would be done as follows:

		H		E		L		L		O	message
	7	(H)	4	(E)	11	(L)	10	(L)	14	(O)	message
+	23	(X)	12	(M)	2	(C)	11	(K)	11	(L)	key
=	30		16		13		21		25		message + key
=	4	(E)	16	(Q)	13	(N)	21	(V)	25	(Z)	message + key (mod 26)
		E		Q		N		V			→ ciphertext

If a number is larger than 25, then the remainder after subtraction of 26 is taken in modular arithmetic fashion. This simply means that if the computations "go past" Z, the sequence starts again at A.

The ciphertext to be sent to Bob is thus "EQNVZ". Bob uses the matching key page and the same process, but in reverse, to obtain the plaintext. Here the key is subtracted from the ciphertext, again using modular arithmetic:

		E		Q		N		V		Z	ciphertext
	4	(E)	16	(Q)	13	(N)	21	(V)	25	(Z)	ciphertext
-	23	(X)	12	(M)	2	(C)	10	(K)	11	(L)	key
=	-19		4		11		11		14		ciphertext - key
=	7	(H)	4	(E)	11	(L)	11	(L)	14	(O)	ciphertext - key (mod 26)
		H		E		L		L		O	⟶ message

Similar to the above, if a number is negative then 26 is added to make the number positive.

Thus Bob recovers Alice's plaintext, the message "HELLO". Both Alice and Bob destroy the key sheet immediately after use, thus preventing reuse and an attack against the cipher. The KGB often issued its agents one-time pads printed on tiny sheets of "flash paper"—paper chemically converted to nitrocellulose, which burns almost instantly and leaves no ash [29].

The classical one-time pad of espionage used actual pads of minuscule, easily concealed paper, a sharp pencil, and some mental arithmetic. The method can be implemented now as a software program, using data files as input (plaintext), output (ciphertext) and key material (the required random sequence). The XOR operation is often used to combine the plaintext and the key elements, and is especially attractive on computers since it is usually a native machine instruction and is therefore very fast. However, it is difficult to ensure that the key material is actually random, is used only once, never becomes known to the opposition, and is completely destroyed after use. The auxiliary parts of a software one-time pad implementation present real challenges: secure handling/transmission of plaintext, truly random keys, and one-time-only use of the key.

3. CONCLUSION

The paper explains various types of polyalphabetic substitution ciphers. If we have theoretical knowledge of ciphers in detail then we can develop and improve various encryption and decryption algorithms. The main innovation in this paper is that it helps from any cryptographic attack as we have deep knowledge for our secure and safe data through encryption and decryptions algorithms.

REFERENCES

[1] Digital Economy Promotion Agency, under the Administrative Supervision of the Minister of Digital Economy and Society, 2016.
[2] Avialable at: “http://www.sipa.or.th”
[3] D. Vest, M. Long, L. Thomas, and M. E. Palmquist, "Relating Communication Training to Workplace Requirements: The perspective of New Engineers", IEEE Trans on Prof. Commun., Vol. 38, No. 1, 1995, pp. 11-17.
[4] GraduateCareer(GCA), 2016.
[5] Available at: “http://www.graduateopportuities.com”
[6] GURU99, "Agile Model and Methodology: Guide for Developers and Testers”, 2017.
[7] Available at: “http://www.guru99.com”
[8] H. L. Carol, S. Mary, and R. M. Nancy, “The Carnegie Mellon University Master of Software Engineering Specialization Tracks”, in Proceeding of 9th Conference on Software Engineering Education, 1998, p 100-118.
[9] L A. Jessica, H. Richard, T. Corel, C. Steve, L. Jessica, W. Anneliese, . W. Mark, and M. W. Julia, "Communicating Sustainability: Sustainability and Communication in the Engineering, Science, and Technical Communication Classrooms", in IEEE International professional Communication Conference, 2008, pp. 1-7.
[10] P. Phil, "What Employers Want from Students: A Report from OOPSLA",SIGCSE Bulletin, Vol.31, No.2, 1999,pp. 69-70.
[11] R. Susan, and C. Michael, "Communication Learning Outcomes from Software Engineering Professionals: A Basis for Teaching Communication in the Engineering curriculum", in 39th ASGE/IEEE Frontiers in Education Conference W1E-1 Session, 2009, Oct 18-21.
[12] The Joint Task Force on Computing Curricula Association for Computing Machinery(ACM) and IEEE Computer , Society, “Computer Science Curricula 2013 Curriculum Guidelines for Undergraduate Degree Programs in Computer Science”, in A cooperative Project of ACM, IEEE, and IEEE Computer Society, 2013.