Implication of Genetic Algorithm in Cryptography to Enhance Security

Muhammad Irshad Nazeer, Ghulam Ali Mallah, Noor Ahmed Shaikh
Shah Abdul Latif University, Khairpur, Sindh, Pakistan

Rakhi Bhatra, Raheel Ahmed Memon, Muhammad Ismail Mangrio
Department of Computer Science Sukkur IBA University, Sukkur, Sindh Pakistan

***Abstract*—In today's age of information technology secure transmission of information is a big challenge. Symmetric and asymmetric cryptosystems are not appropriate for high level of security. Modern hash function based systems are better than traditional systems but the complex algorithms of generating invertible functions are very time consuming. In traditional systems data is being encrypted with the key but still there are possibilities of eavesdrop the key and altered text. Therefore, key must be strong and unpredictable, so a method has been proposed which take the advantage of theory of natural selection. Genetic Algorithms are used to solve many problems by modeling simplified genetic processes and are considered as a class of optimization algorithms. By using Genetic Algorithm the strength of the key is improved that ultimately make the whole algorithm good enough. In the proposed method, data is encrypted by a number of steps. First, a key is generated through random number generator and by applying genetic operations. Next, data is diffused by genetic operators and then logical operators are performed between the diffused data and the key to encrypt the data. Finally, a comparative study has been carried out between our proposed method and two other cryptographic algorithms. It has been observed that the proposed algorithm has better results in terms of the key strength but is less computational efficient than other two.**

***Keywords*—Secure transmission; symmetric cryptosystems; invertible functions; genetic algorithms; efficient encryption**

I. Introduction

Recently, secure data transmission over network has become a vital and critical issue due to increased demand of digital media transmission and unauthorized access of important data [1]. Cryptography uses mathematical techniques for information security, data integrity, confidentiality, non-repudiation and authentication. Cryptography is based on concepts of Encryption and Decryption [2]. When data is sent from sender to receiver, the data is converted to some unreadable form called encryption of data and at receiver side data is again converted to its original form called decryption of data. Both encryption and decryption process require the key. For protection of valuable information from unlawful imitation, eavesdropper's attack and modification, different types of cryptographic algorithms are designed. There are two major types of such algorithms: symmetric cryptography [3] and asymmetric cryptography [4]. In asymmetric key cryptography two different keys are used, one for encryption called public key and one for decryption called private key. Only one same key is used in symmetric scheme.

The applications of both schemes differ due to efficiency of scheme; symmetric scheme is mostly used for encryption of data due to its high performance while asymmetric is often used for digital signature and distribution of key. Moreover, no any symmetrical ciphering technique such as AES, DES, Advanced AES, and IDEA has taken any benefit from most recent advances in information processing technology. Various kinds of modern data encryption techniques [2], [5] are found in the literature. Genetic Algorithms (GAs) [6] are among such techniques.

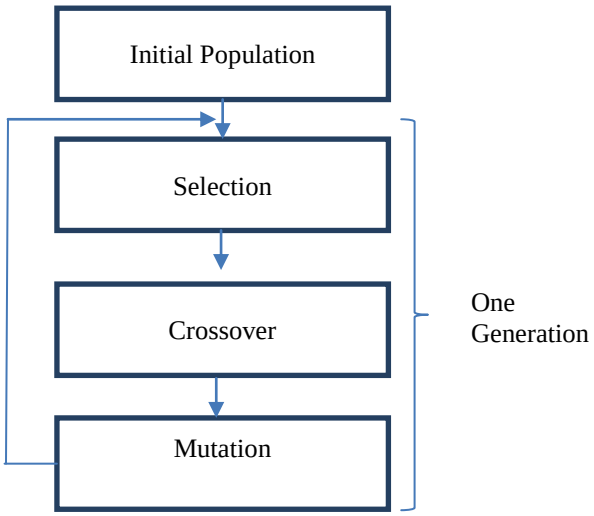

Fig. 1. Flow chart of genetic algorithm.

GA is kind of adaptive search algorithms which make use of the mechanics of natural selection and genetics. GA is part of Evolutionary Algorithms; which are used to solve optimization problems with the help of biological mechanism like selection, crossover and mutation [7]. Fig. 1 shows the process of solving optimization problems using Genetic Algorithms.

The key idea of GA is to imitate the randomness of the nature where natural selection process and behavior of natural system make population of individuals able to adapt the surrounding. We can say the survival and reproduction of the individuals is supported by exclusion of less fitted individuals. The population is generated in such a way that the individual with the highest fitness value is most likely to be replicated and

This work has been sponsored by the Higher Education Commission of Pakistan through Indigenous Ph D fellowships program.

unfitted individual is discarded based on threshold set by an iterative application of set of stochastic genetic operators [8].

Genetic Algorithm performs following operations to transform the population to new population based on fitness value.

A. Crossover

Crossover is a genetic operator which joins two chromosomes to form a new chromosome. The newly generated child chromosome is composed of chromosomes from each parent.

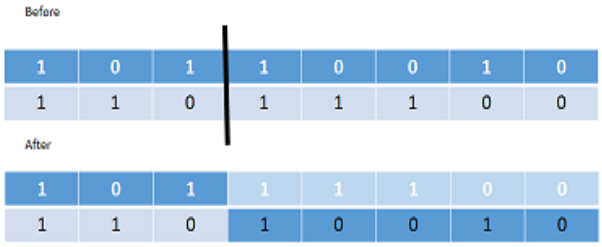

Fig. 2. Single point crossover.

Crossover is classified as single point, two point and uniform crossover. In Single Point only one crossover point is selected to generate new child (Fig. 2).

In Two Point crossover two crossover points are selected to generate new child (Fig. 3). In Uniform crossover bits are selected uniformly from each (Fig. 4) [8].

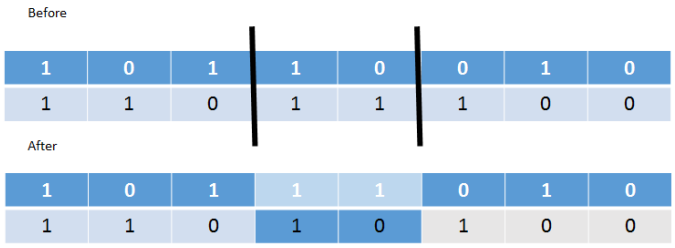

Fig. 3. Two point crossover.

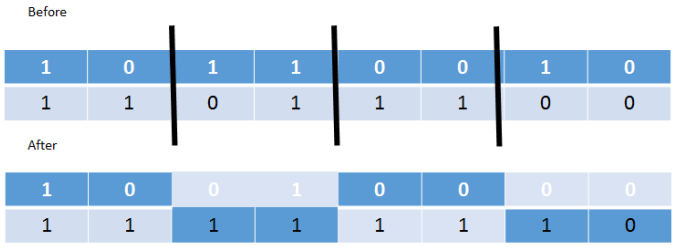

Fig. 4. Uniform crossover.

B. Mutation

In mutation after crossover at least one bit in each chromosome is changed (Fig. 5) [9]. This is performed to reflect the effect of surrounding in natural genetic process. There are two major types of Mutation i-e Flipping of Bits and Boundary Mutation. In Flipping of Bits one or more bits are converted into 0 to 1 or 1 to 0. In Boundary Mutation randomly upper or lower block in swapped in chromosome [9].

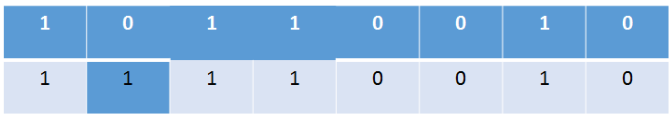

Fig. 5. Mutation.

C. Selection

In selection, chromosomes are chosen from the population for generation of new population. The selection is based on fitness value, higher the value more is the chances to be selected. Selection is classified as Roulette-wheel Selection, Tournament Selection; Truncation Selection [8].

D. Fitness Function

This is very important function of Genetic Algorithm because good fitness functions are useful for exploring the search space efficiently and bad fitness functions are confined to local optimum solution. Fitness Function can be categorized as Constant fitness function and Mutable fitness function [9].

Key Selection in cryptography is kind of selection problem and when we consider selection then; the key with highest fitness and randomness is selected. The applications of Genetic Algorithm are also in search heuristic problems, which make the GA a reliable algorithm for key generation and data encryption.

The opinion, which, we are following in this paper, is that if the quality (randomness) of the pseudorandom numbers generated for keys is good then the keys generated will always be non-repeating and purely random and ultimately increase the security and strength of keys.

Our major research question for this research is how can we get benefit of computational intelligence especially the genetic Algorithm to optimize the Cryptosystems? If so what will be the performance of such kind of solutions?

II. Literature Review

With the help of GA most of the research has been done by different researchers in the area of data encryption and key generation. Some of the work is defined in this section.

Jhingran et al. [7] conducted survey on applications of genetic algorithm in the field of cryptography.

Hassan et al. [10] have used the concept of encryption and decryption with the help of GA and RSA. First the key was generated with the help of GA and then generated key was used in RSA to encrypt the data. In this way the strong key was generated that was non-repeating too and this was not easy to break. This algorithm is better in terms of key strength than DES, AES, and RSA, etc. Sindhuja et al. [11] has given a symmetric key cryptosystem by applying GA. Key matrix and text matrix were added to create an additive matrix and then substitution cipher was applied on additive matrix to create the intermediate cipher. Crossover and Mutation were then applied on intermediate cipher to encrypt the data. This method is simple and easy to implement.

Aarti Soni et al. [12] proposed a new algorithm in which pseudorandom number generator was used to generate the key. The random number generator used the current time of

computer for random numbers. Then genetic operations were performed on random numbers. Finally selected key was used in AES symmetric algorithm to encrypt the image. The benefits of this algorithm were increased efficiency, less computational time and irregularity of key. The same method of key generation was also followed by Sania Jawed et al. [13] but in this, fitness value was calculated by applying Frequency and Gap test along with hams distance between the two binary keys. This algorithm was implemented in Java technology where 100 chromosomes, 0.5 mutation rate, 2.5 crossover rate were selected for the algorithm.

Narendra K. Pareek et al. [14] used the GA for encryption of gray scale images. The performance analysis of scheme revealed that the algorithm possesses the good statistical results, key sensitivity and can handle the plaintext attack, brute force attack, entropy attack and differential attack. Kirshna et al. [15] proposed cryptographic algorithm by using genetic function. In this algorithm substitution matrix and double point crossover was used to encrypt the data. This algorithm was implemented in Xilinx 13.2 version and verified using Spartan 3e kit. Almarimi et al. [1] dealt with security of electronic data over network. The proposed algorithm integrated the GA and pseudorandom sequence for encryption and decryption of data. Random sequence was obtained by using nonlinear shift register. Time and speed of algorithm was calculated for observing results.

Swati Mishra et al. [8] worked to generate a best fit key which could make code difficult to crack. Fitness of key was calculated by Pearson coefficient of autocorrelation. Two keys public and private were generated by using random number generator, crossover and then mutation. Finally Gap and Frequency tests were applied to select the best sample of key. The process was repeated until there was no best key. C++ programming was used to implement the algorithm and frequency was tested by chi-square test.

Ankit et al [9] generated the key for stream cipher with the help of natural selection process. The genetic operations were repeated until fitness value of any chromosome is less than threshold. Once completed the final selection of key was done through GA. Selected key was unique and non-repeating.

Kalaiselvi et al [16] discussed the need of adaptive and dynamic cryptographic algorithm to reduce computational cost and enhance security. In this paper two enhanced AES cryptosystems were proposed by using GA in SP boxes. AES was modified to accommodate the nonlinear Neural Network in SP network. This scheme ensured the increased security against timing attacks and reduction of computational time.

Subhajit et al [17] encrypted an image by using genetic algorithm. Then statistical test were performed to visualize the feasibility of solution.

The work done by researchers has impressive results but each research work has used some existing cryptographic algorithm in combination with genetic operators. Our motivation is to create novel cryptographic algorithm with the help of Genetic operations, which is easy to implement and secure in terms of key strength and attack time.

III. Proposed Algorithm

The proposed algorithm is named as Genetic Crypto and is divided into three major steps, i.e. Key Generation, Data diffusion and Data Encryption (Fig. 6).

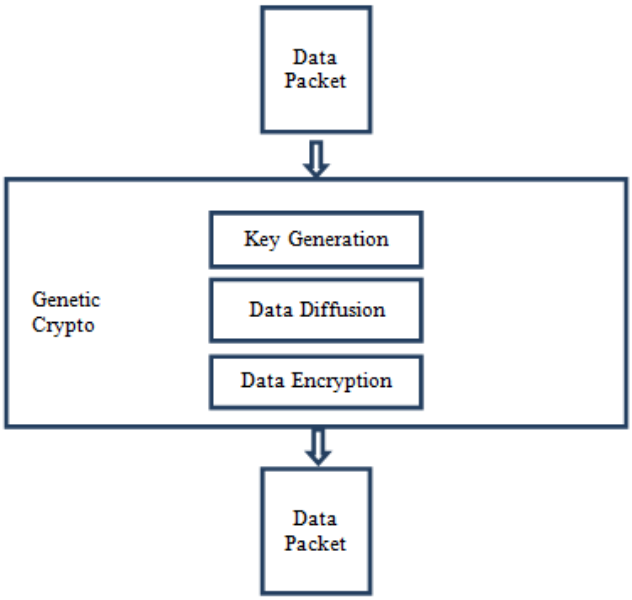

Fig. 6. Genetic crypto flow diagram.

The genetic operators are used in both key generation and data diffusion. Initial population is generated through random number generator. For simplicity one point crossover and bit filliping techniques are used for Crossover and Mutation respectively. Fitness value of key is calculated through Shannon Entropy because entropy is one of important feature of randomness. This algorithm is implemented in C# programing language, .net framework 4.5 in Visual Studio 2012. The interface and example result is shown in Fig. 7.

E. Key Generation: Key will be of 80-128 bits.

1) Sixteen random characters are generated with the help of random number generator from A-Z.

2) Each randomly generated character is converted to binary format (8 bits).

3) The result is stored in 2D array data structure.

4) Sixteen prime random numbers are generated from 0-100.

5) Each randomly generated number is converted to binary format (8 bits).

6) The result is stored in 2D array data structure.

7) Eight random numbers from 1 to 7 are generated for crossover points.

8) The numbers are stored in array data structure.

9) One point crossover is performed by taking one parent from array of random prime number and one parent from array of random characters. The crossover point is identified from the array of random numbers generated in step 1.8.

*10)*Step 1.9 will be repeated until there is parent left for crossover.

*11)*For Mutation, bit flipping mutation is used in which first and last bit of each chromosome is inverted; means 0 will be converted to 1 and vice versa.

*12)*Step 1.11 will be repeated for all the child chromosomes.

*13)*After Mutation, Fitness function of each chromosome is calculated through Shannon Entropy.

*14)*Chromosomes with the Shannon Entropy of greater than 0.95 will be merged and selected as key. If there is no any.

*15)*Chromosome with entropy greater than 0.95 then the whole process will be repeated again until there is no best fit key.

F. Diffusion of Original Text

1) Data is converted to binary format.

2) Binary data will be segmented into blocks. Each block size is 8 bits and number of blocks (chromosomes) is size of data/8.

3) The result is stored in 2D array data structure.

4) Eight random numbers from 1 to 7 are generated for crossover points.

5) The numbers are stored in array data structure.

6) One point crossover is performed between adjacent parents in array of binary data. The crossover point is identified from the array of random numbers generated in step 2.5.

7) For Mutation, bit flipping mutation is used in which first and last bit of each chromosome is inverted.

G. Encryption

1) Length of key and length of data is calculated first. If any of them has fewer bits than the other, 0s will be appended from left to make the length of data and key equal.

2) Logical XOR operation will be performed between diffused data and key bit wise.

3) The resulting set of bits is encrypted data

Some of the limitations of our work are:

a) Randomness purely depends on the random number generator and it may be pseudo random number generation. It is just limited to 16 characters.

b) Length of key and data is subject to design consideration.

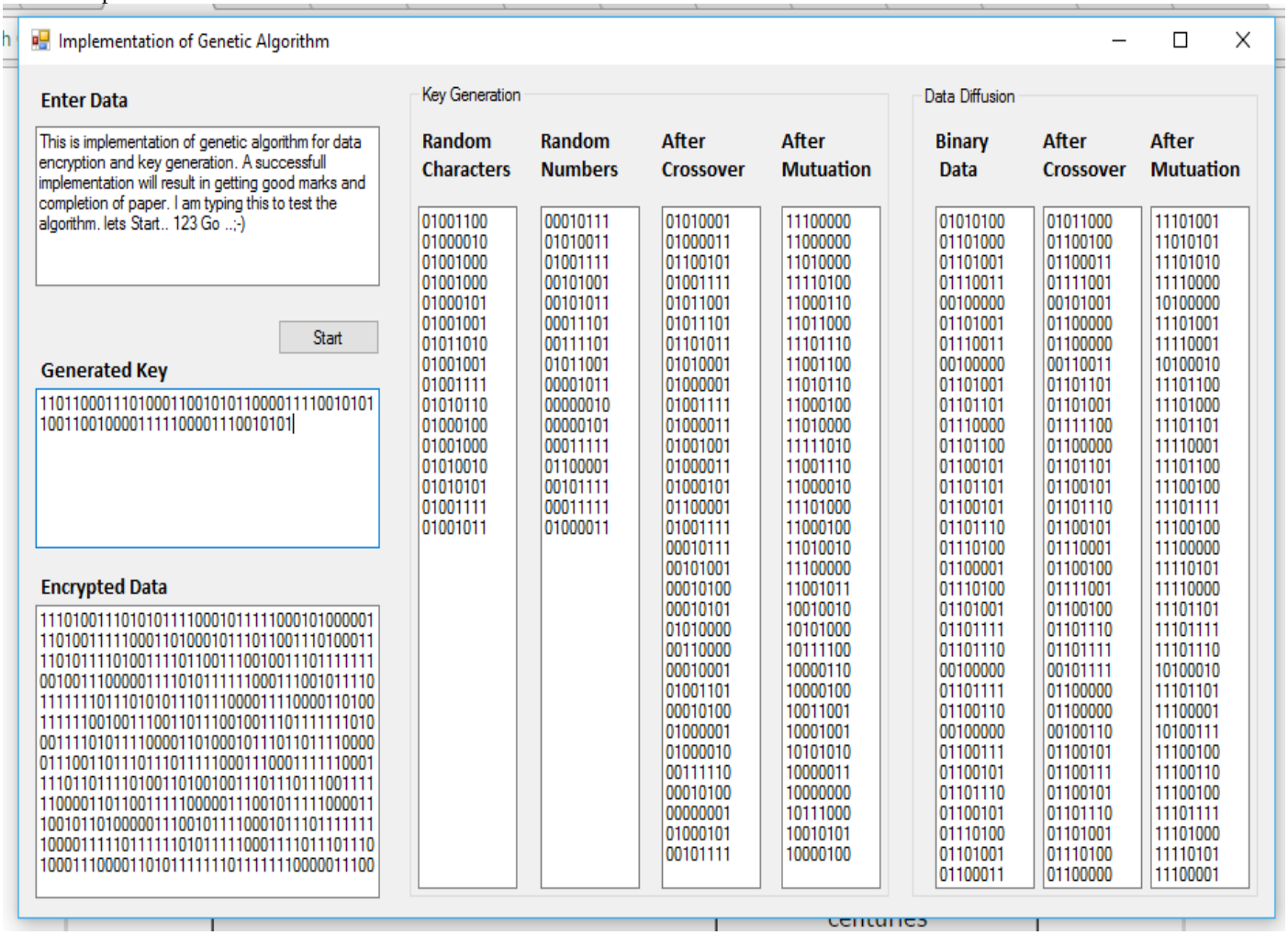

Fig. 7. Results of implementation of proposed algorithm.

IV. Results and Discussions

The proposed algorithm (Genetic Cipher) is compared with DES and AES symmetric key cryptosystems in terms of encryption, decryption time and key strength. The key strength is categorized by key search space size means how many alternative keys can be tried to break the cipher, Attack Scenario means how much time is required by eavesdropper to attack on data. The Encryption and decryption are calculated by implementing the algorithm and key strength is in terms of attack time is calculated with help of GRC [1] Interactive Brute Force key "Search Space" Calculator.

TABLE I. Comparison with Other Algorithms

	DES	AES	Genetic Cipher
Encryption Time	068907 mm	084440 mm	27069 mm
Key Search Space Size	$4.85 * 10^{28}$ Keys	$2.31 * 10^{57}$ Keys	$1.11 * 10^{120}$ Keys
Attack Time (1000 k/s)	15.41 thousand trillion days	7.34 hundred million trillion days	3.53 hundred billion trillion days

Table I shows that Encryption time of DES and AES is 068907mm and 084440 mm respectively while Encryption time of Genetic Cipher is 27069 mm, which is higher than both. The complex cryptographic algorithms with high provision of security are much better than simple algorithm with less security in cryptography. This point is evidenced by measure of key strength. In both categories key search space and attack time the Genetic Cipher requires much higher time to break than DES and AES.

To see performance improvement we consider the encryption time, size of the key search space and attack time. In Fig. 8, we plot the time taken by our algorithm and compare with the time taken by DES and AES. There is a significant improvement as encryption time is lesser, Search space is vast and Attack Time is much higher than AES and DES. In this graph we took log of the Search Space and Attach time in order to improve visibility of the plot.

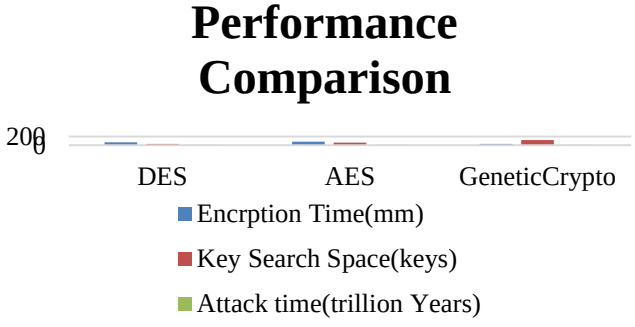

Fig. 8. Performance with respect to DES and AES.

V. Conclusions and Future Work

In this paper we have adopted a new way to encrypt the data i-e using GA. First a key of length between 80 and 128 is generated by applying genetic operations on randomly generated characters and prime numbers. Shannon Entropy is used to calculate the fitness value of each chromosome. After key generation, data is diffused again by applying crossover and mutation on data. At last key and diffused data are XORed for encryption. The result shows that although the proposed algorithm take little longer encryption time than DES and AES but the key strength is better than the other two compared algorithms.

In future we will prepare to improve this algorithm for multimedia encryption like images, video and audio. Efficiency in terms of time will be considered first. From the evaluation point of view, we will compare this genetic cipher with other cryptographic algorithms. Also, we can use more statistical techniques for evaluation of key randomness.

[1] https://www.grc.com/haystack.htm

References

[1] A. Almarimi, A. Kumar, I. Almerhag, and N. Elzoghbi, "A NEW APPROACH FOR DATA ENCRYPTION USING GENETIC Original Image Pseudorandom Binary Sequence Generator using GA and Decryption Decrypted Image," Computer (Long. Beach. Calif)., pp. 2–6, 2014.

[2] D. R. Stinson, Cryptography: Theory and Practice, vol. 30. 2005.

[3] J. Daemen and V. Rijmen, The Design of Rijndael: AES - The Advanced Encryption Standard. 2002.

[4] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," Commun. ACM, vol. 21, no. 2, pp. 120–126, 1978.

[5] W. M. H. Company, Modern Cryptography: Theory and Practice, vol. 170, no. 2. 2003.

[6] D. E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning. 1989.

[7] R. Jhingran and A. Prof, "A Study on Cryptography using Genetic Algorithm Vikas Thada Shivali Dhaka," Int. J. Comput. Appl., vol. 118, no. 20, pp. 975–8887, 2015.

[8] S. Mishra and S. Bali, "Public key cryptography using genetic algorithm."

[9] A. Kumar and K. Chatterjee, "An efficient stream cipher using Genetic Algorithm," 2016 Int. Conf. Wirel. Commun. Signal Process. Netw., pp. 2322–2326, 2016.

[10] A.-K. S. O. Hassan, A. F. Shalash, and N. F. Saudy, "MODIFICATIONS ON RSA CRYPTOSYSTEM USING GENETIC OPTIMIZATION," Int. J. Res. Rev. Appl. Sci., vol. 19, no. 2, p. 150, 2014.

[11] S. K and P. D. S, "A Symmetric Key Encryption Technique Using Genetic Algorithm." .

[12] A. Soni and S. Agrawal, "Using Genetic Algorithm for Symmetric key Generation in Image Encryption," Int. J. Adv. Res. Comput. Eng. Technol., vol. 1, no. 10, pp. 2278–1323, 2012.

[13] S. Jawaid and A. Jamal, "Article: Generating the Best Fit Key in Cryptography using Genetic Algorithm," Int. J. Comput. Appl., vol. 98, no. 20, pp. 33–39, Jul. 2014.

[14] N. K. Pareek and V. Patidar, "Medical image protection using genetic algorithm operations," Soft Comput., vol. 20, no. 2, pp. 763–772, 2014.

[15] G. M. K. and V. Lakshmi, "A Proposed Method for Cryptographic Technique by Using Genetic Function," Int. J. Emerg. Eng. Res. Technol., pp. 1–7, 2015.

[16] K. Kalaiselvi and A. Kumar, "Enhanced AES cryptosystem by using genetic algorithm and neural network in S-box," in 2016 IEEE International Conference on Current Trends in Advanced Computing, ICCTAC 2016, 2016.

[17] S. Das, S. N. Mandal, and N. Ghoshal, "Diffusion and Encryption of Digital Image Using Genetic Algorithm," in Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014, 2015, pp. 729–736.