International Journal of Latest Trends in Engineering and Technology
Special Issue SACAIM 2017, pp. 034-039
e-ISSN:2278-621X

DNA CRYPTOGRAPHY ALGORITHM USING GENETIC OPERATORS

Anushree Raj[1] & Dr. Rio G L D'Souza[2]

Abstract-The traditional cryptographic algorithm lacks high level of security. This paper intends to propose a DNA Cryptography Algorithm using Genetic operators based on existing works. Novel DNA cryptography algorithm is based upon a secure symmetric key generation and DNA encoding procedure. This paper is proposed based on various traditional DNA based encryption techniques. This technique focuses on secure data transmission in network by providing a strong encryption decryption mechanism. DCA binaries any plain text and follows steps like block construction, crossover, mutation and DNA sequencing to generate an encrypted file (cipher text) at the sender side. On other side, DCA performs decryption on receiver side where it follows steps like binaries the DNA sequence (cipher text), mutation, crossover, block construction, binary form of file and generate a decrypted file (original data). By using default key, the plain text is transformed to final cipher as a DNA base sequence and visa versa. The DCA undergoes two levels of security, firstly it uses a symmetric key which is implemented at encryption and decryption and secondly uses the concept of DNA encoding to ensure another level of data security.
Keywords: DNA cryptography, symmetric encryption, genetic operatons.

1. INTRODUCTION:

In this new era of secure data transmission over network, it is a major concern to protect data from the hands of hackers or against illegal access and modification. Lot of techniques and algorithms are identified and many solutions are proposed by researchers to ensure privacy on data transmission. The proposed work is based on the idea of bi-serial DNA encryption mechanism [1]. A brief introduction on cryptography, genetic operators and DNA technology, also various techniques and algorithms used for DNA cryptography is investigated.

Cryptography is one of the schemes to prevent unauthorised access to data. In the proposed paper the plaintext ASCII values is first converted to its binary form, a combination of 0s and 1s, blocks are generated upon which the genetic operators crossover and mutation is applied, which is then converted into a DNA sequence. DNA coding is based on four nucleotides for encrypting binary to DNA sequence such as adenine 'A', cytosine 'C', thymine 'T' and guanine 'G' [2]. The cipher text so formed is completely different from the original plain text, which is not easily detectable by unauthorized users.

Genetic algorithms contain bio inspired operators such as mutation, crossover and selection. They are used to optimize and search problems by generating high-quality solutions. In this paper crossover and mutation operation are used to support encryption decryption of plain text to convert a binary code into a DNA sequence which reduces the time complexity.

1.1 Cryptography

Cryptography is a technique used to make a communication unreadable to all except of the deliberate receiver(s). It ensures to send data in puzzled form so as to preserve privacy of the data.

There are two types of cryptographic schemes based on the key used:

A. Symmetric Cryptography: It's a technique where same key is used for encryption and decryption. The key is shared between both the communicating party's sender and the receiver. Symmetric key cryptography helps to achieve high performance. For e.g. AES, IDEA, DES, etc.

B. Asymmetric Key Cryptography: It's a technique where two different keys are used. Key for encryption is known as the public key, and private key for decryption. For e.g. RSA, Diffie - Hellman

1.2 Genetic algorithms

Genetic Algorithm (GA) is based on biological evolutionary theories and is often used to solve optimization problems. GA comprises of a set of individual elements (the population) and a set of biologically inspired operator. A simple GA uses following operators to transform a population into new population:

GA uses three main types of rules at each step to create the next generation from the current population:

Selection rules select the individuals, called parents that contribute to the population at the next generation.

Crossover rules combine two parents to form children for the next generation.

Mutation rules apply random changes to individual parents to form children

Here in our algorithm we have used only crossover and mutation operators.

[1] Department of Computer Science and Engineering, St Joseph Engineering Collage, Mangalore, Karnataka, India
[2] Department of Computer Science and Engineering, St Joseph Engineering Collage, Mangalore, Karnataka, India

1.3 DNA

DNA is a nucleic acid that contains the genetic instructions. The four bases found in DNA are adenine (A), cytosine (C), guanine (G) and thymine (T). A gene is a sequence of DNA that contains genetic information of all living organisms [3].

There are two ways to realize DNA technology - DNA coding and DNA cryptography.

DNA cryptography is a relatively new paradigm that has attracted great interest in the field of information security. It has parallel computing properties to perform the encryption and decryption. DNA cryptography stores huge amount of data in small volume with the combination of only these four letters A, C, G, and T. These bases form the structure of DNA strands by forming hydrogen bonds with each other to keep two strands intact [4]

DNA coding technology is used to convert binary data to DNA sequence. Binary data can be encoded in DNA by using sequence of alphabet. It is known that DNA sequences contain four basic letters Adenine (A), Cytosine (C), Guanine (G) and Thymine (T).

2. LITERATURE SURVEY

Kazuo Tanaka et al. [5] presented cryptographic algorithm based on public key as an one way function distribution. This technique enhances for encryption and decryption. The public key is generated and same is encoded as DNA sequence. They used control process and PCR extension to decode the agreed DNA sequence.

Sherif T. Amin et al. [6] this technique uses a DNA cryptographic approach based symmetric key algorithm. The key in DNA sequences are obtained from genome and stored large DNA sequence in compact space. This approach has huge storing capabilities compared to other conventional cryptography algorithm.

Guangzhao Cui et al. [7] this paper explains DNA coding, PCR amplification and DNA synthesis and encryption process. The PCR amplification two primers pair was used as key and does not design by sender and receiver. This encryption algorithm is used to increase security purpose. This method can be used to generate different cipher text, which can prevent fraud as PCR primers. This encryption scheme exhibits highly effective privacy.

Lai Xin-she et al. [8] explained novel generation key scheme based on DNA cryptographic approach. This technique uses matrix operation to increase computational speed. They generated key expansion matrix M and generate encryption between two key using XOR operations. This paper uses the DNA sequence as a randomized database; reduce the computation and influence of matrix operation to the computed speed.

Xing Wang, et al. [9] in this paper the author has applied new technique to work cryptography with DNA computing and RSA algorithm is used to connect with DNA computing to encrypt message efficiently. This paper introduced a new encryption algorithm combine with RSA algorithm. This new method of parallel computing is a new method of computation.

Lai, XueJiaet al. [10] proposed DNA sequence as asymmetric encryption and signature method with DNA technology matrix is obtained for encoding the image. Divide the DNA sequence matrix into block and additional operation is performed between block. In this paper original image are jumbled by addition and complement operation, which generates a highly secure secret key of encryption algorithm. This algorithm resists exhaustive attack, statistical attack and differential attack.

Deepak Kumar et al. [11] presented secret data writing using DNA sequence. Paper focused on DNA computing, DNA sequence, which provides large storage capacity and extraordinary information density. Author present encryption and decryption algorithm based on one time pad technique which ensures security on the data in DNA sequence. Steganography is used to hide message in double strand DNA sequence microdots. Author designed data hiding algorithm by using DNA sequence and traditional cryptography. This algorithm is very simple and efficient.

Yunpeng Zhang et al. [12] proposed DNA cryptographic approach based on DNA digital coding and DNA fragment assembly. They provide high security analysis and prove that the algorithm has high confidential strength. This paper exhibits DNA technology based symmetric encryption algorithm. DNA technology has low energy consumption and high storage capacity.

Wang Zhong et al. [13] proposed a new index based symmetric algorithm. This algorithm encrypts plain text using block cipher and index of string. Algorithm converts each character into ASCII code and according to the nucleotide sequence convert into DNA sequence. This algorithm stores position as a cipher text. The researchers have proved efficiency and time complexity of this algorithm through simulation and theoretical analysis.

Tushar Mandge et al. [14] proposed DNA cryptographic approach based on matrix manipulation for making data much secure. They used mathematical manipulation and scrambling in cycles to make data non readable. XOR operation is

performed with the initial key. The benefit for this proposed algorithm is that it always generates different cipher text for same plain text and key.

Monika Borda et al. [15] presented DNA secret writing techniques of bio molecular computation and different algorithm for cryptography and steganography. In this paper author used XOR operation, DNA chromosomes indexing for encoding message. Algorithm uses bioinformatics toolbox and not implementing laboratory experiments.

Amritha Veetil. [16] Introduced a new encryption technique where the genetic operators crossover and mutation are used to encrypt messages, so as to protect the data during transmission. It provides a better encryption technique which is difficult for cryptanalysis. Frequent binary and decimal conversions increases strength of the method and provide high security.

Fatma, [17] proposed a new DNA cryptographic algorithm which used the key features of DNA and amino acid coding to overcome limitations of the classical One Time Pad (OTP) cipher. A significant feature of the proposed algorithm is that; it is considered an encryption and hiding algorithm at the same time. The proposed algorithm also enhances the security level of OTP cipher.

Poornima G, [18] proposed a use Genetic Algorithm which generates an asymmetric key pair which is entirely a new approach unlike the RSA and DES algorithms. It ensures a highly secure key which restricts an unauthorized user.

Mohammadreza, [19] proposed a DNA cryptosystem concepts based on the classic Vigenere cipher. It demonstrates five steps to encrypt and decrypt binary information based on DNA cryptography. The proposed cryptosystems focus on the properties of DNA technology and probability theory.

Mayank, [20] proposed a novel encryption algorithm is suggested based on "BI-SERIAL DNA ENCRYPTION ALGORITHM (BDEA)", which is a bi-serial DNA architecture used to preserve information in a secure manner. This cryptographic approach concentrates on size of cipher text and works only with text data. This paper proposes an encryption scheme through the hex code generation and subsequent MD5 generation of original data to encrypt data for security.

3. PROPOSED ALGORITHM:

Implementation of proposed work:

3.1 DNA Coding:

Consider the alphabetical sequence of the DNA bases A, C, G and T and number them 0, 1 , 2 and 3. We construct a DNA encoding table such that:

S. No	DNA Bases	Binary Conversion
1	A	00
2	C	01
3	G	10
4	T	11

3.2 Encryption process:

1. Data in the file in its ASCII form is converted into their corresponding binary form.
2. Generate binary blocks of 8 bits each, by dividing the binary string into blocks of 8 bits each B1, B2, B3…….. Bn. Where, 'n' is number of blocks generated.
3. Produce a decimal key between the range 0011 to 1277, where the binary form of first three digits denotes the second parent for crossover and fourth digit represents the crossover point and mutation bit.
4. Use binary form of first three digits and perform crossover for all the n blocks and the generated child blocks be C1, C2, C3 ……. Cn.
5. Identify the left out block as L1, L2, L3….. Ln which will be later used for crossover in decryption process.
6. Perform mutation on bit number denoted in key and generate M1, M2, M3 …….. Mn.
7. Convert all the blocks into binary form, into ASCII form and then into its text equivalent.
8. Convert the binary data along with the left out into its DNA sequence and store the result (encrypted data) in a file.

3.3 Decryption process:

1. Convert the encrypted DNA sequence file into its binary form.
2. Generate binary blocks of 8 bits each B1, B2, B3…….. Bn. Where, 'n' is number of blocks generated.
3. Apply mutation for the bit number given in key and generate M1, M2, M3 …….. Mn.

4. Apply crossover with the left out blocks L1, L2, L3….. Ln to get the new child C1, C2, C3 ……. Cn.
5. Convert this binary string to ASCII form and then into text file (decrypted data).

4. EXPERIMENTAL RESULT OF PROPOSED ALGORITHM:

4.1 Encryption:

Step 1: Convert the ASCII into binary form.

T								**I**								**M**								**E**							
0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1

Step 2: Generate blocks of 8 bits.

T								**I**								**M**								**E**							
0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1

Step 3: Given default key = 0354, 035 represent the parent-2 and 4 represent the crossover point and mutation bit.

Step 4: Perform crossover generate the child

T								**I**								**M**								**E**							
0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1
0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	1
0	1	0	1	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	1

Step 5: Identify the left out blocks to use for decryption as a DNA sequence

0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	1	0	0	1	0	1
A			**G**			**C**		**A**			**G**			**G**		**C**			**A**			**G**			**T**			**C**			**A**			**G**			**A**			**C**		

Step 6: Perform mutation on 4th bit of every block

0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1

Step 7: Covert binary to ASCII form, and to text

0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1
C			**A**			**A**		**T**			**C**			**C**		**A**			**T**			**C**			**C**		**A**			**T**		

The encrypted file consists **{CAATAGCACCATAGGCCCATAGTCCCATAGAC} with key=0354**

4.2 Decryption:

Step 1: Covert the DNA sequence into binary form applying DNA encoding.

C			**A**			**A**		**T**			**C**			**C**		**A**			**T**			**C**			**C**		**A**			**T**		
0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1

Step 2: Generate blocks of 8 bits

C			**A**			**A**		**T**			**C**			**C**		**A**			**T**			**C**			**C**		**A**			**T**		
0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	1	0	0	1	1

Step 3: Perform mutation on 4th bit

0	1	0	1	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	1

Step 4: Perform crossover for each block with the corresponding left out as parent 2

C								**S**								**S**								**S**							
0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	1	0		1	0	1	0	0	1	0	0	1	0	1
0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1

Step 5:

0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	1	0	1
T								**I**								**M**								**E**							

4.3 Analysis:

4.3.1 Algorithm validation:

The proposed algorithm encrypts the plaintext and decrypts the cipher text. Two different plain text is chosen to check the algorithm for its effectiveness. The results ensure data is sensitive to proposed algorithm.

Key	1234	1234
Plain text	ENCRYPTION	CRYPTOGRAPHY
Cipher text	CCGT CCGT CCGT CAGT CAGT CAGT CAGT CCGT CCGT CCGT	CCGT CAGT CAGT CAGT CAGT CCGT CCGT CAGT CCGT CAGT
Decrypt text	ENCRYPTION	CRYPTOGRAPHY

4.3.2 Key sensitivity analysis

Using the proposed algorithm with two different keys the sample plain text are tested for cryptography. The cipher text obtained seems to generate unpredictable DNA sequence for two different keys. Hence the proposed algorithm is sensitive to key used.

Key	1234	0534
Plain text	ENCRYPTION	ENCRYPTION
Cipher text	CCGT CCGT CCGT CAGT CAGT CAGT CAGT CCGT CCGT CCGT	CCCC CCCC CCCC CACC CACC CACC CACC CCCC CCCC CCCC
Decrypt text	ENCRYPTION	ENCRYPTION

4.3.3 Data sensitivity analysis

The proposed algorithm is tested for plain text with same key but slight change in plain text. The DNA sequence generated cannot predict the plain text since it depends on the genetic operators which are used for encryption and decryption. The cipher text is not easily detected by an unauthorized user. Hence the algorithm is sensitive to the plain text.

Key	Plain text	Cipher text
1234	ENCRYPTION	CCGT CCGT CCGT CAGT CAGT CAGT CAGT CCGT CCGT CCGT
1234	ENCRYPDION	CCGT CCGT CCGT CAGT CAGT CAGT CAGT CCGT CCGT CCGT
1234	ENSRYPTION	CCGT CCGT CAGT CAGT CAGT CAGT CAGT CCGT CCGT CCGT

5. CONCLUSION:

In today's world data loss through the illegal access is one of the most concerned issues. Providing security is on the priority list therefore a performance measure produced between traditional cryptography algorithm and genetic algorithm in order to validate genetic operators in the field of cryptography has been done [21]. The proposed method in this paper is simple and easy to implement in cryptographic system. The analysis shows that the algorithm is sensitive to key and plaintext. The proposed algorithm is very effective and efficient in encrypting, hiding, transmitting and preventing privacy attacks. This algorithm converts huge data into DNA sequence, which stores large message in compact volume which reduces time complexity. Crossover and mutation operators of genetic algorithm are used for cryptography which provides high security to the transmitted data. Since the data is first converted to its binary form, as future work the proposed system can be implemented for different types of data like text, image, audio or video.

6. REFERENCES:

[1] D. Prabhu , M. Adimoolam, "BI-SERIAL DNA ENCRYPTION ALGORITHM (BDEA)", submitted on 13 JAN 2011, ARXIV: 1101.257

[2] Mohit Rusia, Hemant Makwana, "Review on DNA Based Encryption Algorithm for Text and Image Data", Dept. of IT, DAVV, IET Indore, India, IJERT, Vol. 3 Issue 1, January – 2014, ISSN: 2278-0181

[3] Mona Sabry, Mohamed Hashem and Taymoor Nazmy, ―Three Reversible Data Encoding Algorithms based on DNA and Amino Acids' Structure‖, International Journal of Computer Applications (0975 – 8887) Volume 54– No.8, September 2012.

[4] Borda, Monica, and Olga Tornea, "DNA secret writing Techniques," In Communications (COMM), 8th IEEE International Conference on, pp. 451-456, (2010)

[5] Kazuo Tanaka, Akimitsu Okamoto, and Isao Saito, "Public-key system using DNA as a one-way function for distribution". Bios stems 81, 1, pp. 25-29, (2005).

[6] Sherif T. Amin, MagdySaeb and El-Gindi Salah, "A DNA-based implementation of YAEA encryption algorithm," In Computational Intelligence, pp. 120-125, (2006).

[7] Cui, Guangzhao, Liming Qin, Yanfeng Wang, and Xuncai Zhang, "An encryption scheme using DNA technology," In Bio-Inspired Computing: Theories and Applications, IEEE International Conference on, pp. 37-42, (2008).

[8] Lai Xin-she, Zhang Lei, "A novel generation key scheme based on DNA". Computational Intelligence and security, IEEE, International conference on 13-17 Dec. (2008).

[9] Xing Wang,Qiang Zhang "DNA computing-based cryptography". Key Laboratory of Advanced Design and Intelligent Computing (Dalian university), Ministry of education, Dalian, 116622, China IEEE in 2009.

[10] Lai, XueJia, "Asymmetric encryption and signature method with DNA technology," Science China Information Sciences 53.3, page 506-514, (2010).

[11] Deepak Kumar, and Shailendra Singh, "Secret data writing using DNA sequences," In Emerging Trends in Networks and Computer Communications (ETNCC), IEEE International Conference on, pp. 402-40, (2011).

[12] Yunpeng Zhang, Bochen Fu, and Xianwei Zhang, "DNA cryptography based on DNA Fragment assembly," In Information Science and Digital Content Technology (ICIDT), IEEE International Conference on, vol. 1, pp. 179-182, (2012).

[13] Wang Zhong, Zhy Yu, "Index-based symmetric DNA encryption algorithm". Image and Signal Processing (CISP), 2011 4th International congress on image and signal processing, 15-17 Oct.(2011).

[14] Tushar Mandge, Vijay Choudhary. "A DNA encryption technique based on matrix manipulation and secure key generation scheme". Information Communication and Embedded Systems (ICICES), International Conference on 21-22 Feb. (2013).

[15] Olga Tornea, and Monica E. Borda, "Security and complexity of a DNA-based cipher," IEEE Roedunet International Conference (RoEduNet), 11th, pp. 1-5, (2013).

[16] Amritha Thekkumbadan Veetil, "An Encryption Technique Using Genetic Operators", INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 4, ISSUE 07, JULY 2015

[17] Fatma E. Ibrahim, "A Symmetric Encryption Algorithm based on DNA Computing," International Journal of Computer Applications (0975 – 8887) Volume 97– No.16, July 2014

[18] Poornima G. Naik, "Asymmetric Key Encryption using Genetic Algorithm," International Journal of Latest Trends in Engineering and Technology (IJLTET) Vol. 3 Issue 3 January 2014

[19] Mohammadreza Najaftorkama, "A Method to Encrypt Information with DNA-Based Cryptography," International Journal of Cyber-Security and Digital Forensics (IJCSDF) 4(3): 417-426 2015 (ISSN: 2305-0012)

[20] Mayank Kumar Rusia, "A Novel DNA Based Symmetric Encryption Algorithm for Various Data Formats by Applying Hashing Algorithm," International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 5, Issue 5, May 2017

[21] Delman, B., "Genetic Algorithms in Cryptography", M.S. in Computer Engineering, Rochester Institute ofTechnology, Rochester, New York, July(2014)