# Analysis of Different Cryptosystems Using Meta-Heuristic Techniques

Aditi Bhateja[1]
Ambedkar Institute of Advanced Communication Technologies & Research, Delhi, India
aditibhateja89@gmail.com

***Abstract*—With the increasing usage of internet, the need of securing the information is also getting more important. The most widely used technique to make the information secure is cryptography. Cryptanalysis is a method to break the unreadable cipher text without having the key. This paper describes a method of deciphering encrypted messages of Vigenere cipher cryptosystems, Simple substitution cryptosystems and LFSR based cryptosystems by some nature inspired meta-heuristic techniques like swarm intelligence and evolutionary techniques.**

***Keywords—Vigenere cipher; Simple Substitution cipher; LFSR; PSO; fitness function; genetic algorithm***

## I. INTRODUCTION

Cryptography is a method of secret writing such that an unauthorized person cannot read the message and thus the integrity of the message is not violated. The strength of any cryptosystem lies in the key used to encrypt the text. A cryptosystem based on the key can be broadly classified in two categories viz. symmetric and asymmetric cryptosystems. A symmetric cryptosystem uses one key both for encrypting the plain text and decrypting the cipher text. While, an asymmetric cryptosystem uses two different keys, one for encryption and one for decryption. Symmetric cryptosystems can be further classified as monoalphabetic and polyalphabetic cryptosystems. In a monoalphabetic cryptosystem, each character is replaced with some another character with a fixed replacement structure. Simple substitution is an example of monoalphabetic cryptosystem in which the substitution of each alphabet is fixed. If alphabet "*s*" in plain text is replaced by alphabet "*a*", then every time we see the alphabet "*s*", we will replace it by "*a*". In polyalphabetic ciphers an alphabet can be replaced by any other alphabet without any fixed structure. Vigenere cipher is the most common, widely used polyalphabetic cipher. On the basis of number of characters to be encrypted, a cryptosystem is classified as stream cipher cryptosystem and block cipher cryptosystem. A stream cipher encrypts the text bit by bit whereas a block cipher encrypts a block or a byte at a time. Cryptanalysis on the other hand is breaking of the cryptic text without any knowledge of the key. A cryptanalyst can break almost any cipher with brute force attack. But since, the time complexity to break some cryptosystems is so high that it is practically infeasible to break the system. No doubt, technologies have advanced to make a secure cryptosystem but there are some technologies to break these cryptosystems. This paper gives a brief overview of some nature inspired meta-heuristic techniques and their applications in analysing some cryptosystems namely vigenere cipher, simple substitution and Linear Feedback Shift Register based cryptosystem.

In 2000, Mehmet *et al.* used Kasiski test and Index of coincidence concept [1] for analyzing vigenere cipher. In 2009, vigenere cipher was analyzed by T. Purusothaman *et al*. using dictionary attack [2]. Aditi *et al*. analysed Vigenere cipher with PSO with Markov chain random walk [3]. Uddin, Mohammad Faisal used Particle Swarm Optimization for analyzing Simple Substitution Ciphers [4]. Transposition cipher was analyzed by Heydari *et al.* [5] using genetic algorithm. They analyzed transposition cipher for key length upto 25. Simplified-DES and simplified AES was analyzed by Vimalathithan *et al*. using PSO based Computational Intelligence technique [6]. Properties of LFSR [7] like feedback period, time complexity and statistical behaviour were described by Faheem Masoodi *et al*. in 2012. They analyzed the security paradigms of LFSR and different techniques for security.

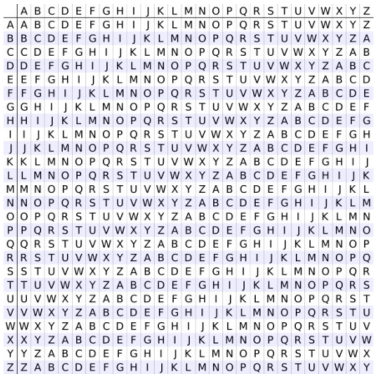
## II. CRYPTOSYSTEMS

A cryptosystem intakes the plain text and the key and outputs the cipher text, which is unreadable, and difficult to understand. Different cryptosystems are available in today's market. We will be analysing three most popular cryptosystems viz. vigenere, simple substitution and LFSR based cryptosystems.

### *A. Vigenere cipher cryptosystem*

Blaise de vignere proposed a polyalphabetic cipher system in sixteenth century which was named after him as vignere cipher. The encryption and decryption of the text is based on the vignere square or vigenere table or tabula recta shown in Figure 1.

ISBN No. 978-1-4799-3914-5/14/$31.00 ©2014 IEEE



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

Figure 1: Vignere square

The basic principle of encrypting a plaintext $P = p_1, p_2, p_3\ldots$, knowing the key $K{=}k_1, k_2\ldots$ to get the cipher text $C = c_1, c_2\ldots$ is $C_\mathrm{i} = (P_i + K_i)$ mod 26. The key elements are repeated if the size of the key is less than the size of the plain text. To get the plain text, from the cipher text, knowing the key, operation is performed is $P_\mathrm{i} = (C_i - K_i)$ mod 26. The complexity to break a vigenere cipher is $O(26^m)$, where $m$ is the key length. An example of enciphering the text with vigenere cipher is illustrated below

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Plain text | M | Y | C | O | U | N | T | R | Y |
| Key | S | E | C | R | E | T | S | E | C |
| Cipher text | E | C | E | F | Y | G | L | V | A |

### B. Simple Substitution

Simple substitution is a monoalphabetic substitution cipher, in which there is a fixed structure to replace the alphabets of the plain text to form the cipher text. The following example shows the way to form the cipher text

Let the key to encrypt the text is:

QWERTYUIOPASDFGHJKLMNBVCZX

This means that alphabet A in the entire plain text is replaced by alphabet Q (i.e. A →Q). Similarly, B →W, C →E etc. Time complexity of breaking simple substitution cipher by brute force is of order $O$ (26!).

### C. Linear feedback Shift Register based cryptosystem (LFSR)

LFSR is a shift register that performs an exclusive-OR (XOR) to some of its outputs together to form a feedback path. LFSRs are also used for pseudo random generation of bit patterns when provided with a suitable seed. An $n$ stage LFSR is maximum length if some initial states will results a sequence that repeats every $2^n$ - 1 bit. The pseudo random series generated act as the key and is used to encrypt the text. Cipher text is obtained by first converting plain English text to a binary string and then XORing it with corresponding bits of key. A primitive polynomial of degree 4, is taken to form the cipher text as shown below in table I

TABLE I: LFSR BASED ENCRYPTION

| Plain text | T | H | I | S | Y | R |
|---|---|---|---|---|---|---|
| Binary plain text | 10100 | 01000 | 01001 | 10011 | 11001 | 10010 |
| Key | 11110 | 10110 | 01000 | 11110 | 10110 | 01000 |
| Binary Cipher text | 01010 | 11110 | 00001 | 01101 | 01111 | 11010 |
| Cipher text | J | N | A | M | O | Z |

## III. NATURE INSPIRED META-HEURISTIC TECHNIQUES

To find an optimum solution of a problem, the first task is to formulate the problem correctly. Once an optimization problem is correctly formulated, the optimum solution can be found using appropriate mathematical techniques. For stochastic algorithms, two types of optimization techniques are well known: heuristic and meta-heuristic. Heuristic means 'to find'. Good solutions of a hard optimization problem can be found, but getting an optimal solution is not a guarantee. There are chances that heuristic algorithms might trap to local optimum solutions. To avoid being trapped in local optimum solutions and to get global best optimum solution, advancements to heuristic algorithms are made and are called meta- heuristic algorithms. Meta heuristic techniques are thus an improvement over traditional heuristic techniques and perform better compared to them. Thus, meta-heuristic algorithms intend to be suitable for global optimization. Some of the nature inspired meta-heuristic techniques are swarm intelligence and evolutionary techniques. Author studied two prominent optimisation techniques namely particle swarm optimization (PSO) and genetic algorithm (GA) for analysing above mentioned cryptosystems.

### A. Particle Swarm Optimization

PSO was pioneered by Kennedy and Eberhart. In 1995, they developed an optimization technique [8] based on intelligent movement of swarms. According to this algorithm, number of particles or agents flies around a search space for the best solution. The flowchart for PSO algorithm is shown in figure 2. Every particle updates its position and velocity according to the position and velocity of the best solution (*gbest*) and its own experience (*pbest*). Position and velocity of particles are updated as per the following equations

... (1)

... (2)

Here,
$v(i,t)$ : velocity of $i^{th}$ particle at iteration $t$
$wf$ : weighting function
$C_1$: Self confidence
$C_2$: swarm confidence
$r_1$ and $r_2$ : random numbers ($0 < r_1, r_2 < 1$)
$x(i, t)$ : current position of $i^{th}$ particle at iteration $t$
$pbest(i)$ : personal best of $i^{th}$ particle
$gbest$ : global best position of any particle.

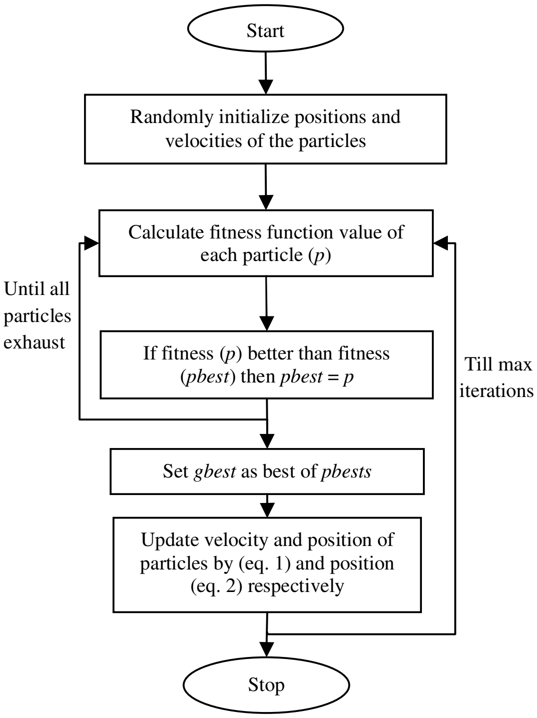

Start
Randomly initialize positions and velocities of the particles
Calculate fitness function value of each particle (p)
Until all particles exhaust
If fitness (p) better than fitness (pbest) then pbest = p
Till max iterations
Set gbest as best of pbests
Update velocity and position of particles by (eq. 1) and position (eq. 2) respectively
Stop

Figure 2: Flow chart of Particle Swarm Optimization

### B. Genetic Algorithm

Originally developed by John Holland in 1975, Genetic Algorithm (GA) [9] is a type of optimization technique which is driven by the two factors of 'Natural selection' and 'Genetic Inheritance' proposed by Darwin while observing evolution in nature. A new better population is evolved from the previous population using operators like selection, recombination, mutation etc. Selection of parent chromosomes is generally done by roulette wheel method. Modification of chromosomes includes crossover and mutation.

Crossover can be one point crossover or two point crossover, depending on the type of the problem. Crossover and mutation operations are performed to form offsprings from the parent chromosome for the next generation. To form the new population, fitness function value of each of the chromosome is evaluated. Among the offsprings and parent chromosomes, only those chromosomes are selected which have higher fitness function value. Figure 3 shows the basic evolutionary cycle.

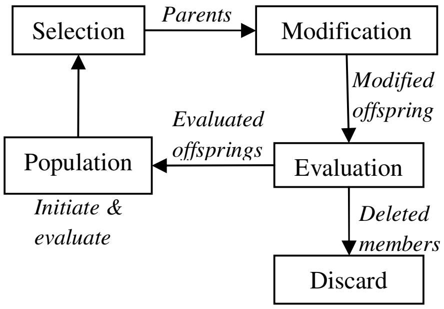


Figure 3: The evolutionary cycle

The Pseudo code for Genetic Algorithm is described below

```
procedure GA
{
        irt = 0;
        initialize random population P(t);
        find fitness value of P(t);
        until (done)
        {
         itr = itr + 1;
        select parent chromosomes ($p_1$, $p_2$);
        recombine parent chromosomes
        mutate ($p_1$, $p_2$) to form offsprings ($O_1$, $O_2$);
        find fitness of ($O_1$,$O_2$);
        select best among ($p_1$, $p_2$, $O_1$, $O_2$);            }
}
```

For each of the meta heuristic techniques used to analyze the cryptosystem, fitness function is defined on the basis of English language characteristics i.e. frequencies of the monograms and high frequent bigrams according to the following equation

$$\text{Fitness} = \alpha \times \sum_{i=1}^{26} |SM(i) - OM(i)| + \beta \times \sum_{i=1}^{25} |SB(i) - OB(i)| \text{ ... (3)}$$

Where,
$SM$ ($i$) is standard frequency of $i^{th}$ monogram in normal English.
$OM$ ($i$) is observed frequency of $i^{th}$ monogram in decrypted text.
$SB$ ($i$) is standard frequency of $i^{th}$ bigram in normal English.
$OB$ ($i$) is observed frequency of $i^{th}$ bigram in decrypted text.

## IV. OBSERVATIONS AND RESULTS

The analysis of different cryptosystems viz. Vigenere cipher cryptosystem, Simple substitution cryptosystem and Linear Feedback Shift Register based cryptosystem were analyzed by different meta-heuristic techniques in

MATLAB. Vigenere cipher was analyzed both with PSO and GA while Simple substitution cipher and LFSR based cryptosystem were analyzed with Genetic Algorithm. A plain text of 5000 characters was taken. Out of that randomly a text of 600 characters was chosen and the cipher text was made with each of the cryptosystem. The results are shown in table II, table III and table IV. For analysis of vigenere cipher, simple substitution cipher and LFSR based cryptosystem; mutation rate selected was 0.2, 0.1 and 0.02 respectively and cross over rate selected for these cryptosystems was 0.8, 0.9 and 0.8 respectively. Fitness function parameters α and β were selected as 0.23 and 0.77. All the three cryptosystems were analyzed for 600 iterations.

TABLE II: ANALYSIS OF VIGENERE CIPHER

| Vigenere Cipher Cryptosystem | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Min correct key characters recovered** | | | | | **Max correct key characters recovered** | | | | |
| Key size | 5 | 10 | 15 | 20 | 25 | 5 | 10 | 15 | 20 | 25 |
| GA | 4 | 8 | 12 | 17 | 19 | 5 | 10 | 14 | 18 | 22 |
| PSO | 4 | 7 | 8 | 11 | 13 | 5 | 10 | 13 | 15 | 17 |

TABLE III: ANALYSIS OF SIMPLE SUBSTITUTION

| Simple Substitution Cryptosystem | | |
|---|---|---|
| | **Min. correct key characters recovered** | **Max. correct key characters recovered** |
| Key size | 26 | 26 |
| GA | 17 | 20 |

TABLE IV: ANALYSIS OF LFSR BASED CRYPTOSTEM

| LFSR based cryptosystem | | | |
|---|---|---|---|
| **Degree of polynomial** | **Total no. of bits in the key** | **Min. no. of correct key bits recovered** | **Max. no. of correct key bits recovered** |
| 13 | $2^{13}$-1 | $2^{10}$-1 | $2^{12}$-1 |
| 17 | $2^{17}$-1 | $2^{13}$-1 | $2^{15}$-1 |
| 19 | $2^{19}$-1 | $2^{16}$-1 | $2^{17}$-1 |
| 23 | $2^{23}$-1 | $2^{18}$-1 | $2^{20}$-1 |

## V. CONCLUSION

In this paper different cryptosystems using nature inspired meta-heuristic techniques like particle swarm optimization and genetic algorithm were analyzed in MATLAB. Vigenere and LFSR based cryptosystems were analyzed for different key sizes. From the results of vigenere cipher, we can conclude that genetic algorithm proves to be a better technique than PSO.

## REFERENCES

[1] M. E. Dalkilic and C. Gungor, "An interactive cryptanalysis algorithm for the Vigenere Cipher," Advances in Information Systems, Springer, pp. 341-351, 2000.

[2] T. Purusothaman, V. Gopalakrishnan, S. Arumugam, V. Palanisamy, S. Balraja, G. Parvathavarthini, and G. ManiKandan, "Cryptanalysis of vigenere cipher using genetic algorithm and dictionary analysis," presented at the Proceedings of the IASTED International Conference, 2009.

[3] A. Bhateja, A. K. Bhateja, and S. Kumar, "Cryptanalysis of Vigenere Cipher using Particle Swarm Optimization with Markov chain random walk," International Journal on Computer Science and Engineering, vol. 5(5), pp. 422-429, 2013.

[4] M. F. Uddin and A. M. Youssef, "Cryptanalysis of simple substitution ciphers using particle swarm optimization," presented at the Evolutionary Computation,CEC, IEEE Congress on. IEEE, 2006.

[5] M. Heydari, G. L. Shabgahi, and M. M. Heydari, "Cryptanalysis of Transposition Ciphers with Long Key Lengths Using an Improved Genetic Algorithm," World Applied Sciences Journal vol. 21(8), 2013.

[6] R. Vimalathithan and M. L. Valarmathi, "Cryptanalysis of simplified-DES using computational intelligence," presented at the WSEAS Transactions on Computers, 2011.

[7] M. Faheem, S. Alam, and M. U. Bokhari, "An Analysis of Linear Feedback Shift Registers in Stream Ciphers," International Journal of Computer Applications, vol. 46, 2012.

[8] J. Kennedy and R. Eberhart, "Particle swarm optimization," Proceedings of IEEE International Conference on Neural Networks, vol. 4, pp. 1942-1948, 1995.

[9] J. Holland, "Genetic algorithms," Scientific american, vol. 267(1), pp. 66-72, 1992.