Text Encryption Using Genetic Algorithm

[1] Dr.Alsadig Mohammed Abdallah; [2] Dr.Mandour Mohammed Ibrahim

[1] Information Technology Department, College of Computer and Information Sciences,
AL-Imam Muhammad ibn Saud Islamic University, Kingdom of Saudi Arabia

[2] Information Technology Department, College of Computer and Information Sciences,
AL-Imam Muhammad ibn Saud Islamic University, Kingdom of Saudi Arabia

Abstract - At this time with the evolution of technology and the spread of computer networks to the great expansion in the online networks around the world, making all transactions based on the use of computers, such as storing important information on the various storage media, changing text messages, images and multimedia across the network, electronic banking, e-commerce and electronic payment operations using (credit cards, visa card, MasterCard), making the information vulnerable to espionage and hack by vandals and hackers, so it must find ways to protect this information because of the great importance. There are many different text encryption methods to convert text from plain text to cipher text. In this research we will introduce the method that using Genetic Algorithms (GA) which is used to produce encryption method based on (Crossover and Mutation). The proposed encryption method in this study has been tested on some texts and we have got excellent results.

Keywords - *Genetic Algorithm, Crossover, Mutation, Cryptography, Hackers*

1. Introduction

Safe and secure transfer of data is a stipulation in all domains ranging from two people's conversation to the country's defense and military. Data and transmission over the network is at threat from the hackers and the attackers spread everywhere. Data can be stolen, changed, corrupted or lost. Thus network security is indispensable. The major role of network security lies in avoiding the tampering of data transmitted across the network [1].

Process of encrypting data in-order to restrict the access of data to only authorized person is referred to as cryptography. Cryptography offers efficient solution to protect sensitive information including personal data security, internet security, military communication security, etc [2].

Cryptography is the art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text[2] , so Cryptography is essential for protecting information as the importance of security is increasing day by day with the advent of online transaction processing and e commerce [3].

2. Background

The encryption methods are important as in today's world to provide security to data is on everyone's priority list therefore cryptography schemes are needed for preventing unauthorized access to data, such as genetic algorithm. [4]

Studies mentioned that Genetic algorithm for cryptography is better than typical algorithms for cryptography. Genetic algorithms contain process/operations such as mutation, crossover and selection. They are used to optimize and search problems by generating high-quality solutions.

The method discussed below generate different key by pseudorandom function for each block of data according to which the crossover is applied and in order to provide more security mutation is also applied.

Since the proposed method is being applied to the binary data therefore it can be used to encrypt different type of file.

3. Types of Cryptography

3.1 Symmetric key cryptography:

In this type, the sender and the receiver use the same key for encryption and decryption and hence the key is shared between them.

Example: DES (Data Encryption Standard) and AES (Advanced Encryption Standard).

3.2 Asymmetric key cryptography:

The technique is also known as public key cryptography. In this type, the sender and the receiver use different keys for encryption and decryption and hence the key is not shared between them [5]. Example: RSA (Ron Rivest, Adi Shamir and Leonard Adleman).

There are two classical techniques for encrypting data, which are used singly or combination in virtually every cryptographic algorithm. Substitution involves the systematic replacement of bytes in the data by a cipher byte according to some algorithm. Being relatively easy to automate both mechanically and electronically, substitution has been very widely used in both government commercial cryptographic systems. The broad category of cryptographic algorithm is transposition, in which the relative order of bytes make up the data is permuted according to some rule. This easy to automate, although the advent of microprocessors led to its being incorporated into a number of modern cryptographic systems [6]. genetic algorithms are a class of stochastic search algorithms based on biological evolution. Given a clearly defined problem to be solved and a binary string representation for candidate solutions [7].

The relative merits of crossover, mutation, and other genetic operators have long been debated in the literature of genetic algorithms. The traditional view is that crossover is primarily responsible for improvements in fitness, and that mutation serves a secondary role of reintroducing alleles that have been lost from the population. This view is consistent with the notion that evolutionary progress is best made by combining building blocks (or schemata) from high-fitness individuals [8].

But the roles of building blocks and of crossover have become increasingly controversial in recent years. Several researchers have presented new theoretical arguments and empirical results showing that mutation can be more useful than was previously thought (for example, [Shaffer and Eshelman 1991. The mutation/crossover debate has produced a variety of insights about the nature of genetic algorithms, and there is more yet to be discovered.

In the adaptive genetic algorithm (AGA), the probabilities of crossover and mutation, pc and p m, are varied depending on the fitness values of the solutions. High-fitness solutions are `protected', while solutions with subaverage fitnesses are totally disrupted. By using adaptively varying pc and p, we also provide a solution to the problem of deciding the optimal values of pc and pm, i.e., pc and pm need not be specified at all. The AGA is compared with previous approaches for adapting operator probabilities in genetic algorithms. The Schema theorem is derived for the AGA, and the working of the AGA is analyzed [8]. the crossover operator randomly chooses a crossover point where two parent chromosomes 'break', and then exchanges the chromosome parts after that point. As a result, two new offspring are created. For example, the chromosomes X6 and X2 could be crossed over after the second gene in each to produce the two offspring. Offspring are created as exact copies of each parent. For example, the parent chromosomes X2 and X5 may not cross over. Instead, they create the offspring that are their exact copies. Mutation, which is rare in nature, represents a change in the gene. It may lead to a significant improvement in fitness, but more often has rather harmful results. So why use mutation at all? Holland introduced mutation as a background operator (Holland, 1975). Its role is to provide a guarantee that the search algorithm is not trapped on a local optimum. The sequence of selection and crossover operations may stagnate at any homogeneous set of solutions [9]. first we need to represent the Text as binary with affixed length , after that we choose a pair of row information to apply GA Operations .

4. A GA applies the following major steps

Step 1: Represent the problem variable domain as a chromosome of a fixed length, choose the size of a chromosome population N, the crossover probability pc and the mutation probability pm.

Step 2: Define a fitness function to measure the performance, or fitness, of an individual chromosome in the problem domain. The fitness function establishes the basis for selecting chromosomes that will be mated during reproduction.

Step 3: Randomly generate an initial population of chromosomes of size N: x1; x2; . . . ; xN

Step 4: Calculate the fitness of each individual chromosome:f(x1),f(x2),........f(xN)

Step 5: Select a pair of chromosomes for mating from the current population.

Parent chromosomes are selected with a probability related to their fitness. Highly fit chromosomes have a higher probability of being selected for mating than less fit chromosomes.

Step 6: Create a pair of offspring chromosomes by applying the genetic operators – crossover and mutation.

Step 7: Place the created offspring chromosomes in the new population.

Step 8: Repeat Step 5 until the size of the new chromosome population becomes equal to the size of the initial population, N.

Step 9: Replace the initial (parent) chromosome population with the new (offspring) population.

Step 10: Go to Step 4, and repeat the process until the termination criterion is satisfied. As we see, a GA represents an iterative process. Each iteration is called a generation. A typical number of generations for a simple GA can range from 50

to over 500 [5]. The entire set of generations is called a run. At the end of a run, we expect to find one or more highly fit chromosomes.
Generally, a Genetic Algorithm consists of three basic operations.

- ❖ Selection
- ❖ Crossover
- ❖ Mutation

The first step consists of searching individuals for Reproduction [9].

5. The Proposed Encryption Method

In this paper we have try to Prepare a method to data encryption based on symmetric key. The randomness involved in crossover and mutation is exploited in generation of a one time symmetric key. A permutation factor is randomly generated and applied to the successive blocks of text to make the algorithm more unpredictable for the intruder.

Step 1: Extract two 8-byte blocks from the text file to be encrypted. Let the two blocks be represented by

b0	**b1**	**b2**	**b3**	**b4**	**b5**	**b6**	**b7**

c0	**c1**	**c2**	**c3**	**c4**	**c5**	**c6**	**c7**

where each bi and ci is a character in a file.

Step 2: Perform crossover operation. Generate two random numbers in the range 0-7. Let the two random numbers generated be 2 and 5.

Perform the crossover between two crossover points generated above. ⇣

b0	**b1**	**b2**	**b3**	**b4**	**b5**	**b6**	**b7**

c0	**c1**	**c2**	**c3**	**c4**	**c5**	**c6**	**c7**

The blocks after performing crossover operation are

b0	**b1**	**c2**	**c3**	**c4**	**b5**	**b6**	**b7**

c0	**c1**	**b2**	**b3**	**b4**	**c5**	**c6**	**c7**

Step 3: Perform mutation operation. Generate two random numbers in the range 0 to 7. Let the two random numbers generated be 1 and 7.

Hence, Mutation Point1 = 1

Mutation Point2 = 7

Perform mutation operation on two blocks obtained in Step 1.

b0	**128-b1**	**c2**	**c3**	**c4**	**b5**	**b6**	**128-b7**

c0	**128-c1**	**b2**	**b3**	**b4**	**c5**	**c6**	**128-c7**

Step 4: Generate the permutation factor randomly in the range 1-4 Let the permutation factor be 3.

Step 5: Generate a random key based on crossover points, mutation points and crossover factor generated above.

Hence the symmetric key in an octal form is

2	**5**	**1**	**7**	**2**

Each octal digit in a symmetric key can be represented using 3 bits. Hence a symmetric key in a binary format is given by

In the above example, c = m = 2.

Hence Key Length = 15

Hence, Crossover Point1 = 2

Crossover Point1 = 5

0	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**1**	**1**	**1**	**1**	**0**	**1**	**0**

Hence the length of the key is 15 bits in our case which depends on the number of crossover points and mutation points. The length of the symmetric key can be computed using a general formula

Key Length = 3 (c + m + 1)

Where c-> No. of Crossover points

m-> No. of Mutation points

The same operation is repeated for the next set of blocks after permuting the digits to the left of the permutation factor of a symmetric key by a permutation factor.

The new key generated is

1	7	2	5	2

Applying the same operation on the next two blocks, we get,

Step 1: Extract next two 8-byte blocks from the text file to be encrypted. Let the two blocks be represented by

d0	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**d7**

f0	**f1**	**f2**	**f3**	**f4**	**f5**	**f6**	**f7**

Step 2: Perform crossover operation. New crossover points are,

Crossover Point1 = 1

Crossover Point1 = 7

d0	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**d7**

f0	**f1**	**f2**	**f33**	**f4**	**f5**	**f6**	**f7**

d0	**f1**	**f2**	**f3**	**f4**	**f5**	**f6**	**d7**

f0	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**f7**

Step 3 : Perform mutation operation. New mutation points are,

Mutation Point1 = 2

Mutation Point2 = 5

d0	**f1**	**128-f2**	**f3**	**f4**	**128-f5**	**f6**	**d7**

f0	**d1**	**128-d2**	**d3**	**d4**	**128-d5**	**d6**	**f7**

6. Conclusion

A cryptography algorithm for encryption and decryption of data which uses the operations of genetic algorithm. Encryption of binary data is successful and we have satisfied our goals. Thus, we conclude that applying operations of Genetic Algorithm to provide security to the data in a file is possible.

References

1-(Sania Jawaid, Anam Saiyeda, Naba Suroor . Selection of Fittest Key Using Genetic Algorithm and Autocorrelation in Cryptography, Science and Education publishing

2- Amritha Thekkumbadan Veetil, An Encryption Technique Using Genetic Operators, INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 4, ISSUE 07, JULY 2015

3- Aarti Soni, Suyash Agrawal,Using Genetic Algorithm for Symmetric key
Generation in Image Encryption.

4- P Srikanth; Abhinav Mehta; Neha Yadav; Sahil Singh; Shubham Singhal, Encryption and Decryption Using Genetic Algorithm Operations and Pseudorandom Number

5- Mohammed Alhusainy ,image encryption using Genetic Algorithms, information technology journal,2006

6- A Guide to Intelligent Systems,Michael Negnevitsky -Second Edition,2005

7- A Comparison of Crossover and Mutation in Genetic Programming,Sean Luke And Lee Specto

8-Adaptive probabilities of crossover and mutation in genetic algorithms,M. Srinivas Dept. of Comput. IEEE).

9- Symmetric Key Encryption using Genetic Algorithm,Dr. Poornima G. Naik,Mr. Girish R. Naik ,2014).