To Design a Genetic Algorithm for Cryptography to Enhance the Security

Dr. Dilbag Singh
Associate Professor,
Department of Computer Science & Applications
Chaudhary Devi Lal University, Sirsa

Pooja Rani
Department of Computer Science & Applications
Chaudhary Devi Lal University, Sirsa

Dr. Rajesh Kumar
Assistant Professor
Govt. Daronacharya College, Sector-4, Gurgaon

Abstract-In today's age of information technology secure transmission of information is a big challenge. Traditional symmetric and asymmetric methods are not suitable when the needed level of security is high. Hash function based systems are although better than traditional methods but are still inadequate in many cases due to their algorithmic complexity as they need the invertible functions to generate hashes which are time consuming and complex. Digital signature is the new concept in field of cryptography but are much complex in implementation and increase server overhead. In present study, a genetic algorithm for cryptography has been proposed to find an optimized solution for a problem. In the proposed algorithm, the concept of genetic algorithm has been incorporated within cryptography algorithm to get an optimized solution and within minimum possible time.

Keywords: Cryptography, Genetic Algorithm, Encryption, Decryption.

I. Introduction

Security, integrity, non-repudiation, confidentiality, and authentication services are becoming the most challenging issues in today's age of information technology. For secure communication be mean of electronic cryptography is used. Cryptography is concerned with encoding and decoding of information to ensure the security. Cryptographic techniques are used to protect individual privacy as well as commercial secrets [2]. Genetic algorithms are based on the mechanics of natural selection and natural genetics. Genetic algorithm belongs to the family of evolutionary algorithms, along with genetic programming, evolution strategies, and evolutionary programming [19].Genetic algorithm considers an optimization problem as the environment where feasible solutions are the individuals living in that environment. Degree of adaptation of an individual to its environment is equivalent to the fitness function evaluated on a solution [23]. Genetic algorithm begins with a randomly generated set of individuals. Once the initial population has been created, the genetic algorithm enters in a loop. Result of each iteration gives a new population is produced by applying a certain number of stochastic operators to the previous population. Each such iteration output as new generation. An intermediate population of n-parent is created by applying selection operator. To produce these parents 'n' independent extractions of an individual from the old population is performed. Probability of each individual being extracted should be linearly proportional to the fitness of that individual [23].

Once the intermediate population of parents is produced, individual for the next generation will be created using reproduction operators. These operators can involve one or more parents. An operator that involves just one parent, simulating asexual reproduction, is called a mutation operator. When more than one parent is involved, sexual reproduction is simulated, and the operator is called recombination. The genetic algorithm uses two reproduction operators usually the crossover and mutation [3].

After crossover each individual has a small chance of mutation. The purpose of the mutation operator is to simulate the effect of transcription errors that can happen with a very low probability when an individual is mutated. A standard mutation operator for binary strings is a bit inversion means '0' would mutate into '1' and vice versa.

Cryptography and cryptanalysis could be considered to meet these criteria. However, cryptanalysis is not closely related to the typical genetic algorithm application areas [13].

II. Objectives of the Study

Study has been carried out with objectives:

1. To study the Cryptographic Algorithms and Genetic Algorithms.
2. To design a new Genetic Algorithm for Cryptography.
3. To analyze the experimental results.

III. Research Methods

Present research is experimental in nature as study has been carried out in simulated environment. In this research MATLAB 7.8.0 is used as simulation platform. Security of data sharing over web which has been studied in simulated environment. Present study is also analytical as facts have been analyzed to assess its performance and throughput of the proposed algorithm.

IV. Design and Analysis of Proposed Algorithm:

In present section, genetic algorithm of cryptography is proposed. This section also presents the data flow diagram for encryption and decryption of data transferred over the web. Genetic algorithms are all about crossover, mutation, selection operation and calculating the fitness number using some fitness function. Genetic algorithms are actually the optimization algorithms used in artificial intelligent systems. In present study, the concept of genetic algorithms has been used in cryptography so as to find an optimized solution within an optimized time.

4.1 Encryption Process

Encrypting process emulates the working of the crossover operator using pseudorandom sequence. Steps for the data encryption are as follows:

1. Generate the sequence of pseudorandom binary numbers using the random number generator.
2. Convert the binary pseudorandom sequence into decimal pseudorandom sequence ranging from 0 to 7 as Yn .
3. Read 16 consecutive bytes from the data file.
4. Initialize l=0
5. Initialize m=0
6. Modify the consecutive bytes using byte substitution, as per AES standard.
7. Take two consecutive bytes of the data stream as P1 and P2.
8. Perform crossover on two consecutive bytes of the data stream as Q1 and Q2 by using the number Yi.
9. Encrypt data as C1 and C2, where,

$Xi = Yi \oplus (Yi << 4)$

$Xi+1 = Yi+1 \oplus (Yi+1 << 4)$

$C1 = Q1 \& Xi$

$C2 = Q2 \& Xi+1$

10. l= l+2 and m= m+1 repeat steps 6 to 9 until l<= 16
11. Repeat steps 5 to 10 until m<=5
12. Again perform the byte substitution over the encrypted 16 consecutive bytes.
13. Repeat steps 3 to 12 until end of the data.

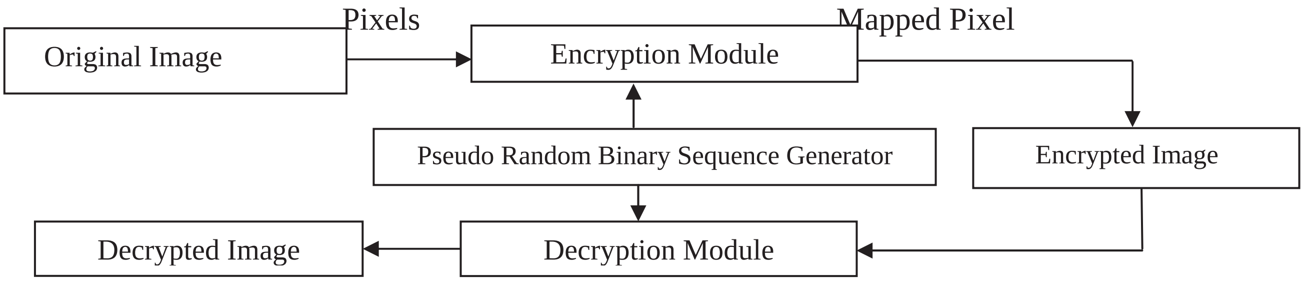

Figure 1 Block Diagram of the Proposed Method

Firstly the original image dimensions will enter into the encryption module as shown in figure 1. Then the proposed algorithm will run on that image or dimensions to create an encrypted image dimensions using a key generated by pseudorandom binary sequence generator. Then the data is sent over the communication channel in that encrypted format so that no third party other than sender and receiver can misuse it. After that at the receiver end the decryption module runs over that encrypted format of image using the same key generated by pseudo random binary sequence generator and find the original image. Thus the algorithm provides a better idea for secure data transmission and communication over web so that no unauthorized person can misuse anyone's personal data.

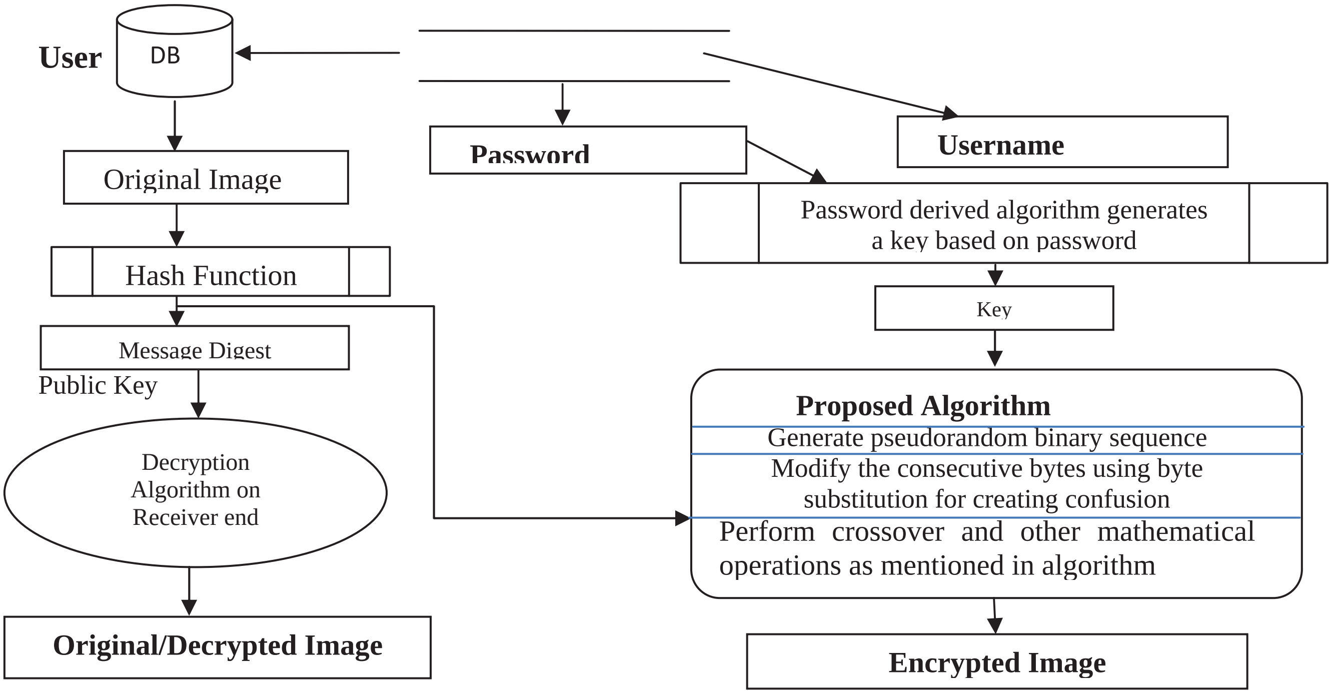
User
DB
Original Image
Hash Function
Message Digest
Public Key
Decryption Algorithm on Receiver end
Original/Decrypted Image
Password
Username
Password derived algorithm generates a key based on password
Key
Proposed Algorithm
Generate pseudorandom binary sequence
Modify the consecutive bytes using byte substitution for creating confusion
Perform crossover and other mathematical operations as mentioned in algorithm
Encrypted Image

Figure 2 DFD for Secure Data Transfer over Web

4.2 Decryption Process:

Steps for decryption are just reversal of the encryption. First, generate the pseudorandom sequence and then use this pseudorandom sequence and crossover operator to decrypt the data.

4.3 Analysis of Proposed Algorithm:

Proposed algorithm has been analyzed using two different types of analysis. Analysis has been made by drawing the graphs of encrypted and original data dimensions and then comparing them. Secondly the proposed algorithm has been analyzed using the time analysis to find the estimated time that the algorithm will take to encrypt or decrypt the data.

4.4 Analysis of the Encryption Process

In present analysis various images and some textual form of data have been used for analysis purpose. Firstly, original image has been plotted and then corresponding encrypted dimensions have been shown using the histograms. After that, comparative graph have been plotted to differentiate them. On the horizontal axis of graph the individual pixel color values has been shown and on vertical axis the frequency of occurrence of same color value in that image has been taken. In textual form data, the horizontal axis shows per byte value of the data and vertical axis shows the same.

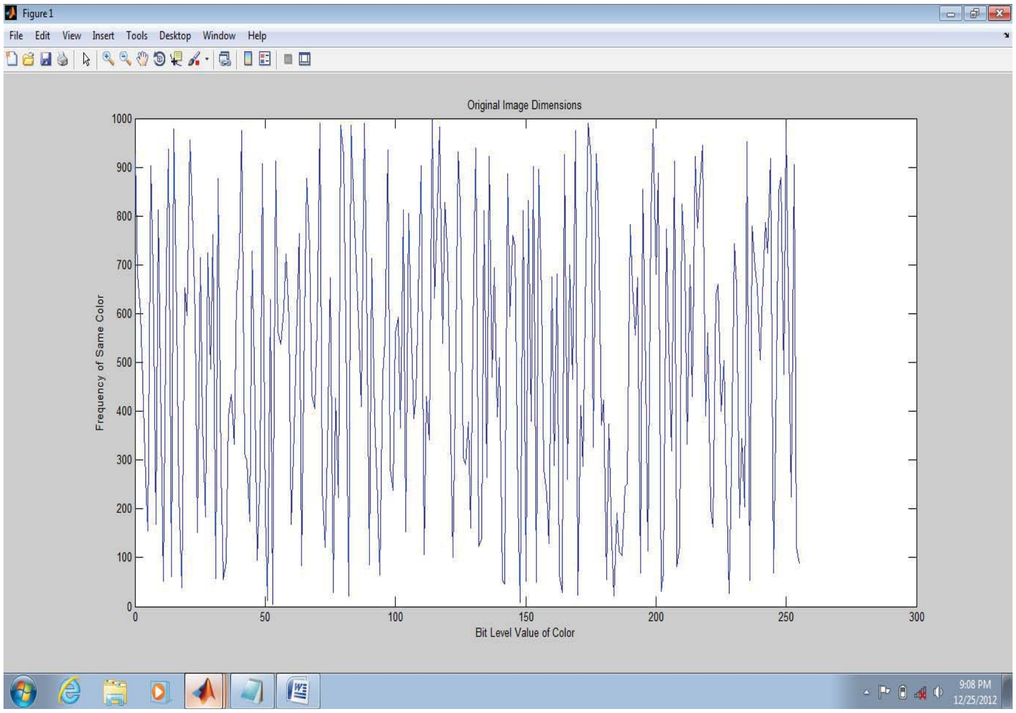

Figure 3 Dimensions of Original Image

Figure 3 shows the dimensions of the original image which is to be encrypted. The image dimensions are calculated by dividing the image colors on basis of bit level gray scale which are shown along horizontal axis. The '0' value shows darkest means absolute black color and value 255 indicates the brightest color. The vertical axis on the graph shows how much the image is found at any particular brightness level.

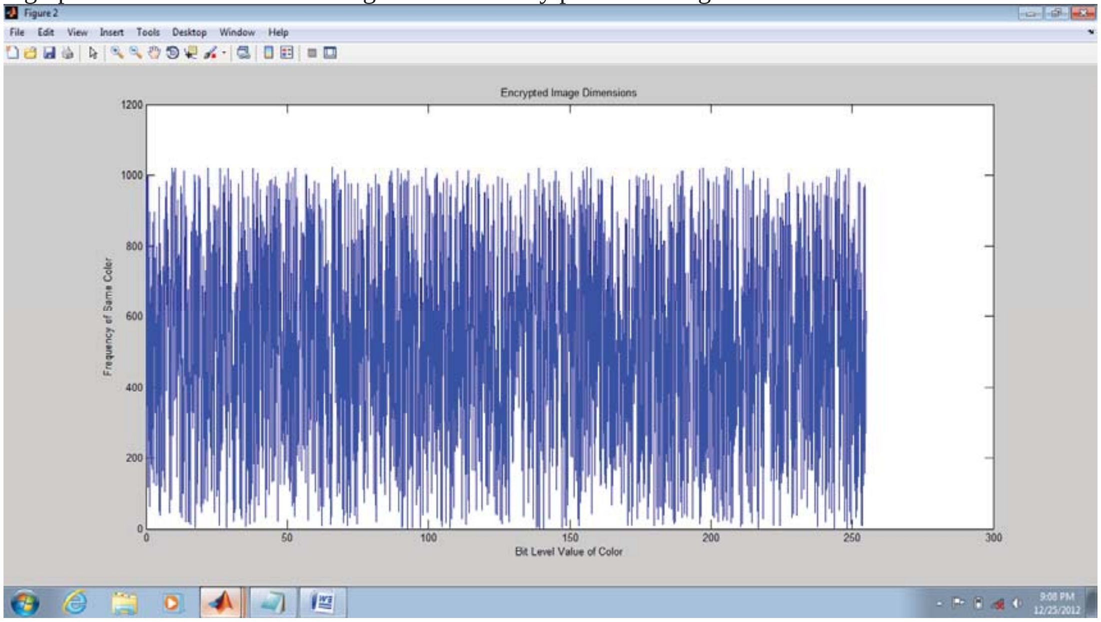

Figure 4 Dimensions of Corresponding Encrypted Image

Figure 4 shows the image color dimensions according to bit level gray scale after the decryption of image in 3. The image dimensions are calculated on basis of bit level gray scale where value '0' indicates absolute black and 255 indicates absolute white and is taken along horizontal axis. The vertical axis on the graph shows how much the image is found at any particular brightness level means how much pixels of the image have that particular brightness level.

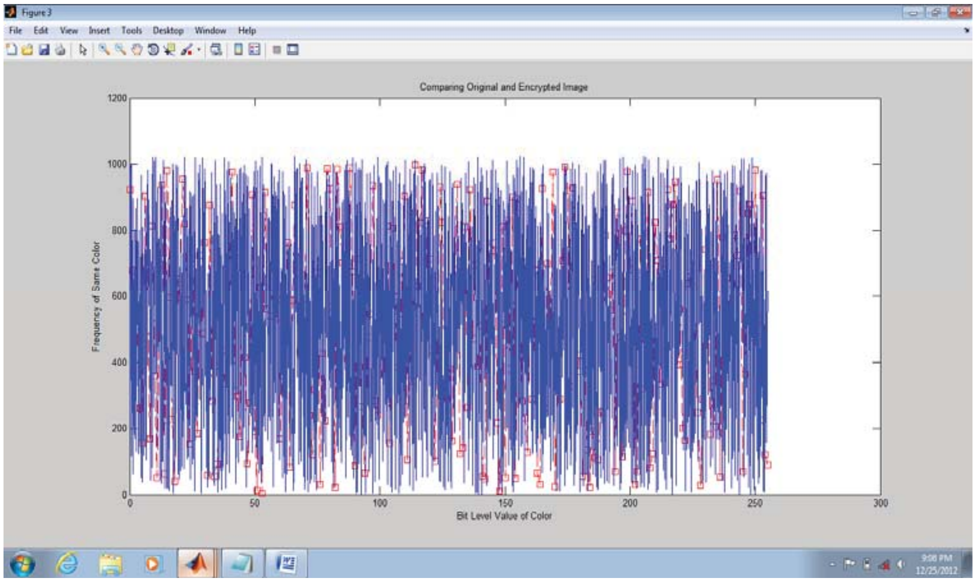

Figure 5 Dimensions of Original and Encrypted Image

The figure 5 shows a comparative graph between the original image dimensions and its corresponding image dimensions. The horizontal line on the graph have value scale from 0-255 which shows the brightness level of the pixel color and vertical line on the graph shows the frequency of an occurrence of a particular brightness level in an image. The dotted line is showing the original image dimensions and the solid line is showing the encrypted image dimensions. It can be easily seen from the graph that the encrypted image dimensions are totally different from that of original image.

It is clear from figures, the encrypted image is nearly uniform and quite different from the original image and hence it does not provide any clue to employ any type of statistical attack on the proposed image encryption. Without the knowledge of the pseudorandom sequence no one will be able to extract the message**.**

4.5 Time Analysis of the Encryption

The speed is an important factor for a good encryption algorithm. The encryption/decryption rate for several gray scale images of different size have been measured here. The algorithm speed has been measured for many different images of different sizes. The average time taken by the algorithm for different size of images is shown in table 1.

Image Size (in pixels)	Bits/Pixels	Average Encryption/Decryption time(in sec)
128 X 128	8	0.008-0.013
128 X 128	8	0.010-0.015
256 X 256	8	0.031-0.035
256 X 256	8	0.033-0.036
512 X 512	8	0.071-0.108

Table 1: Average ciphering speed of gray scale images

The images of different sizes have been taken for time analysis to measure the time that the algorithm takes to encrypt or decrypt an image. The images are defined as 8 bit per pixel. It is found here that the algorithm takes only a few milliseconds to encrypt an image of standard size. Thus, the algorithm is much efficient and provides a much better throughput. It provides a better solution within the least time.

V. CONCLUSION

In present study a cryptographic algorithm has been designed using the concept of genetic algorithms. This algorithm enhances the quality, efficiency and effectiveness of the algorithm being used for the cryptography. From the experimental results, it can be easily seen that the present algorithm has achieved the objective set in the present study. Statistical analysis shows that the dimensions of original data and the encrypted are totally different. Also the histogram of the encrypted image is nearly uniform and is quite different from the original image; hence, it does not provide any clue to employ any type of statistical attack on the proposed image encryption technique. Time analysis results show the throughput rate of the proposed method and it is found that the algorithm gives a much better and acceptable throughput rate. The conventional encryption techniques are not a feasible solution in terms of their throughput rate. Hence, a secure encryption technique with high throughput has been designed for real time data transmission. From experimental results it is clear that the encryption results are completely random and very sensitive to the parametric fluctuation that makes transferring of secret information highly safe and highly reliable.

REFERENCES

[1] Data Encryption Standard (DES)," National Bureau Standards FIPS Publication" pages 46-50, 1977.

[2] Douglas, R. Stinson, "Cryptography- Theory and Practice", CRC Press, 1995.

[3] Tragha A., Omary F., Kriouile A., "Genetic Algorithms Inspired Cryptography", A.M.S.E Association for the Advancement of Modeling & Simulation Techniques in Enterprises, Series D : Computer Science and Statistics, November 2007.

[4] Husainy M., "Image Encryption using Genetic Algorithm", Information Technology Journal pages 516-519, 2006.

[5] Goldberg D.E., "Genetic algorithms in search optimization & Machine Learning", Addison-Wesley. Publishing Company, 1989.

[6] Clark a., "Modern Optimization Algorithms for Cryptanalysis", In Proceedings of the Second Australian and New Zealand Conference on Intelligent Information Systems, pages 258-262,1994.

[7] Clark A. & Dawson Ed., "A Parallel Genetic Algorithm for Cryptanalysis of the Polyalphabetic Substitution Cipher", pages 129-138,1997.

[8] Clark A., Dawson Ed. & Nieuwland H., "Cryptanalysis of Polyalphabetic Substitution Ciphers Using a Parallel Genetic Algorithm", In Proceedings of IEEE International Symposium on Information and its Applications, pages 17-20, 1996.

[9] Matthews R.A.J., "The use of Genetic Algorithms in Cryptanalysis", pages 187-201, 1993.

[10] Spillman R.,"Cryptanalysis of knapsack ciphers using Genetic Algorithms", pages367-377, 1993.

[11] Yaseen I. F. T., & V.Sahasrabuddhe H.," A genetic algorithm for the cryptanalysisof Chor-Rivest knapsack public key cryptosystem (PKC)", In Proceedings of Third International Conference on Computational Intelligence and Multimedia Applications, pages 81-85, 1999.

[12] http://www.hku.hk/bse/bbse3002/Research-Methodology.pdf

[13] http://www.garykessler.net/library/crypto.html

[14] http://www.obitko.com/tutorials/genetic-algorithms/crossover-mutation.php

[15] http://www.obitko.com/tutorials/genetic-algorithms/selection.php

[16] http://www.obitko.com/tutorials/genetic-algorithms/ga-basic-description.php

[17] http://www.luminous-landscape.com/tutorials/understanding-series/understanding- histograms.shtml

[18] http://www.garykessler.net/library/crypto.html

[19] Spillman R.,"Cryptanalysis of knapsack ciphers using Genetic Algorithms", pages 367-377, 1993.

[20] Deo Narsingh, "System Simulation with Digital Computer", PHI Learning Pvt. Ltd., 01-Aug-2004.

[21] www. Mathworks.com

[22] http://www.mathworks.in/videos/getting-started-with-matlab-68985.html

[23] Sivanandam S. N. & Deepa S. N. "Introduction to Genetic Algorithm", pages 31-80.