ISSN No. 0976-5697

Volume 3, No. 3, May-June 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Breaking of Simplified Data Encryption Standard Using Evolutionary Algorithm

Lavkush Sharma*
Department of Computer Science & Engineering
Faculty of Engineering & Technology, R.B.S College
Agra, India
lavkush07@yahoo.com

Nidhi Sharma
Department of Computer Application
G.I.C.T.S
Gwalior, India
nidhi.gicts@gmail.com

Abstract: Cryptanalysis of ciphertext, plaintext by using evolutionary algorithm has gained so much interest in recent years. In this paper we used a Genetic algorithm with improved crossover operator (Ring Crossover) for cryptanalysis of S-DES and then results are compared with Simple Genetic Algorithm. Here the cipher text attack is considered and several keys are generated in the different run of the genetic algorithm on the basis of their cost function value which depends upon frequency of the letters. The results on the S-DES indicates that, this is a promising method and can be adopted to handle other complex block ciphers like DES, AES

Keywords: Cryptanalysis, Ciphertext attack, Simplified Data Encryption Standard, Genetic algorithm, Key search space

I. INTRODUCTION

A cipher is a secret way of writing in which plaintext is encrypted into ciphertext by using a key. Those who know the key can easily decrypt the ciphertext back into the plaintext. Cryptanalysis is the study of breaking ciphers that is finding the key or converting the ciphertext into the plaintext without knowing the key. Many cryptographic systems have a finite key space and, hence, are vulnerable to an exhaustive key search attack. Yet, these systems remain secure from such an attack because the size of the key space is such that the time and resources for a search are not available. Optimization techniques have got a significant importance in determining efficient solutions of different complex problems. One such problem is to break S-DES. This paper considers cryptanalysis of S-DES. In the brute force attack, the attacker tries each and every possible key on the part of cipher text until desired plaintext is obtained. A brute force approach may take so much time to guess the real key which is used to generate a cipher text. On the other hand optimization technique can be used for the same purpose. Genetic algorithm is an evolutionary algorithm that works well and takes less time to break cipher as compared to Brute force attack.

The remaining paper is organized as follows: Section II discusses the earlier works done in this field. Section III presents overview of S-DES and Section IV gives the overview of Genetic Algorithm. Experimental results are discussed in Section V. Conclusion are presented in section VI at last References are given in section VII.

II. RELATED WORK

In the last few years, so many papers have been published in the field of cryptanalysis. In the recent years G Poonam [1, 2] presented the use of memetic algorithm and genetic algorithm to break a simplified data encryption standard algorithm. Nalini [3] used efficient heuristics to attack S-DES. In 2006 Nalini used GA, Tabu search and Simulated Annealing techniques to break S-DES. R.Spillman etc. showed that Knapsack cipher [4] and substitution ciphers [5] could be attacked using genetic algorithm. Clark [6] also presented important analysis on how different optimization techniques can be used in the field of cryptanalysis. Matusi [7] showed the first experimental cryptanalysis of DES using a linear cryptanalysis technique. Vimalathithan [8] used GA to attack Simplified-DES. Lavkush Sharma [9] also used Genetic Algorithm to break S-DES.

In this paper, a Genetic Algorithm with improved parameters is used to break S-DES.A population of keys is generated and their fitness is calculated by using efficient fitness function. At the end, we will find the key in less time.

III. S-DES

In this section we will provide the overview of S-DES Algorithm. Simplified DES, developed by Professor Edward Schaefer of Santa Clara University is an educational rather than a secure encryption algorithm. The S-DES [10, 11] encryption algorithm takes an 8-bit block of plaintext and a 10-bit key as input and produces an 8-bit block of ciphertext as output. The S-DES decryption algorithm takes an 8-bit block of ciphertext and the same 10-bit key used to produce that ciphertext as input and produces the original 8-bit block of plaintext. The encryption algorithm involves five functions: an initial permutation (IP); a complex function labeled f_K, which involves both permutation and substitution operations and depends on a key input; a simple permutation function that switches (SW) the two halves of the data; the function f_K again; and finally a permutation function that is the inverse of the initial permutation (IP^{-1}).

The function f_K takes as input not only the data passing through the encryption algorithm, but also an 8-bit key. S-DES uses a 10-bit key from which two 8-bit sub keys are generated. In this, the key is first subjected to a permutation (P10). Then a shift operation is performed. The output of the shift operation then passes through a permutation function that produces an 8-bit output (P8) for the first sub key (K_1). The output of the shift operation also feeds into another shift and another instance of P8 to produce the second sub key (K_2).

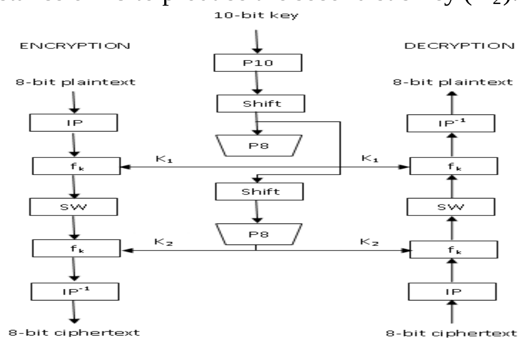

Figure 1: Simplified Data Encryption Algorithm

A. *Initial and Final Permutations*

The input to the algorithm is an 8-bit block of plaintext, which we first permute using the IP function IP= [2 6 3 1 4 8 5 7].This retains all 8-bits of the plaintext but mixes them up. At the end of the algorithm, the inverse permutation is applied; the inverse permutation is done by applying, IP^{-1} = [4 1 3 5 7 2 8 6] where we have $IP^{-1}(IP(X))$ =X.

B. *The Function f_K:*

The function f_k, which is the complex component of S-DES, consists of a combination of permutation and substitution functions. The functions are given as follows.

Let L, R be the left 4-bits and right 4-bits of the input, then,

$$f_K (L, R) = (L \text{ XOR } f(R, key), R)$$

Where XOR is the exclusive-OR operation and key is a sub -key. Computation of f(R, key) is done as follows.

a. Apply expansion/permutation E/P= [4 1 2 3 2 3 4 1] to input 4-bits.
b. Add the 8-bit key (XOR).
c. Pass the left 4-bits through S-Box S0 and the right 4-bits through S-Box S1.
d. Apply permutation P4 = [2 4 3 1].

The two S-boxes are defined as follows:

$$S0 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} 0\ 1\ 2\ 3 \\ \begin{bmatrix} 1 & 0 & 3 & 2 \\ 3 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 \\ 3 & 1 & 3 & 2 \end{bmatrix} \end{array}$$

$$S1 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} 0\ 1\ 2\ 3 \\ \begin{bmatrix} 0 & 1 & 2 & 3 \\ 2 & 0 & 1 & 3 \\ 3 & 0 & 1 & 0 \\ 2 & 1 & 0 & 3 \end{bmatrix} \end{array}$$

Figure 2.Working of S-Box

The S-boxes operate as follows:

The first and fourth input bits are treated as 2-bit numbers that specify a row of the S-box and the second and third input bits specify a column of the S-box.

The entry in that row and column in base 2 is the 2-bit output.

C. *The Switch Function:*

The function f_K only alters the leftmost 4 bits of the input. The switch function (SW) interchanges the left and right 4 bits so that the second instance of f_K operates on a different 4 bits. In this second instance, the E/P, S0, S1, and P4 functions are the same. The key input is K2.

IV. GENETIC ALGORITHM

The genetic algorithm [12, 13] is a search algorithm based on the natural selection and on "survival of the fittest". The main idea is that in order for a population of individuals to adapt to some environment, it should behave like a natural system. This means that survival and reproduction of an individual is promoted by the elimination of useless traits and by rewarding useful behavior. The genetic algorithm belongs to the family of evolutionary algorithms. An evolutionary algorithm maintains a population of solutions for the problem at hand. The population is then evolved by the iterative application of a set of stochastic operators. The simplest form of genetic algorithm involves three types of operators: selection, crossover and mutation.

A selection operator is applied first.

a. ***Selection:*** This selection operator selects chromosomes in the population for reproduction. The better the chromosome, the more times it is likely to be selected to reproduce.
b. ***Crossover***: Crossover selects genes from parent chromosomes and creates a new offspring. The simplest way to do this is to choose randomly some crossover point and everything before this point is copied from the first parent and then is copied from the second parent, everything after a crossover point
c. ***Mutation:*** After a crossover, mutation is performed. This is to prevent falling all solutions in population into a local optimum of solved problem. Mutation changes randomly the new offspring. In binary GA we can switch a few randomly chosen bits from 1 to 0 or from 0 to 1.

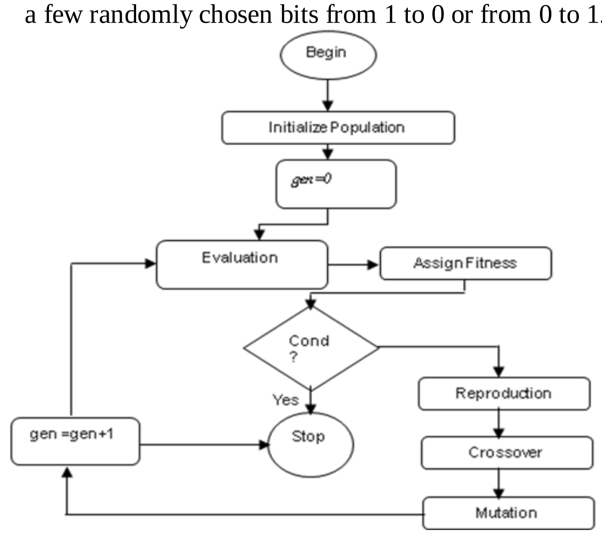

Figure 3: Flow chart of Genetic Algorithm

In this paper, we are using Ring crossover operator [14].In ring crossover two parents such as parent1 and parent2 are considered for the crossover process, and then combined in the form of ring, Later; a random cutting point is decided in any point of ring. The children are created with a random number generated in any point of ring according to the length of the combined two parental chromosomes. With reference to the cutting point, while one of the children is created in the clockwise direction, the other one is created in direction of the anti-clockwise. Then swapping and reversing process is performed in the Ring Crossover operator.

The procedure to carry out the cryptanalysis using GA with ring crossover and by simple G.A in order to break the key is as follows

a) Input: ciphertext, and the language statistics.
b) Randomly generate an initial pool of solutions (keys).
c) Calculate the fitness value of each of the solutions in the pool using equation (1).
d) Create a new population by repeating following steps until the new population is complete
   i. Select parent (keys) from a current population according to their fitness value (the better fitness, the bigger chance to be selected). Here Tournament selection is used.
   ii. With a crossover probability cross over the parents to form new offspring (children's). In our genetic algorithm.
   iii. For each of the children, perform a mutation operation with some mutation probability to generate new children.
   iv. Place new children in the new population
e) Use new generated population for a further run of the algorithm
f) If the end condition is satisfied, stop, and return the best solution in current population

A. Cost Function:

Equation (1) is a general fitness function used to determine the suitability of a assumed key (k). Here, A denotes the language alphabet (i.e., for English, [A... Z, _], where _ represents the space symbol), K and D denote known language statistics and decrypted message statistics, respectively, and the u, b, and t denote the unigram, digram and trigram statistics respectively; α, β and γ are the weights assigning different weights to each of the three statistics where $\alpha + \beta + \gamma = 1$. In view of the computational complexity of trigram, only unigram and digram statistics are used.

$$\begin{aligned} \mathbf{C}^{\mathbf{K}} &= \alpha\, \Sigma(i\ \varepsilon\ \tilde{A})\ |\ K(i)^{u} - D(i)^{u}\ | \\ &+ \beta\, \Sigma(i, j\ \varepsilon\ \tilde{A})\ |\ K(i, j)^{b} - D(i, j)^{b}| \\ &+ \gamma\, \Sigma(i,j,k\varepsilon\tilde{A})|\ K(i,j,k)^{t} - D(i,j,k)^{t}| \qquad (1) \end{aligned}$$

V. RESULTS AND DISCUSSION

Our objective in this paper is to compare the results obtained from simple Genetic Algorithm with the Genetic Algorithms with improved parameters. The experiments were conducted on Core 2 Duo system. There are a variety of cost functions used by other researchers in the past. The most common cost function uses gram statistics. Some use a large amount of grams while others only use a few. Equation 1 is a general formula used to determine the suitability of a proposed key. A number of experiments have been carried out by giving different inputs and applying genetic algorithm with ring crossover and by simple G.A for breaking Simplified Data Encryption Standard. The results are shown in table 1.the choice of the Genetic Operators play a vital role in GA and are described below:

GA Parameters with Ring Crossover:

The following are the GA parameters used during the experimentation:

Population Size: 100
Selection: Tournament Selection operator
Crossover: Ring Crossover
Crossover: .85
Mutation: .02
No. of Generation: 50

GA Parameters with single point Crossover:
The following are the GA parameters used during the experimentation:
Population Size: 100
Selection: Tournament Selection operator
Crossover single point Crossover
Crossover: .85
Mutation: .02
No. of Generation: 50

Table 1: Comparison of Genetic Algorithm with Ring Crossover and Simple Genetic Algorithm.

S. No	Amount of Cipher Text	No. of bits matched using GA with R.C	No. of bits matched using Simple G.A	Time Taken by GA with Ring crossover (M)	Time Taken by Simple G.A (M)
1.	200	5	5	4.7	4.9
2.	400	4	3	2.1	2.5
3.	600	7	6	1.9	2.2
4.	800	8	7	3.1	3.4
5.	1000	9	7	2.6	2.8
6.	1200	9	8	2.1	2.5

From the above table, it is found that both GA with ring crossover works better than simple G.A in terms of time taken as well as obtaining number of key bits.

VI. CONCLUSION

In this paper, we have used a Genetic algorithm with Ring crossover and other operators for the cryptanalysis of Simplified Data Encryption Standard. We found that Genetic Algorithm with ring crossover is better than Simple G.A algorithm for cryptanalysis of S-DES. Although S-DES is a simple encryption algorithm, GA with Ring Crossover method can be adopted to handle other complex block ciphers like DES and AES.

VII. REFERENCES

[1]. G Poonam, Memetic Algorithm Attack on Simplified Data Encryption Standard algorithm, proceeding of International Conference on Data Management, February 2008, pg 1097-1108

[2]. Garg Poonam, Genetic algorithm Attack on Simplified Data Encryption Standard Algorithm, International journal Research in Computing Science, ISSN1870-4069, pp. 139-149, 2006.

[3]. Nalini, Cryptanalysis of S-DES via Optimization heuristics, International Journal of Computer Sciences and network security, vol 6, No 1B, pp. 240-246, Jan 2006.

[4]. Spillman, R.: Cryptanalysis of Knapsack Ciphers Using Genetic Algorithms.Cryptologia XVII (4), 367–377 (1993)

[5]. Spillman, R., Janssen, M., Nelson, B., Kepner, M.: Use of A Genetic Algorithm in the Cryptanalysis of simple substitution Ciphers. Cryptologia XVII (1), 187–201 (1993)

[6]. Clark A and Dawson Ed, “Optimization Heuristics for the Automated Cryptanalysis of Classical Ciphers”, Journal of Combinatorial Mathematics and Combinatorial Computing, Vol.28, pp. 63-86, 1998.

[7]. M. Matsui, Linear cryptanalysis method for DES cipher, Lecture. Notes Computer Science. 765 (1994) 386–397.

[8]. Vimalathithan.R, M.L.Valarmathi, “Cryptanalysis of S-DES Using Genetic Algorithm”, International Journal of Recent Trends in Engineering, Vol 2, No.4, pp.76-79, November 2009

[9]. Lavkush Sharma, B.K Pathak, R.G Sharma “Breaking of S-DES Using GA” G.J.C.S.T, Volume 12 Issue 5 Version 1.0 March, pp 55-59, 2012.

[10]. William Stallings, Cryptography and Network Security Principles and Practices, Third Edition, Pearson Education Inc. 2003.

[11]. Schaefer E, “A Simplified Data Encryption Standard Algorithm”, Cryptologia, Vol .20, No.1, pp. 77-84, 1996.

[12]. Kalyanmoy Deb, Multi-objective Optimization using Evolutionary Algorithms, John Wiley and Sons, 2001.

[13]. A,Michalewiez and N. Attia."Evolutionary optimization of constrained problems." InProc. 3rd annu. Conf. on Evolutionary Programming. 1994. pp 98-108

[14]. Yılmaz Kaya, Murat Uyar, Ramazan Tekdn,” A Novel Crossover Operator for Genetic Algorithms: Ring Crossover”.

[15]. Davis,L. “Handbook of Genetic Algorithm”,Van Nostrand Reinhold, New York,1991

[16]. D. E. Goldberg,"Genetic algorithms in search. optimization and Machine Learning. Reading. M.A. addison - Wesley.1989.

[17]. Z. Michalewicz. "Genetic algorithms+ Data structures = Evolution programs 3rd ed. New York. springer,1996.

[18]. N.Koblitz, “A Course on number theory and Cryptography”, Springer-Verlag New York,Inc., 1994.

[19]. Alfred J. Menezes. Menezes, Alfred J. Handbook of Applied Cryptography, CRC, 1997.

[20]. R. Toemeh, S. Arumugam, Breaking Transposition Cipher with Genetic Algorithhm Electronics and Electrical Engineering,ISSN 1392 – 1215 2007. No. 7(79)

[21]. N.Koblitz, “A Course on number theory and cryptography”, Springer-Verlag New York, Inc., 1994