ISSN No. 0976-5697

Volume 2, No. 1, Jan-Feb 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

# An Improved Playfair Cipher Cryptographic Substitution Algorithm

Ravindra Babu Kallam*
Professor in Computer Science Engineering
Vivekananda Institute of Technology &Science SET, JNTUH
Kareemnagar, AP, India
rb_kallam@yahoo.com

Dr.A. Vinaya Babu
Director, Admissions
Jawaharlal Nehru Technological University
Hyderabad, AP, India
avb1222@gmail.com

Dr. S. Udaya Kumar
Deputy Director, Professor in CSE
SNIST, JNTUH
Hyderabad, AP, India
uksusarla@regiffmail.com

Sikharam Swetha
Computer Science Engineering
AZCET, JNTUH
Mancherial, AP, India
swethasikharam@yahoo.in

***Abstract:*** The goal of this research is to find the efficient and most widely used cryptographic algorithms form the history, investigating one of its merits and demerits which have not been modified so far. Perception of cryptography, its techniques such as transposition & substitution and Steganography were discussed. Our main focus is on the Playfair Cipher, its advantages and disadvantages. Finally, we have proposed a few methods to enhance the playfair cipher for more secure and efficient cryptography.

***Keywords:*** substitution; Transposition; Mono alphabetic; poly alphabetic; cryptography; decryption; encryption; frequency of letters;

## I. INTRODUCTION

From the beginning of human society, people have been very much concerned with the privacy of their communications. In contemporary societies, the growing use of computer has made the security of digital files of outmost concern against those users with malevolent intentions, especially on the internet. To protect the digital files either in the computer or in the transmission [4], the scientists have proposed and improved many algorithms, which are known as cryptographic and Steganography algorithms.

Cryptography is the science which deals with all the means and methods for converting an intelligible message into an unintelligible or secret form and for reconverting the secret form into the intelligible message by a direct reversal of the steps used in the original process [1]. Steganography conceals the existence of the message. A Simple form of the Steganography, but one that is time consuming to construct, is one in which an arrangement of words or letters within an apparently innocuous text spells out the real message[6].

Even though encryption is very powerful among these two, the cryptanalysts are very intelligent and they were working day and night to break the ciphers. To make a stronger cipher it is recommended that to use: More stronger and complicated encryption algorithms with more number of rounds, Keys with more number of bits (Longer keys), secure transmission of keys [9].

## II. CRYPTOGRAPHIC SYSTEMS

In his paper F. Ayoub [2] mentioned that, Cryptographic systems are used to provide privacy and authentication in computer and communication systems. As shown in Figure- 1, encryption algorithms encipher the Plaintext, or clear messages, into unintelligible cipher text or cryptograms using a key.

A deciphering algorithm is used for decryption or decipherment in order to restore the original information. In general, the enciphering and deciphering keys need not be identical [5].

Eavesdropping is the interception of messages by a third party monitoring a Communication channel. Anyone trying to break (solve) a cipher is called a cryptanalyst.

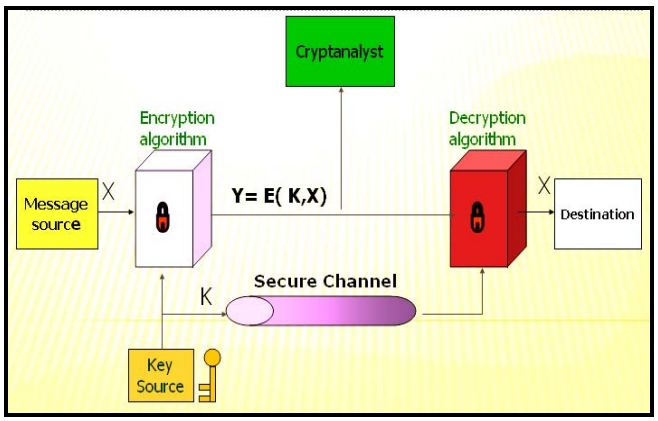


Figure1. General Secrecy System

All Cryptographic algorithms[3][5] are based on two general principals: substitution, in which each element in the plaintext (bit, letter, and group of bits or letters) is mapped into another element and in transposition, the elements of the plaintext have simply been re-arranged in different order; their position with relation to each other have been changed .

### A. *Transposition Cipher*

Transposition ciphers are block ciphers that change the position (or the sequence) of the characters or bits of the input blocks. To encipher, the plaintext is broken into n symbols and a key specifies one of (n!—1) possible permutations.

Deciphering is accomplished by using an inverse permutation which restores the original sequence [10].

Transposition ciphers preserve the frequency distribution of single letters but destroy the diagram and higher-order distributions.

Transposition ciphers are often combined with other ciphers to produce a more secure product cipher.

The simplest such cipher is the rail fence technique, in which the plain text is written down as a sequence of diagonals and then read off as a sequence of rows.

For example, to encipher the message — "meet me after the toga party" with a rail fence of depth 2, we write the following:

m e m a t r h t g p r y
↘ ↘ ↘ ↘ ↘ ↘ ↘ ↘ ↘ ↘ ↘
e t e f e t e o a a t

The encrypted message is:

MEMATRHTGPRYETEFETE OAAT

A more complex scheme is to write the message in a rectangle, row by row and read the message column by column. But permute the order of the columns. The order of the columns then becomes the key to the algorithm. Example:

| | |
|---|---|
| Key: | 4 3 1 2 5 6 7 |
| Plain text: | a t t a c k p |
| | o s t p o n e |
| | d u n t i l t |
| | w o a m x y z |

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

### *B. Substitution ciphers*

In the Substitution Techniques the letters of the plain text are replaced by other letters or by numbers or symbols [2]. Caesar invented a substitution method in which, each letter of the alphabet is replaced with the letter standing three places further down the alphabet.

Example: Note that the alphabet is wrapped around, so that the letter following Z is A. We can define the transformation by listing all possibilities, as follows:

Plain: a b c d e f g h I j k l m n o p q r s t u v w x y z

Cipher: D E F G H I J K L M N O P Q R S T U V W X Y Z A B C

If the plain text is: 'meet me after the party'

The cipher text is: PHHW PH DCWHU WKH SDUWB

It is observed that, if we assume the algorithm is known, and then the brute force analysis is possible; to overcome this problem it is recommended to use the algorithm with large number of key.

With only 26 possible keys Caesar cipher is far from secure. A dramatic increase in the key space can be achieved by allowing an arbitrary substitution. If, instead, the cipher line can be any permutation of the 26 alphabetic characters, then there are 4 x 1026 possible keys and would seem to eliminate brute – force technique for cryptanalysis. Such an approach is referred to as a mono alphabetic substitution cipher.

There is however another type of attack. If the crypt analyst knows the nature of the plain text, then the analyst can exploit the regularities of the language. To see how such a cryptanalysis might proceed, we have a particular example [4].The cipher text is to be solved is.

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXU
DBMETSXAIZVUEPHZHMDZSHZOWSFPAPPDTSVPQU
ZWYMXUZUHSXEPYEPOPDZSZUFPOMBZWPFUPZHM
DJUDTMOHMQ

As a first step, the relative frequency of the letters can be determined and compared to a standard frequency distribution for English, such as is shown in the fig- 2 [2][6]. Proportional fonts only for special purposes, such as distinguishing source code text

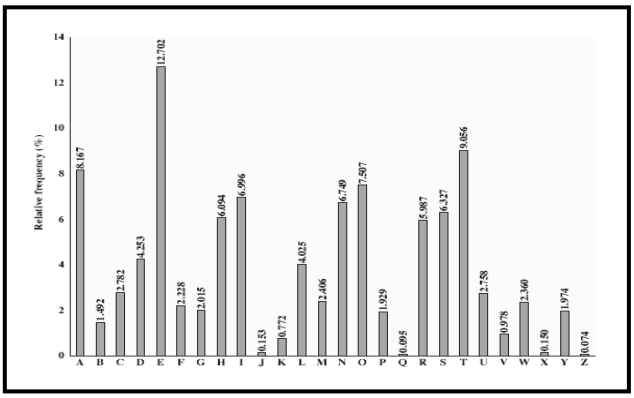


Figure-2: Relative Frequency of Letters in English Text.

It is observe in the above example, the frequency occurrence of cipher letters P is 13.33 is near to e and Z 11.67 is nearer to t, but it is not certain which is which. Mono alphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet. Proceeding with trial and error finally gets the plain text for the above cipher text is: it was disclosed yesterday that several informal but direct contacts have been made with political representatives of the viet cong in moscow .

A counter measure is to provide multiple substitutes, Known as homophonic substitution cipher(HSC), for a single letters. The great mathematician Carl Friedrich Gauss believed that he had devised an unbreakable cipher using homophones. Bale cipher[6] is an example of Homophonic cipher .

One of the well known cryptographic algorithms which was successfully used as the standard field system by British Army in World War I and still enjoyed considerable use by the U. S. Army is the Playfair Cipher[4][8].

## III. EXISTING PLAYFAIR ALGORITHM

It is best known multiple-letter encryption algorithm, based on the use of a 5X5 matrix (Table-I) of letters constructed using keyword. In this case, the keyword is MONARCHY. The matrix is constructed by filling in letters of the keyword from left to right and from top to bottom, and then filling in the remainder of the matrix with the remaining letters in alphabetic order.

Table I. A 5X5 matrix for Playfair cipher

| *M* | *O* | *N* | *A* | *R* |
|---|---|---|---|---|
| C | H | Y | B | D |
| E | F | G | I/J | K |
| L | P | Q | S | T |
| **U** | V | W | X | Z |

The letters I / J count as one letter. Plaintext is encrypted two letters at a time according to the rules:

Repeating plaintext letters that would fall in the same pair are separated with a filler letter, such as x, so that balloon would be enciphered as ba lx lo on.

Plaintext letters that fall in the same row of the matrix are each replaced by the letter to the right, with the first element of the row circularity following the last. For example, ar is encrypted as RM

Plain text letters that fall in the same column are each replaced by the letter beneath, with the top element of the row circularly following in the last .for example, mu is encrypted as CM.

Otherwise, each plaintext letter is replaced by the letter that lies in its own row and the column occupied by the other plaintext letter. thus, hs becomes BP, and ea becomes IM (or JM, as the enciphered wishes).

If the plain text is "ba lx lo on", its corresponding cipher text will be "IBSUPMNA" or "JBSUPMNA".

The playfair cipher is a great advance over simple mono alphabetic ciphers. For this there are only 26 letters there are 26X26=676 digrams, so that identification of individual digrams is more difficult.

Further more, the relative frequencies of individual letters exhibit a much greater range then that of digrams, making frequency analysis much more difficult. For this reasons the playfair cipher was for a long time considered unbreakable [3].

***In the above algorithm we have observed some drawbacks and given solution for the same. The problems we have noticed are:***

The plain text chosen for conversion should have even number of characters otherwise we can not divide the given text into the pairs of characters.

Example: if the plain text is" frequency" it will be divided as "fr eq ue nc y", because it is having " 9 " characters the last character 'y' is single character and the inventor of this algorithm have not considered this problem.

In the 5X5 matrix, because we have only 25 entries, the letters I / J counted as one letter. If we encrypt the plain text which is having the letter I/J and when we decrypt the cipher text at the receiving end, the receiver will be under ambiguity whether to consider I or J in his text, because the meaning can be changed with the change of the letter.

**Example1:** Iamb and Jamb, these are the two words having different meanings in the dictionary.

**Example 2:** If the plain text is " **Jam**", after encryption when it is decrypted it can be "**I am**" instead of *Jam*. It is observed that the meaning of the word will be changing at the decryption end because of the ambiguity of I/J letters and also because of the space character.

In Defense services, each message is very critical and have a lot of risk involves, in such a case the receiver should not have a choice to select a letter in the text, they should obey their superior order, otherwise that may leads to lot of problems.

Day by day the technology got changing a lot with rapid speed, many surveys saying that, the usage of computer is massively increasing. Hence stronger security algorithms need to be invented or the existing algorithms should be updated for providing more security to the information either in the PC or in the transmission [7].

To meet the current requirements, we have Enhenced the Playfair algorithm.

## IV. IMPROVED PLAYFAIR ALGORITHM

The first problem we have noticed in the existing algorithm is, it does not suit for the text having odd number of characters.

Our proposal for this is, append a letter 'X' right to the last letter in the text, so that the number of characters in the text will become even and can be encrypted.

Consider a plaintext having odd number of characters, for example "frequency". It can be divided in to pairs as " fr eq ue nc yx". Using table-2 we can decrypt the plain text as " INGHOKMYBV".

For decryption same procedure can be used as before and after decryption discard the last letter so that we can get actual text.

From the Figure-2, we can also notice that, the frequency of the letter I is 6.996 and J is 0.153. These are widely used letters in normal text [10] and considered as a single letter in the playfair Table-I. It may leads to the confusion at the receiving side whether to use I/J for decryption.

To redeuce the ambiguity at the receiving end; it is better to combine the lees frequency letters as a one letter in the Table-I rather then using I/J as single letter. So, that the less frequency letters appears very rare in the text and hence we can reduce the confusion level while decrypting at the receiving end.

For this we recommend to combine Q (0.095) and Z (0.074) as a single letter in the playfair Table I.

For constructing the Table-2 we arranged the 26 letters in 5X5 matrix by considering the frequency of letters from the Figure-II. Use the keyword monarchy and fill the keyword characters from left to right and top to bottom in the matrix, then fill the least frequency letters Q/Z count as one letter, after that fill the remainder of the matrix with the remaining letters in alphabetic order.

Table II. A 5X5 matrix for Enhanced Playfair cipher

| *M* | *O* | *N* | *A* | *R* |
|---|---|---|---|---|
| C | H | Y | Q/Z | B |
| D | E | F | G | I |
| J | K | L | P | S |
| **T** | U | V | W | X |

Because the character Q and Z appears very rare in the plain text, the receiver faces very less confusion while decrypting.

Consider a plaintext "quit", by using the same procedure it will be encrypted as "HWDX" and can be decrypted as "quit"

with out any ambiguity at the receiving end. Because, if you consider 'Z' in place of 'Q' it do not have any meaning.

## V. RESULTS

With our enhanced play fair cipher algorithm it is observed that, it can be used for the plaintext with either even or odd number of characters for conversion.

It is also noticed that, by combining less frequency character 'Q/Z' instead of 'I/J' the ambiguity at the receiving end will be reduced. Examples for the same were explained in the required place in paper.

## VI. CONCLUSION

Mainly focused on Playfair Cipher Cryptographic algorithm, its merits and de merits were discussed. To solve the de merits in playfair we have proposed and explained few methods with examples. For this, we have reconstructed the Table-I and named it as Enhanced Playfair Cipher Table-II. Finally we named it as: 'An Improved Play Fair Cipher Cryptography Substitution Algorithm'. The present version of the play fair algorithm will consider only text in English for conversion; we can also extend it to numbers and symbols for wide range of use.

## VII. ACKNOWLEDGMENT

The first authors like to thank the Chairmen Sri. Anada Rao, C.O.O Dr. Z. J. Khan and the Principal Prof. Venkatarami reddy, VITS SET for their overwhelming support all along and Vice Chairmen for providing all the facilities to complete the task. We also like to thank IJARCS for allowing us to use its template.

## VIII. REFERENCES

[1] Beker. H and Piper. F, "Communications Survey: a survey of Cryptography", IEE Proc. A, 357-376, 1982.

[2] Dennie Van Tassel , " Cryptographic techniques for computers: Substitution methods", Vol.6.pp.241-249, Pergamon press 1970, Britain.

[3] Denning, D., F. Ayoub , " Cryptographic techniques and network security", IEEE proceedings, Vol 131, 684-694,Dec 1984.

[4] G. J. Simmons, "Symetric and asymmetric encryption", ACM Compute Surveys, 305-330, 11, 1979.

[5] Kallam Ravindra Babu, " A Survey on cryptography and Steganography methods for information security", IJCA (0975-8887), Volume 12- No.2, November 2010. **(Article in a journal).**

[6] Linda S R utledge, " A Survey of Issue in Computer Network Security", Computers & Security 5, 296-308, Elsevier Science Publishers B.V (North- Holland), 1986.

[7] Michael Willet, " Cryptography Old and New", Computers and Security, North-Holland, 0167-4048 / 82 / 0000-0000 / 177-186, 1982.

[8] Simons, " Cryptography" Encyclopedia Britannica , 5th edition, 1993.

[9] Data Security", ACM Compute Surveys, 227-250, 11, 1979.

[10] William Stallings, Cryptography and network security, 5th Impression, 2008.

 