A Hybrid Computational Intelligence-based Technique for Automatic Cryptanalysis of Playfair Ciphers

Maiya Din*, Saibal K. Pal†, S.K. Muttoo‡, and Sushila Madan‡

*DRDO-HQrs, DF&MM, DRDO Bhawan, Delhi, India
†DRDO-Scientific Analysis Group, Delhi, India
‡University of Delhi, Delhi, India
*E-mail: anuragimd@gmail.com

ABSTRACT

The Playfair cipher is a symmetric key cryptosystem-based on encryption of digrams of letters. The cipher shows higher cryptanalytic complexity compared to mono-alphabetic cipher due to the use of 625 different letter-digrams in encryption instead of 26 letters from Roman alphabets. Population-based techniques like Genetic algorithm (GA) and Swarm intelligence (SI) are more suitable compared to the Brute force approach for cryptanalysis of cipher because of specific and unique structure of its Key Table. This work is an attempt to automate the process of cryptanalysis using hybrid computational intelligence. Multiple particle swarm optimization (MPSO) and GA-based hybrid technique (MPSO-GA) have been proposed and applied in solving Playfair ciphers. The authors have attempted to find the solution key applied in generating Playfair crypts by using the proposed hybrid technique to reduce the exhaustive search space. As per the computed results of the MPSO-GA technique, correct solution was obtained for the Playfair ciphers of 100 to 200 letters length. The proposed technique provided better results compared to either GA or PSO-based technique. Furthermore, the technique was also able to recover partial English text message for short Playfair ciphers of 80 to 120 characters length.

Keywords: Playfair cipher; Cryptanalysis; Swarm intelligence; Multiple particle swarm optimization; MPSO; Genetic algorithm

1. INTRODUCTION

In cryptography, the techniques are studied for securing information from unauthorized persons/intruders. Cryptology consists of two branches known as cryptography and cryptanalysis. Cryptography is a science of making communication unintelligible to others except the authorized recipient. In cryptanalysis, cryptanalytic attacks are applied for breaking cipher text to obtain its plain text without knowledge of applied secret key. The study of classical ciphers helps in understanding the basic principles of cryptography. The classical ciphers can be seen as building blocks of modern cryptosystems. At remote locations, where modern cipher machines are not available, conventional encryption techniques are still being used.

In cryptography, Playfair cipher is a well-known multiple-letter encryption cipher invented by Charles. The name of the cipher comes in crypto literature from the name of the lord Playfair who strongly promoted the cipher. The encryption process is based on arrangement of 25 English letters in a 5 x 5 table without repetition in a way that letters ‘J’ and ‘I’ are located in the same cell of the table.

Swarm intelligence (SI) is an important branch of artificial intelligence. The SI is applied for finding solution of many practical and complex optimization problems, and to find an approximate optimal solution of a crypto problem, a suitable Computational Swarm algorithm is applied. Cryptanalysis of any cipher system can be formulated as an optimization task.

2. LITERATURE REVIEW

With the availability of powerful computers and data analysis techniques, specific parts of the cryptanalysis process can be automated. Computational swarm algorithms and Evolutionary algorithms may be some mechanisms of automating the process. Soft-computing and CSI-based algorithms are very useful in cryptanalysis of Classical ciphers, Stream ciphers and Block ciphers. Benjamin applied evolutionary algorithms in cryptanalysis of Playfair cipher but could not get good results due to implementation issues. The work of Laskari, et al. is considered as an important effort for solving crypto problems by applying computational intelligence (CI). Authors formulated three discrete optimization problems as crypto-problems of public key cryptosystem and solved them by applying evolutionary methods. They also applied PSO as optimization technique in cryptanalysis of DES. Authors concluded that CI methods could be employed effectively for analysis of weak cryptosystems. Cowan made efforts to break short Playfair ciphers with the simulated annealing algorithm. The author suggested simulated annealing for solving short Playfair crypts. Negara proposed an evolutionary approach for cryptanalysis of classical Playfair cipher and concluded that generic evolutionary schemes are suitable in solving crypto problems. Hammood applied memetic algorithm
for breaking Playfair ciphers. Noor et al. presented an automatic cryptanalysis of Playfair cipher using compression. Vimalathithan et al. proposed cryptanalysis of simplified-AES using PSO method by applying cipher-text only attack. Ahmed et al. applied MPSO to break transposition cipher system by using multiple swarms instead of single swarm to determine the best key in less number of trials.

Din et al. proposed cuckoo search based swarm method in analysis of LFSR- based cipher to find initial states of considered LFSR. The authors also applied GA for computing initial states of considered three LFSRs of Geffe Generator based stream cipher. Divide and conquer approach is used in formulating the cryptanalytic attack. Authors also applied binary PSO (BPSO)-based Swarm technique in cryptanalysis of Geffe Generator and obtained encouraging results. Din et al. also developed and tested PSO-based technique for solving RC4-based crypts. Mehrotra et al. used Chaos in Grey Wolf Optimizer (CGWO) for prime factorization, which was further used in cryptanalysis of public key cryptosystem like RSA. Bansal & Deep proposed modified binary PSO for solving Knapsack Problem and Multidimensional Knapsack Problem.

The reviewed literature and reported cryptanalytic results motivated the authors to propose a hybrid technique based on MPSO and GA for breaking the Playfair ciphers by automatic cryptanalysis.

3. DETAILS OF PLAYFAIR CRYPTOSYSTEM

The Playfair cryptosystem is a symmetric encryption scheme based on digram substitution. The scheme, invented by Charles Wheatstone, encrypts digrams of letters compared to single letter substitution in SST cipher. This cryptosystem is more complex than Vignere and SST cipher systems. The Playfair cryptosystem is cryptographically stronger than other classical cryptosystems. It is harder to break the system since frequency analysis of N-Gram features is not successful in this scheme.

The Playfair scheme utilizes a 5 x 5 table containing any keyword. The Playfair key table is generated by filling the letters of the keyword (dropping alphabet ‘J’ and any repeating letters) initially, and then by filling the remaining positions with rest of the alphabets from ‘A’ to ‘Z’ (Fig. 1) using keyword ‘BACKGROUND’.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>A</th>
<th>C</th>
<th>K</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>O</td>
<td>U</td>
<td>N</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>H</td>
<td>I</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>P</td>
<td>Q</td>
<td>S</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Key table of Playfair.

The selected keyword is filled in the top row of the table from left to right. The remaining letters (without repeating any letter and dropping letter ‘J’) can be filled in the 5 x 5 table, comprising the Playfair key. In order to encrypt a plaintext message, it is divided into digrams and letter ‘J’ is replaced by ‘I’ in the entire message. Less frequent letters (‘X’ or ‘Q’) is inserted within each doublet (digram with same letters, e.g. ‘AA’, ‘TT’). In case of odd number of letters in the modified message, one of the aforementioned less frequent letters is added at the end of the message. If both letters of the pair occur on the same row they are respectively replaced with their immediate right letter employing circular right shift. If both letters of the pair occur on the same column, they are respectively replaced with the letter immediately below that letter employing circular down shift. Otherwise, when the letters occur in diagonally opposite manner in the table they are replaced with letters placed at the corners of the rectangle defined by the plaintext pair in such a way that the first letter of the encrypted pair and first letter of plaintext pair lies on the same row and similarly for second letters of both pairs.

Playfair cryptosystem have some weaknesses. Firstly, the repeated plain digrams create corresponding repeated digrams in the crypt. Secondly, digram frequency count in cipher can detect most of the frequent digram in English. Thirdly, frequency of the most frequent cipher monogram matches with frequency of the most frequent monograms in English. The hybrid technique is developed to break Playfair ciphers by exploiting these weaknesses and accordingly the fitness function is formulated based on the most frequent N-Gram features.

4. PROPOSED HYBRID COMPUTATIONAL INTELLIGENCE TECHNIQUE

Many techniques based on computational intelligence have been devised for breaking cryptosystems. A novel hybrid computational intelligence-based technique with coupling MPSO and GA has been proposed for cryptanalysis of Playfair ciphers.

4.1 Genetic Algorithms

Genetic algorithms (GAs) are nature inspired population-based evolutionary algorithms based on survival of the fittest principle. These algorithms are being used for approximate solving optimization problems. In the proposed work, GA is applied on potential solution keys computed by particle swarms (modules PS-1, PS-2, PS-n) as shown in the flowchart (Fig. 7). These keys constitute initial population of chromosomes for GA. The size of population is equal to the number of swarms. Mutation is applied on each chromosome (key) by swapping two randomly selected letters in the key. Auto-crossover is applied by swapping any two randomly selected rows or columns of each parent Playfair Key table as illustrated in the Fig. 2. Auto-crossover operator is easy to implement and provide encouraging results as compared to conventional crossover between two chromosomes in GA module of the proposed technique for breaking Playfair cipher.

Mutation operator of GA module of MPSO-GA technique improve the approximately correct solution keys computed by multiple swarms modules (PS-1, PS-2, PS-n). In most of the Playfair ciphers, GA module provides solution keys including correct Playfair Key with optimal fitness values after convergence.
4.2 Particle Swarm Optimization

Particle swarm optimization (PSO) was a successful computational swarm technique proposed by Eberhart & Kennedy. The technique was initially applied for solving continuous optimization problems, but now discrete PSO algorithm has been used extensively in many other discrete optimization tasks.

The implementation of PSO technique (shown in Fig. 3) is based on Eqs. (1) and (2):

\[ v_{id} = (w \cdot v_{id}) + [c_1 \cdot r_1(p_{id} - x_{id})] + [c_2 \cdot r_2(p_{gd} - x_{id})] \]  
\[ x_{id} = x_{id} + v_{id} \]  

where \( r_1 \) and \( r_2 \): random values lying in \([0, 1]\), \( c_1 \) and \( c_2 \): cognitive and social parameters, \( x_i = (x_{i1}, x_{i2}, ..., x_{in}) \) is \( i^{th} \) particle position; \( p_i = (p_{i1}, p_{i2}, ..., p_{in}) \) represents the previous best particle position and \( p_{gd} \) represents global best position.

\( v = (v_{i1}, v_{i2}, ..., v_{in}) \) is velocity of \( i^{th} \) particle and \( w \) is the inertia weight.

\[
\text{Parent Key: BACKGROUNDFHILMPQSTVWXYZ} \\
\text{After mutation (at Letter Position 3 and 10):} \\
\text{Child Key: BADKGROUPNEFHILOPQSTVWXYZ.}
\]

![Mutation and auto-crossover](image)

**Figure 2. Mutation and auto-crossover:** (a) Mutation, and (b) Auto-Crossover (Swapping of Row-2 and Row-5).

The value of each particle is computed as

\[
\text{Fitness Value} = \sum_{i=1}^{60} \left( f_i \cdot w_i \right) / L \\
\text{Here, } L \text{ is length of the crypt, } f_i \text{ and } w_i \text{ are frequency of } i^{th} \text{ feature and corresponding weight respectively.}
\]

Total 60 N-gram features are selected including 15 of each N-gram (i.e. the most frequent monograms, digrams, trigrams and tetagrams, mentioned in Appendix-1). Position vector of global best particle provide correct solution Key, which gives correct plain message.

![Fitness value for plain text vs crypt](image)

**Figure 4. Fitness value for plain text vs crypt.**

According to Fig. 4, the applied fitness function is able to distinguish the plain text with respect to the corresponding encrypted text clearly.

![Fitness value vs correctness (%) of decrypted text](image)

**Figure 5. Fitness value vs correctness (%) of decrypted text.**

As shown in Fig. 5, the fitness value increases as percent correctness of the decrypted message tends to 100% (i.e. correct solution key). The above graphs help in deciding threshold value of fitness function and storing solution keys with fitness value greater than threshold value.

4.3 Multiple Particle Swarm Optimization

Multiple swarms of particles are taken in the swarm technique. This technique is suitable in case of very large solution space having some clusters of solutions, as in Key space of conventional cryptosystems. In proposed technique, this is assumed that Crypto-Keys have some clusters of similar keys with minor variation at few places. In Playfair cipher,
Keyword: ENCRYPT
Correct Playfair Key: ENCRYPTABDEFGHIJKLMNOPQSVWXYZ

<table>
<thead>
<tr>
<th>Playfair Key cluster:</th>
<th>fitness value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENCRYPTABDEFGHIJKLMNOPQSVWXYZ</td>
<td>1.602</td>
</tr>
<tr>
<td>ENCRABDFGIJKLMNOPQSVWXYZ</td>
<td>1.760</td>
</tr>
<tr>
<td>ENCRYPTABDEFGHIJKLMNOPQSVWXYZ</td>
<td>1.805</td>
</tr>
<tr>
<td>ENCRYPTABDEFGHIJKLMNOPQSVWXYZ</td>
<td>3.057</td>
</tr>
<tr>
<td>ENCRYTADEFGHIJKLMNOPQSVWXYZ</td>
<td>2.215</td>
</tr>
<tr>
<td>ENCRYPTABDEFGHIJKLMNOPQSVWXYZ</td>
<td>1.819</td>
</tr>
<tr>
<td>ENCRYPTABDEFGHIJKLMNOPQSVWXYZ</td>
<td>1.668</td>
</tr>
</tbody>
</table>

Figure 6. Playfair Key cluster.

25 letter keys with variation at few letters of taken Keyword constitute such cluster. Each Particle Swarm represents one Key-cluster as shown in Fig. 6.

Each swarm is simulated by taking NP particles each of 25-dimension corresponding to Key length. PSO parameters (NP, w, C1, C2) are considered similar for each swarm after tuning of the parameters. Number of swarms is determined according to probable clusters of solution Keys with respect to given Playfair crypt. These Keys have fitness value higher than threshold fitness value (lying between 1.8 and 2.2). This threshold value of fitness function is computed according to fitness value for plain English text based on frequency of most frequent monograms, digrams, trigrams and tetragrams.

4.4 Hybrid Computational Intelligence Technique

A hybrid technique is proposed as combination of above mentioned MPSO and GAs for solving Playfair ciphers. In MPSO, each swarm provides one optimal solution key. Therefore multiple solution keys are computed by MPSO technique. By experimentation, it has been observed that in few cases the correct Playfair key is missing in the computed Keys but they are similar to correct Key with variation at 2 to 4 positions in the key as shown in Fig. 6. Genetic Algorithm is suitable to construct correct solution key from partial solution Keys. A novel hybrid technique based on MPSO and GA has been developed and applied in Playfair cryptanalysis. Flowchart of the technique is shown in Fig. 7.

In this technique each particle position represents one Playfair key (Length 25 letters) as ASCII representation of ‘A’ to ‘Z’ (After removing ‘J’). Every Playfair key is a permutation of these letters. These permutation vectors are transformed into real numbers (lying between 0.0 and 1.0) by subtracting 64 from each ASCII value and dividing by maximum number 25. These vectors are taken as position vectors of particles and used in computing their velocity vector. Inverse transformation is applied to get Playfair key (in letter form).

GA is applied on near optimal Playfair keys computed by MPSO. These keys are considered as chromosomes constituting initial population. The size of population is equal to the number of swarms considered in MPSO. Here, mutation is applied on each key by swapping two randomly selected letters in the key. Crossover is applied by swapping any two rows out of the five rows in the Playfair key table. GA module provides solution keys including correct Playfair key with maximum fitness value in the population after convergence. Partial solution keys are converted into optimal solution keys by GA module.

5. ANALYSIS OF EXPERIMENTAL COMPUTED RESULTS

Software implementation of proposed hybrid technique has been carried out in MATLAB Software (Version R2013a) and executed on 2.8 GHz computing machine. The software has been tested on Playfair crypts of English texts, encrypted with different keys generated by different Keywords. A number of experiments were conducted for different values of these parameters and the optimized solutions are reported here. For optimal solution Playfair keys, MPSO-GA parameters: Number of particles (NP) = 30 to 100, Inertia weight (w) = 0.8 to 1.0 and Acceleration constants (C1, C2) = 0.7 to 1.5 for Maximum iterations = 106. The considered GA parameters: Population size = 200 to 500, Mutation Probability = 0.08 to 0.16. Elitism is applied in GA module by keeping 5% to 10% best solution of current population to next generation. Fine tuning of these parameters provided optimal key solutions. The computed results are as shown in Tables 1 and 2.

As per described results in Table-2, for solving taken crypt (Length = 100 letters) with keyword length (K = 7 letters), CPU time taken is 5899.18 second corresponding to MPSO parameters (No. of Swarms = 500, NP = 100, w = 0.85, C1 = 0.9 and C2 = 0.9). The results obtained by applying hybrid technique on Playfair crypts of length 80 to 100 letters with different keys are shown in Figs 8 - 11.
According to cryptanalytic results for short Playfair ciphers encrypted with different keys including keywords of length 4 to 7 letters, the proposed technique is able to compute clear message (length up to 80 letters). In few cases when correct key is not obtained, partial correct message (70% to 80%) is computed by proposed hybrid technique. Time complexity of the hybrid technique depends on time complexity of PSO and GA. Time complexity of PSO algorithm is $O(n \cdot t \cdot \log(n))$, where $n$ is number of particles and $t$ is number of iterations. Time complexity of Genetic algorithm is $O(g \cdot (n \cdot m + n \cdot m + n))$, where $n$ is size of population, $m$ is length of chromosome and $g$ is number of generations.

### 6. Conclusion

Cryptanalysis of Playfair ciphers was carried out by computing the correct secret key. An advanced MPSO and GA-based hybrid computational intelligence technique has been developed to compute the correct solution. According to achieved cryptanalytic results this technique is able to solve the Playfair ciphers in optimal time. As reported in Table 2, proposed technique is able to solve Playfair cipher generated with the key based on 7-letter keyword (e.g., MANGOES) in CPU time 5899.18 s. Parallel implementation of multiple Swarms of MPSO based proposed technique is another research initiative to reduce computational time for cryptanalysis. The developed hybrid technique provides encouraging results compared to Hill Climbing, PSO or GA techniques. The applied mutation and auto-crossover operators of GA module in the MPSO-GA technique improve the approximately correct solution keys computed by multiple swarms. In most of the Playfair ciphers, GA module provides solution keys including correct Playfair Key with optimal fitness values. The technique provides encouraging results to solve short Playfair ciphers of 80 to 120 letters length.

### REFERENCES


doi: 10.1007/978-3-319-91341-4


doi: 10.1007/978-981-13-0589-4_18


doi: 10.1007/978-3-540-71078-3_1


doi: 10.1080/01611190701743658


doi: 10.14429/dsfj.62.778


doi: 10.9790/0661-2005032532


doi: 10.1016/j.pisc.2016.04.098


doi: 10.1007/978-81-322-1768-8 45


CONTRIBUTORS

Mr Maiya Din received his MSc (Computer Science) from University of Allahabad, in 1989 and MBA (Information Technology) from Sikkim Manipal University, in 2012. Presently working as Scientist ‘G’ and Divisional Head at DFMM, DRDO-Hqr, New Delhi. His areas of interest are Cryptography, Information security, Soft computing and computational swarm intelligence. He has published more than 15 research publications in journals and conference proceedings. He is a recipient of Lab Scientist Award of DRDO, in 2014. In the current study, he proposed a hybrid technique based on MPSO and GA in cryptanalysis of Playfair ciphers.

Dr Saibal Kumar Pal received his MSc (Computer Science) from University of Allahabad, in 1991 and PhD from University of Delhi. Presently working as Scientist ‘G’ and Divisional Head at DRDO-Scientific Analysis Group, Delhi. He presently leads R&D teams working on different aspects of cryptography and information security. He has authored 2 books and more than 250 research publications in journals and conference proceedings. He is a recipient of Lab Scientist Award in 2010 and DRDO Scientist of the Year Award in 2012. In the current study, he reviewed proposed technique, implementation and computed results and provided valuable suggestions for improvement.
Prof. Sunil Kumar Muttoo received his MSc (Mathematics), MPhil (Mathematics), PhD (Coding Theory) in 1976, 1978, and 1982, respectively from University of Delhi and MTech (Computer Science and Data Processing) in 1990 from IIT Kharagpur. He has served as Professor & Head, Department of Computer Science, University of Delhi. He has contributed in the field of coding theory, cryptography, steganography and digital watermarking. In the current research work, he provided valuable guidance and suggestions for improving quality of the research paper.

Dr Sushila Madan received her MSc (Applied Mathematics) from IIT Delhi and MTech from BITS Pilani. Presently she is working as an Associate Professor in Computer-Science at Lady Shri Ram College Delhi University. She has authored books on securing mobile transactions, information system auditing and controls, information technology, e-commerce, multimedia and web-technology. She is conferred with ‘Distinguish Woman’ award from Venus International Foundation, Chennai, on 3rd March’ 2018. In the current study, she provided guidance in revision of manuscript according to received comments of the reviewers of Defence Science Journal.

Appendix

Appendix 1. List of the most frequent monograms, digrams, trigrams and tetragrams

<table>
<thead>
<tr>
<th>N-grams</th>
<th>Discriminating features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weights</td>
<td>3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 1, 1, 1</td>
</tr>
<tr>
<td>Digrams</td>
<td>TH, HE, IN, ER, AN, RE, ON, AT, EN, ND, TI, ES, OR, TE, OF</td>
</tr>
<tr>
<td>Weights</td>
<td>3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 1, 1, 1</td>
</tr>
<tr>
<td>Trigrams</td>
<td>THE, AND, ING, ION, TIO, ENT, ATI, FOR, HER, TER, HAT, THA, ERE, ATE, HIS</td>
</tr>
<tr>
<td>Weights</td>
<td>5, 5, 5, 5, 5, 3, 3, 3, 3, 2, 2, 2</td>
</tr>
<tr>
<td>Tetragrams</td>
<td>TION, THEN, THIS, THAT, MENT, THER, OUGH, IGHIT, ATED, WHEN, IOUS, THAN, ALLY, NESS</td>
</tr>
<tr>
<td>Weights</td>
<td>5, 5, 5, 5, 5, 4, 4, 4, 4, 4, 3, 3, 3, 3</td>
</tr>
</tbody>
</table>