IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-6, Issue-4 E-ISSN: 2347-2693

Chaotic Genetic Enhancements to the Modified PlayFair Algorithm

Archi Seth[1], Siddhartha Sankar Biswas[2*]

Dept. of CSE, School of Engineering Sciences and Technology Jamia Hamdard, New Delhi, India
Dept. of CSE, School of Engineering Sciences and Technology Jamia Hamdard, New Delhi, India

**Corresponding Author: ssbiswas1984@gmail.com*

Available online at: www.ijcseonline.org

Received: 28/Mar/2018, Revised: 06/Apr/2018, Accepted: 21/Apr/2018, Published: 30/Apr/2018

***Abstract*—** The central tenets of modern cryptography are data confidentiality, data integrity, authentication, and non-repudiation. There are a number of algorithms being used for confidential transmission of messages from one end to the other. The PlayFair Cipher is a substitution cipher. The classic PlayFair cipher uses 5 * 5 matrix to encrypt pairs of letters (diagrams).The frequency analysis is harder than simple substitution ciphers as there are 25 * 25 = 625 diagrams. But with increase in computing power, classical ciphers have become easy to break. The classical substitution ciphers can be broken by cipher text-only attacks. This paper presents a modified version of the classical PlayFair Cipher. The improvement is done to the original 5 * 5 cipher to modify it to 10 * 9 matrix. The matrix uses uppercase and lowercase English characters, numbers, punctuation marks and some special characters. The security aspect is enhanced by the use of Chaotic Genetic Algorithm to encrypt the cipher text again using Genetic crossover and mutation operations with a Chaotic Pseudo Random Sequence.

***Keywords*—**Component, Chaos Theory, Cryptography, Genetic Algorithm, PlayFair Algorithm, Pseudo Random Sequence

I. INTRODUCTION

Cryptography is a basic building block of Computer Security. It involves the secure transmission of information over insecure communication channels by means of encryption of data. The study and practice of various techniques to hide and protect information from adversaries is the basis of Cryptography. Cryptography involves the encryption and decryption of data

A scheme used for encryption constitutes a cryptographic system or cipher [1]. Techniques used for decrypting a message without any knowledge of the encrypting details fall into the area of cryptanalysis [1]. This paper describes a technique for modifying the classical PlayFair Cipher for enhancing its security using Chaos Theory and Genetic Algorithms.

II. PLAYFAIR CIPHER

The classical PlayFair Cipher uses a 5*5 matrix consisting of uppercase English alphabets with I/J being considered as 1 element. The matrix is formed by putting the letters of the keyword from left to right and top to bottom in the matrix, without considering duplicates. The rest of the matrix is filled by the remaining letters of the uppercase English alphabet. For the encryption of a plaintext, we form digrams (i.e. 2 symbol pairs) of the plaintext. If there are duplicate letters in the diagram(for e.g. 'll' or 'oo' in 'balloons'), another filler letter X is taken instead. For a standalone symbol left in the end, which does not form a diagram, the filler letter (for e.g. X) can be taken.

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

Figure 1. Classical Playfair Matrix

e.g.
Keyword - M A T C H E R
Plaintext - C A B B A G E S
Digrams - CA BX BA GE SX
CipherText - HT KT RT OG QY

A. The working of the Algorithm

Encryption Rules

For each of the digrams of the plaintext encode as follows

1 If the symbols of the digram are in the same row of the PlayFair matrix, then replace the letter of the plaintext with the letter to its right in the matrix, the first element of the row will follow the last circularly.
2 If the symbols of the digram are in the same column of the PlayFair matrix, then replace the letter of the plaintext with the letter below it in the matrix, the top element of the column will follow the bottom-most element circularly.

3 For other digrams, that do not have letters in the same row or column, encode as follows. Each plaintext letter is replaced by the letter that is in its own row and the column of the other letter in the digram.

Decryption Rules
For each of the digrams of the ciphertext, decode as follows

1. If the symbols of the digrams are in the same row of the PlayFair matrix, then replace the letter of the ciphertext with the letter to its left in the matrix, the last element of the matrix will follow the first element circularly
2. If the symbols of the digrams are in the same column of the PlayFair matrix, then replace the letters of the ciphertext with the letter to the top of it in the matrix, the bottom-most element will follow the top-most element circularly.
3. For other digrams, that do not have letters in the same row or column, decode as follows. Each ciphertext letter is replaced by the letter that is in its own row and the column of the other letter in the digram.

Advantages of Playfair Cipher is that in the monoalphabetic cipher there are only 26 letters, in PlayFair there are 26 * 26 = 676 digrams. It is more difficult to identify digrams that individual letters.

III. RELATED WORK

There have been many variants of the Playfair Cipher. It has been extended to 8 * 8 matrix [2] capable of encrypting both the uppercase as well as lowercase alphabets and numeric digits. I and J are considered different alphabets. This cipher is difficult to break as the length of keyword becomes large. This algorithm uses "@"character for adding to a standalone character and for a duplicate character digram it uses "&".

The characteristic of the Playfair cipher is that if the order of the plaintext in a digram is reversed then the ciphertext is also reversed. E.g. if KO is encrypted to DM then OK will encrypt to MD. So this feature can be exploited for cryptanalysis.

There have been several enhancements to the security of the PlayFair Cipher. "A Novel Approach to Security using Extended Playfair Cipher" by Shiv Shakti Srivastava, Nitin Gupta [3] uses a Linear Feedback Shift Register to apply to the ciphertext generated by PlayFair to get a permuted sequence of bits. The limitation of this approach is that LFSR being a linear system is easy for cryptanalysts. The Berlekamp-Massey algorithm can be used to construct an LFSR if we can recover a stretch of LFSR output stream. This constructed LFSR can be used for recovering the plaintext.

Genetic Algorithms can be used to reduce the problem of LFSR-based ciphers. They introduce non-linearity.

IV. GENETIC ALGORITHMS

Genetic Algorithms GAs are adaptive methods which may be used to solve search and optimisation problems. They are based on the genetic processes of biological organism .Over many generations natural populations evolve according to the principles of natural selection and survival of the fittest. By mimicking this process genetic algorithms are able to evolve solutions to real world problems if they have been suitably encoded. For example GAs can be used to design bridge structures for maximum strength weight ratio or to determine the least wasteful layout for cutting shapes from cloth.

The basic principles of GAs were first laid down rigorously by Holland [5].

In nature individuals in a population compete with each other for resources such as food water and shelter. Those individuals which are most successful in surviving and attracting mates will have relatively larger numbers of offspring.

This means that the genes from "fit" individuals will spread to an increasing number of individuals in each successive generation. The combination of good characteristics from different ancestors can sometimes produce "superfit" offspring whose fitness is greater than that of either parent. In this way species evolve to become more and more well suited to their environment.

GAs work with a population of "individuals" .The individual represents a solution to the problem. Each individual is assigned a fitness score according to how good a solution to the problem it is . For example the fitness score might be the strength/weight ratio for a given bridge design. The highly fit individuals are given opportunities to reproduce by cross breeding with other individuals in the population. This produces new individuals as offspring which share some features taken from each "parent". A whole new population of possible solutions is thus produced by selecting the best individuals from the current generation and mating them to produce a new set of individuals. This new generation contains a higher proportion of the characteristics possessed by the good members of the previous generation. By favouring the mating of the more fit individuals the most promising areas of the search space are explored. If the GA has been designed well the population will converge to an optimal solution to the problem.

A. Basic Principles of Genetic Algorithms

Before a GA can be run, a suitable coding (or representation) for the problem must be devised. We also require a fitness function which assigns a figure of merit to each coded solution. During the run parents must be selected for reproduction and recombined to generate offspring.

Coding
A potential solution to a problem is represented as a set of parameters (genes) which are joined together to form a string of values (chromosomes).

Fitness Function
A fitness function is devised for each problem. Given a particular chromosome the fitness function returns a single numerical fitness or figure of merit which is supposed to be proportional to the utility or ability of the individual which that chromosome represents.

Reproduction
During the reproductive phase of the GA individuals are selected from the population and recombined, producing offspring which will comprise the next generation. Having selected two parent their chromosomes are recombined typically using the mechanisms of crossover and mutation.

Crossover takes two individuals and cuts their chromosome strings at some randomly chosen position to produce two head segments and two tail segments. The tail segments are then swapped over to produce two new full length chromosomes. The two offspring each inherit some genes from each parent. This is known as single point crossover.

Mutation is applied to each child individually after crossover. It randomly alters each gene with a small probability (typically 0.001).

V. CHAOS THEORY

Chaos is an inter-disciplinary theory which states that in the seeming randomness of chaotic complex systems, there are patterns, constant feedback loops, self-similarity, repetition, fractals, self-organization, and sensitive dependence on initial conditions [6]. The butterfly effect (BE) describes how a small change in one state of a deterministic non-linear system can result in large differences in a later state, that means a butterfly flapping its wings in Italy can cause a hurricane in Texas [7].

A relationship exists between Chaos and Cryptography [8]

1. Ergodicity and Confusion
2. Sensitivity towards Initial Condition
3. Diffusion
4. Deterministic dynamics and Pseudo-randomness
5. Structure Complexity and Algorithm Complexity

Due to ergodicity property of chaos it is made certain that chaotic variables will visit all state non-repeatedly within a certain range according to its own laws [9]. So, the ergodicity property is used as an optimization mechanism preventing fall into local minimum solution. The sensitiveness to the initial state, makes certain there are not two identical new populations even if the two best fit solutions obtained by sequential evolving procedures are very close [10]. Therefore, population thus obtained not only preserve the best fit chromosome, but the populations are varied.
By using these properties of Logistic Maps and Tent Chaotic Maps we produce chaotic sequences. A great number of such sequences can be obtained by differing the initial condition.
The different chaotic mappings are as follows: [11]

1.Logistic Maps - This map is a 2-degree polynomial:

$$X_{n+1} = \mu X_n (1-X_n) \quad X_n \epsilon(0,1) \qquad (1)$$

When μ is between 0 and 4
Where X ϵ (0, 1) and $3.57 < \mu \leqslant 4$, the system has proven to be chaotic state

2.Tent Maps is also a well-known Chaotic System
It is a discrete-time dynamical system, with equation:

$$X_{n+1} = \mu X_n \text{ if } X_n < ½ \qquad (2)$$
$$\mu (1-X_n) \text{ if } X_n >= ½$$

Where μ is a control parameter when $1 < \mu < 2$ the system displays chaotic behavior.

VI. PROPOSED METHODOLOGY

In this paper, the approach proposed combines the Playfair algorithm encryption and the optimised chaotic genetic algorithm. The string encrypted once with the Playfair cipher is again subjected to encryption using the Genetic Algorithm with a Pseudo-random number generated by 1-D chaotic map. This method proposes using a Playfair Cipher with 10 * 9 matrix. The matrix uses English letters (both lowercase as well as uppercase letters), numbers, punctuation marks, and several special characters [12].

M	**o**	**n**	**a**	**r**	**c**	**h**	**y**	**b**	**d**
E	**f**	**g**	**i**	**j**	**k**	**l**	**m**	**p**	**q**
S	**t**	**u**	**v**	**w**	**x**	**z**	**A**	**B**	**C**
D	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**N**
O	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**
Y	**Z**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
8	**9**		**,**	**.**	**/**	**;**	**‘**	**[**	**]**
<	**>**	**?**	**:**	**{**	**}**	**-**	**=**	**!**	**@**
#	**$**	**%**	**~**	**&**	*****	**(**	**)**	**_**	**+**

Figure 2 – Modified Playfair Matrix 10 * 9

The secret keyword is 'Monarchy'. The lowercase, uppercase letters of the English Alphabet as well as numbers and

punctuation marks are able to be encrypted. A white space character is also introduced, therefore the white space can also be included in the plain text. The padding character to be used is '~'. The duplicate characters in the digrams as well as odd number of characters are padded with '~'.

We use a Pseudo-random number generated which is used as a crossover point, or the Chaotic Crossover operator on Genetic Algorithms that improves the randomness and diversity of the Genetic Algorithm encrypted strings.

The following are the steps for encrypting a binary sequence. [13]

1. The pseudo-random binary sequence is generated using the 1-Dimensional Chaotic map.
2. The pseudo-random binary sequence is converted into a decimal pseudo-random sequence from 0 to 7 as Z_n
3. The string to be encrypted is modified for creating confusion by using byte substitution method
4. Loop
5. Take the two binary sequences to be encrypted A_i and A_{i+1}
6. Do the crossover operation using Zn on the two binary sequences and generate the two new sequences B_i and B_{i+1}
7. Now the two newly generated sequences are encrypted as C_i and C_{i+1}

   $X_i = Z_i \oplus (Z_i <<4)$

   $X_{i+1} = Z_{i+1} \oplus (Z_{i+1} << 4)$

   $C_1 = B_1 \otimes X_i$

   $C_2 = B_2 \otimes X_{i+1}$
8. Perform the byte substitution on the encrypted string again for creating confusion

The following is the process for decrypting a binary sequence

The decryption process will follow the opposite series of steps of the encryption. A pseudo-random sequence generated using chaos for encryption is shared and is used for decryption using the reverse steps.

A. Illustration of the cipher

Let us take the plaintext - This cipher is strong.

Using the 10 * 9 matrix above, and breaking up the text into digrams, the following digrams and ciphertext are obtained.

Th - Uc

is - ev

<space>c - /n

ip - jq

he - Ml

r<space> - n.

is– ev

<space>s - 8u

tr - wo

on - na

g~ - i%

Applying the genetic algorithm with Chaos Uc is the first digram

U – ASCII value - 085 – 0101 0101

c – ASCII value – 099 – 0110 0011

Encryption using Chaotic PseudoRandom Sequence

1. If $\mu = 3.57$ and $x_0 = 0.7$, threshold value $t = 0.5$, the pseudo-random binary sequence using 1-D chaotic map is

   0 1 0 0 1 1 ….
2. The random sequence converted to decimal value using general chaotic models

   $Z_n = 2, 3, 1, 5$
3. $Z_i = 2$ and $Z_{i+1} = 3$

   Using pseudo-random sequence Z_i value, perform crossover operations on the digram binary strings

   A_1 – U – 0101 0101

   A_2 – c – 0110 0011

   After the crossover operation

   B_1 – 0101 0111

   B_2 – 0110 0001

   $X_i = Z_i << 4$ – 00100000

   $X_{i+1} = Z_{i+1} <<4$ – 00110000

   $C_1 = B_1 \otimes X_i$

   $C_1 = 01110111$

   $C_2 = B_2 \otimes X_{i+1}$

   $C_2 = 01010001$

   C_1 – 119 – w

   C_2 – 81 – Q

   So the encrypted digram string is for Uv is wQ

   Correspondingly the entire PlayFair encrypted string is encrypted again using Genetic Algorithm with Chaos

   The decryption process can be illustrated as follows. The decryption process follows the reverse sequence of steps.

1. First the ciphertext is converted to its ASCII binary equivalent. C_1 = w – 01110111 C_2 = Q – 01010001

2. The pseudo-random sequence which was generated using the 1-D chaotic operator was shared and can be used for decryption
   $Z_i = 2$ $Z_{i+1} = 3$
   $X_i = 00100000$
   $X_{i+1} = 00110000$
   $I_1 = C_1 \otimes X_i$
   $I_2 = C_2 \otimes X_{i+1}$
   $I_1 = 01010111$
   $I_2 = 01100001$

3. Performing the crossover on the on the intermediate strings I1 and I2 to retrieve the decrypted digram
   $D_1 = 01010101 – 85 – U$
   $D_2 = 01100011 – 99 – c$

   Now decipher using the modified Playfair matrix with 10 * 9 Characters to get Th.
   Similarly, the whole string is decrypted to get back the original string, This cipher is strong.

B. Encryption Process

The encryption algorithm is shown in the figure below

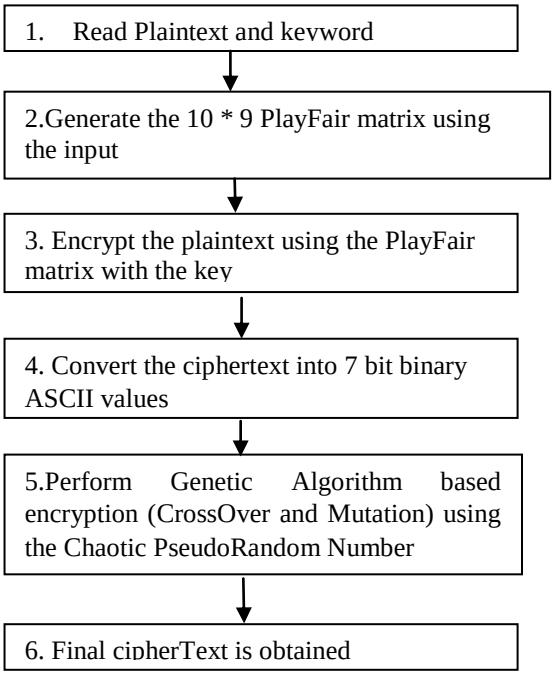

Figure 3 – Proposed Encryption Process

C. Decryption Process

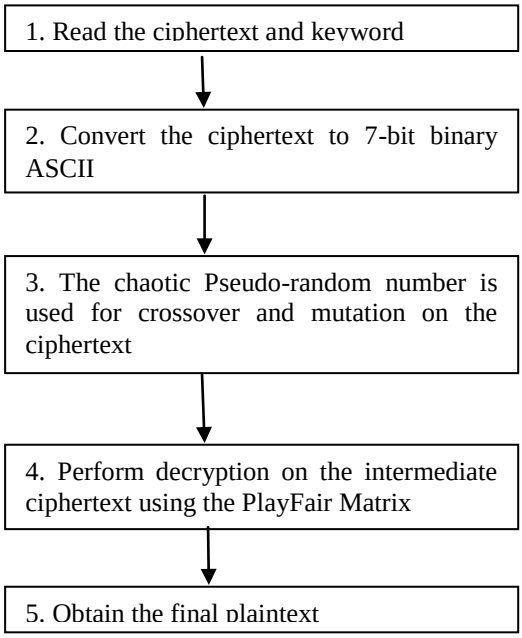

Figure 4 – Proposed Decryption Process

D. Block Diagram of the Cryptographic Process using 1-D Chaotic Map

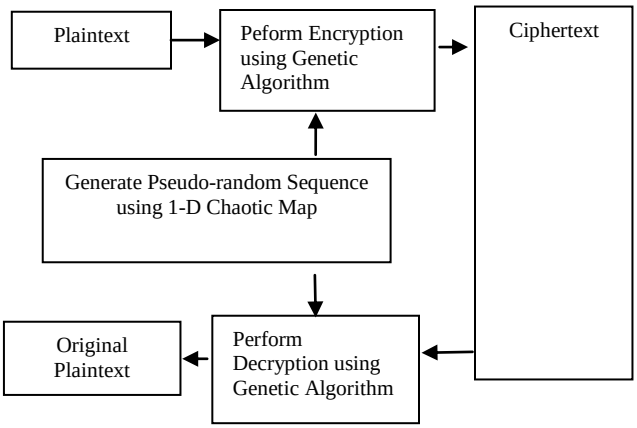

Figure 5–Block Diagram of the Proposed Cryptographic Process

VII. ANALYSIS OF THE PROPOSED APPROACH

1. Ciphertext only Attack - The length of our ciphertext is 22.

22 * 7 = 154 bits

The size of the plaintext to be searched is 2 ^ 154 bits which is quiet a large number. Hence ciphertext only attack becomes difficult.

2. Known PlainText - For the known plaintext attack , not only is the original plaintext encrypted with PlayFair algorithm with a 10 * 9 matrix , but also encrypted again using Genetic algorithm with Chaotic Pseudorandom number. Thus it is difficult to establish any relationship between the plaintext and the final ciphertext.

VIII. CONCLUSION

In this paper we have analysed the various approaches for the PlayFair Cipher and proposed a variant that uses Chaotic Genetic Algorithm [14 and 15]. The proposed approach uses 10 * 9 matrices instead of the classical 5 * 5 matrix. It includes the 26 uppercase letters, 26 lowercase letters, 0-9 numerals and a number of special characters including white space. The modified PlayFair is doubly encrypted with Genetic Algorithm with chaotic pseudorandom operator to improve its security. Thus this proposed algorithm is more robust and secure than the original PlayFair algorithm.

REFERENCES

[1]W Stallings," Cryptography and Network Security - Principles and Practice", Fourth Edition (Pearson Education), USA, pp. 30, 2017

[2] Siddhartha Sankar Biswas,Mohammad Sadiq Nisar Siddiqui and Parul Agarwal, "Genetic Extension of Playfair Cipher Using Modified Matrix", International Journal of Computer & Mathematical Sciences, ISSN 2347 – 8527, Volume 6, Issue 6, June 2017pp. 25-30

[3] S S Srivastava, N Gupta, "A Novel Approach to Security using Extended Playfair Cipher", International Journal of Computer Applications, Volume 20– No.6, April 2011

[4] D Beasley, D R Bull and R R Martin, "An Overview of Genetic Algorithms: Part 1, Fundamentals", University Computing, 15(2) 58-69, 1993

[5]JH Holland, "Adaptation in Natural and Artificial Systems", MIT Press, USA, 1992

[6] G. Boeing, "Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction.", Systems 4, no. 4: 37, 2016.

[7] S Zaminpira and S Niknamian, "How Butterfly Effect or Deterministic Chaos Theory in Theoretical Physics Explains the Main Cause of Cancer", EC Cancer, Volume 2, Issue 5, pp. 227-238, 2017

[8]R. Brown, L. O. Chua, "Clarifying chaos: Examples and counterexamples", International Journal of Bifurcation and Chaos, Volume 06, Issue 02, February 1996

[9]Tan D.,"Application of Chaotic Particle Swarm Optimization Algorithm in Chinese Documents Classification", In the proceedings of the 2010 International Conference on Granular Computing, USA, pp. 763-766, 2010

[10] L J C Zi-xing, L Jian-qin, "A Novel Genetic Algorithm Preventing Premature Convergence by Chaos Operator", Journal of Central South University of Technology, Volume 7, Issue 2, pp 100–103, June 2000

[11]M Javidi and R Hosseinpourfard, "Chaos Genetic Algorithm Instead Genetic Algorithm, The International Arab Journal of Information Technology", Vol. 12, No. 2, March 2015

[12] S Basu ,U K Ray, "Modified Playfair Cipher using Rectangular Matrix", International Journal of Computer Applications (0975 – 8887), Volume 46 Issue No.9, May 2012

[13] A Kumar, M. K. Ghose, "Overview of Information Security Using Genetic Algorithm and Chaos", Information Security Journal: A Global Perspective, Volume 18, 2009

[14] Siddhartha Sankar Biswas, Mohammad Sadiq Nisar Siddiqui1 and Jawed Ahmed , "An Extension of Playfair Cipher Using Modified Matrix", International Journal of Computational Intelligence Research

ISSN 0973-1873 Volume 13, Number 5 (2017), pp. 923-931.

[15]Siddhartha Sankar Biswas, Saman , Md.Tabrez Nafis and Mohammad Sadiq Nisar Siddiqui , "Addendum of Playfair Cipher in Hindi", Advances in Computational Sciences and Technology, ISSN 0973-6107 Volume 10, Number 5 (2017) pp. 977-983.

Authors Profile

Ms. Archi Seth pursued Bachelors of Science from Delhi University and Masters of Computer Applications from Agra University. She is currently pursuing M Tech from Jamia Hamdard University, New Delhi. Her main research area of interest focuses on Cryptography Algorithms and Machine Learning. She has 10 years of Industry experience in Software Design and development.

Dr. Siddhartha Sankar Biswas is Assistant Professor at Department of Computer Science and Engineering, School of Engineering Sciences and Technology, Jamia Hamdard, New Delhi, India. He perused B.Tech. (Information Technology) from Punjab Technical University, Jalandhar, M.Tech. (Computer Engineering) and M.B.A.(Information Technology) from M.D.U., Rohtak and Ph.D. (Computer Engineering) from Jamia Millia Islamia, New Delhi

