International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-5, January 2020

DNA Cryptographic System Utilizing Random Permutation

Shakir M. H. Al-Farraji

Abstract: ***The study of encryption/decryption of information is known as cryptography. The need of protecting information from old years until now is the reason of appearing the process of hiding information from unauthorized people to access it. In this research paper, a cryptographic system is designed by using the DNA computing concepts and random permutation. The proposed system is a block symmetric cipher that uses one initial key in which will be used to generate permutations as many as needed, convert the initial key to DNA key, convert plaintext block to DNA bases. The remaining needed DNA keys are produced through the cipher/deciphering processing. Different operations applied: permute using permutation, modulo and XOR operations to perform the encryption/decryption process. Using the DNA based cryptography enhance the information security and produce highly efficient cipher systems.***

Keywords: ***Cryptography, DNA, DNA computing, DNA Cryptography.***

I. INTRODUCTION

A normal text is called plaintext in cryptography terminology in which anyone reach it can read and understand it even for unauthorized person. To secure the information, encryption information is a suitable way to make it unreadable and incomprehensible and difficult to know its content and details. Encryption is used to hide information from unintended person even to those that can see the encrypted information. Encryption is a process of converting the plaintext to ciphertext. The process of converting the ciphertext to plaintext is called decryption. Both encryption and decryption need a key to process the text. Figure 1, shows the general block diagram of cryptography. There are two types of encryptions: symmetric encryption and asymmetric encryption. Symmetric encryption is also called secret key cryptography which uses the same key for both encryption and decryption. Asymmetric encryption is also known as a public key cryptography, which using two keys, one is called public key for encryption and another key is called the private key for decryption. Confidentiality, integrity, and authenticity are major elements of cryptography [1]. Cryptography can be categorized into three branches: modern encryption, Quantum encryption, and DNA encryption [2]. In this paper, a DNA cryptographic system is proposed that is based on the DNA bases and the random permutations.

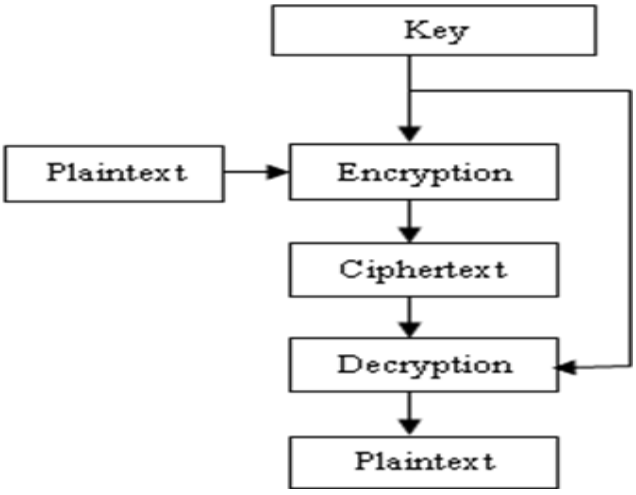

Figure 1. General block diagram of cryptography

II. LITERATURE REVIEW

A. What is DNA

DNA stands for Deoxyribonucleic acid which is sets of nucleotides that contain genetic information. DNA is made up of molecules called nucleotides. Each nucleotide contains a phosphate group, a sugar group and a nitrogen base. Nitrogen bases consists of four types: adenine (A), thymine (T), guanine (G), and cytosine (C). the genetic code or the DNA's instructions is determined by the order of the four bases [3].

B. DNA Computing

Adleman proposed the new research field of bio-computing by using the actual chemistry of DNA to solve problems that are either unfeasible by conventional computers, or need massive amounts of computation. He is the first scientist starting thinking of DNA behavior and relate it to the concept of computability that is presented by Alan Turing, A. Church, and S. Kleene. His experiment of using DNA in solving the Hamilton Path problem was his motivation to think that molecular computers have many attractive properties that are distinct from ordinary computers. It can store a huge amount of data, have the potential of extraordinary energy efficiency, and a powerful parallel processing that current supercomputer can be faster.

Manuscript published on January 30, 2020.
* Correspondence Author
Shakir Al-Farraji*, Department of Computer Science, University of Petra, Amman, Jordan. E-mail: shussain@uop.edu.jo

He stated that many hard problems can be solved by using the DNA computing [4].

C. DNA Cryptography

DNA computing plays main role in the DNA cryptography where the DNA is used as an information carrier and modern biological technology is used as an implementation tool. DNA can be used for computation and cryptography for storing and manipulating information and transmitting that information. Several DNA computing algorithms has been proposed currently such as cryptography, steganography, and cryptanalysis. DNA computing in the area of cryptography is a promising technology that gives potential leading or even unbreakable algorithms [5].

Many researchers work on DNA cryptography and different DNA based encryption system are proposed. R. Biswas et al., proposed an encryption system based on a dynamic DNA base sequences that is assigned to 256 ASCII charters randomly [6]. Karimi and Haider proposed a cipher solution with a new symmetric key generation based upon DNA strands, nucleotides, codons base pair rules, mutation and DNA to mRNA conversion [1]. Tornea and Borda proposed two encryption algorithms based on DNA; one is using public key of binary data followed by its transform in DNA digital sequence; the second algorithm is using XOR one-time-pad cryptographic algorithm, where binary data is transformed in chemically generated DNA structures named tiles [7]. A DNA-based, bio molecular cryptography design using one-time padding (OTP) to encrypt image is proposed by Chen [8]. A pseudo DNA cryptography method that used the biological methods is proposed by Kang Ning and discussed in detail the weakness of the proposed method [9]. A novel generation key scheme based on DNA for creation number table is proposed by Li [10]. The DNA cryptography based on symmetric key exchange is proposed by using one-time pad, symmetric key exchange, and DNA hybridization to minimize the time complexity Tausif [11]. Ochani [12] proposed the solution to transmit a DNA image. They used modified symmetric key encryption with LSB steganography technique, but did not mention about what would be the cover image, and did not provide image compression test with regards to image size and bit quality. Chen proposed a DNA cryptography in which the XOR operation on each bit is carried out independently, thus the encryption/decryption process could be done in a massive, parallel way [13]. T. Mandge and V. Choudhary proposed a DNA encryption technique based on matrix manipulation and secure key generation scheme.

III. PROPOSED SYSTEM

This paper proposes DNA Crypto System utilizes random permutation that generated dynamically from the initial key and encrypt/decrypt data during the encipher/decipher process. It is a symmetric block cipher in which one initial key is needed that is not used directly in the encipher/decipher process. A different block sizes can be used for the plaintext while the DNA block size and permutation are about four times on plaintext block. This system can use up to 64 block size of plaintext; so, the DNA block/permutation size will be up to 256. Different keys are generated during the encipher/decipher process in order to use each one for one block only. Encryption can be made to any file of any data type. The encrypted data consists of the DNA bases only. The XOR operation is applied to the DNA bases (DXOR) shown in Table 1. The DNA Crypto system consists of two types of process: Encryption method and Decryption method which is described in details in this paper.

Table 1: DXOR OF DNA bases

DXOR	A	C	G	T
A	A	C	G	T
C	C	A	T	G
G	G	T	A	C
T	T	G	C	A

A. Encryption Method

Define Reading data for encryption is converted into ASCII values of 8-bits each called a byte. Each byte is divided into four 2-bits which are converted to the four DNA bases: adenine (A), cytosine (C), guanine (G) and thymine (T). The DNA Encryption consists of the following steps:

Step1: input key (k).
Step2: generate M (M is number of input data block+1) random permutations P1, P2, . . ., Pm from the input key (k).
Step3: read one block of plaintext.
Step4: generate DNA key (k).
Step5: preprocess one block of the input data (plaintext) and convert it to DNA bases and rearrange its elements according to permutation Pi.
Step6: Xoring DNA key with the permuted block and rearrange its elements according to permutation Pi+1. The result is the DNA ciphertext block.
Step7: convert the ciphertext to bytes, where each byte consists of the 4 consecutive DNA ciphertext bases.
Step8: add value of each byte to the value at the corresponding position of Pi+1 and take MOD operation, then convert it to character and the result will be considered the input key for the next plaintext block. Go to Step3.

Flowchart of the encryption method is shown in Figure 2.

B. Decryption Method

Decryption process consists of the following steps:

Step1: input initial key (k)
Step2: generate M (M is number of input data block+1) random permutations P1, P2, . . ., Pm from the input key (k).
Step3: read one ciphertext block (Ci).
Step3: generate DNA key
Step4: permute Ci by using Pi+1
Step6: Xoring the permuted Ci with DNA key and rearrange its elements according to permutation Pi. The result is the DNA plaintext block
Step7: convert the DNA plaintext bytes of the original plaintext.

Step8: convert Ci to bytes add value of each byte to the value at the corresponding position of Pi+1 and take MOD operation, then convert it to character and the result will be considered the input key for the next ciphertext block. Go to Step3.

C. Generate permutation

There are several methods available for generating a random permutation. All methods need a seed in order to use it for the generation process. The input key can be used to extract the seed. Two permutations are needed to encrypt one plaintext block; so, two seeds are extracted from the input key in order to generate the needed permutations.

The String Base Random Permutation method (SBRP) is used to generate random permutation as many as needed for the encryption/decryption method [14]. The following is a brief description of this method.

1. SBRP needs an input key and permutation size N.
2. Take the ASCII code of the input key and fill a vector V of size N with these values
3. Apply MOD operation of N on each value in V
4. Keep one of all similar values in V and set others by -1 Replace all -1 values in V with values 1 to N which are not exist in V, the result is a permutation of size N without repetition values

D. DNA key generation

The DNA Key Generation Algorithm is used to generate the DNA key for encryption/decryption from initial key. Any changes to the key even a single bit generate a completely different DNA key. This method is an enhanced to the DNA-Based Cryptographic Key Generation Algorithm [15]. The following is a short description of this algorithm:

1. The input key is used to generate a permutation P of size N out of N! possible permutations.
2. The DNA-based key is generated by using the permutation P as follows:
   2.1. Convert each value of the permutation P to its equivalent binary value (one byte each).
   2.2. Convert each two successive bits to an integer value 0, 1, 2, or 3.
   2.3. A vector V of size 4N is used to store these integer values.
   2.4. Four vectors (V1, V2, V3, and V4) each of size N are made by dividing the vector V.
   2.5. Using the permutation P, permute the four vectors in 2.4 in order to produce a new permuted vector (PV1, PV2, PV3, PV4).
   2.6. The N elements of the DNA key can be calculated as:

For i = 1 To N

$$DNA(i)=(PV1(i)+PV2(N-i+1)+PV3(i)+PV4(N-i+1)) \% 4 \quad (1)$$

Next i

   2.7. Each DNA base is converted to its 2-bits equivalent value in order to produce the DNA key (A as 0→00, C as 1→01, G as 2→10, and T as 3→11).

Figure 3, shows the flowchart of the decryption method.

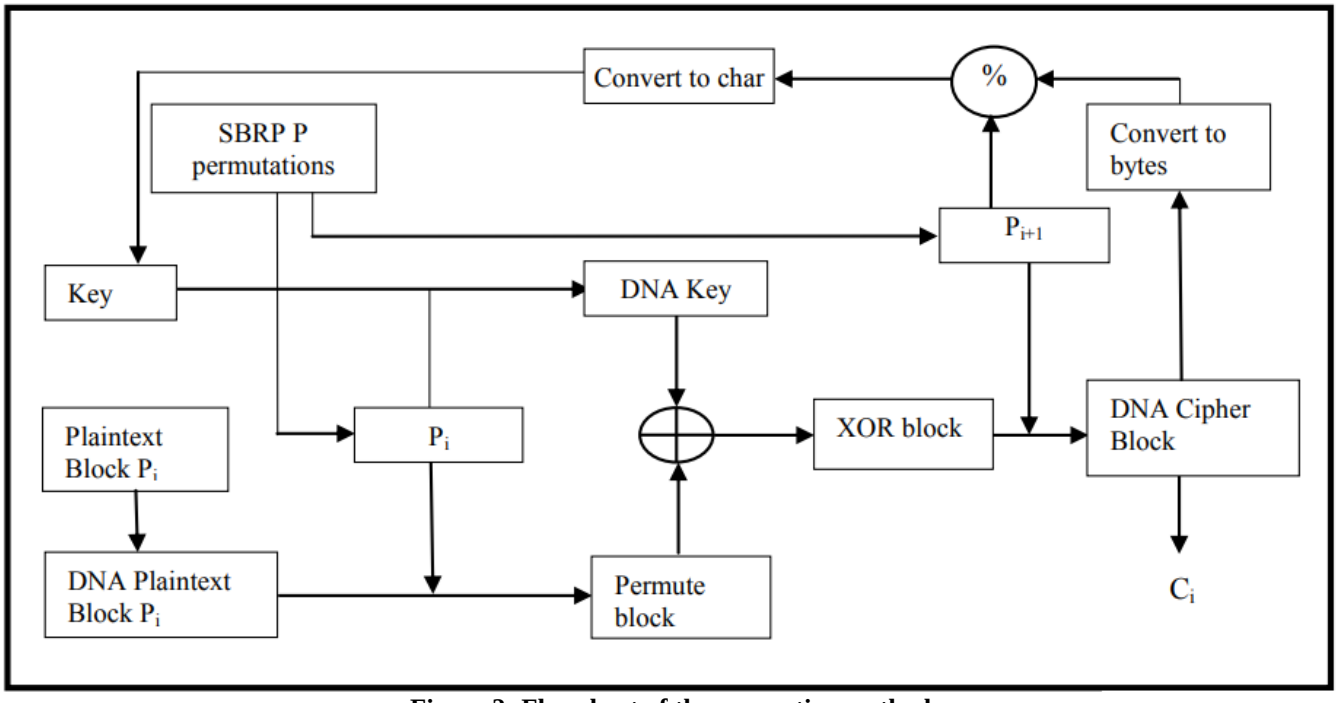

Figure 2: Flowchart of the encryption method

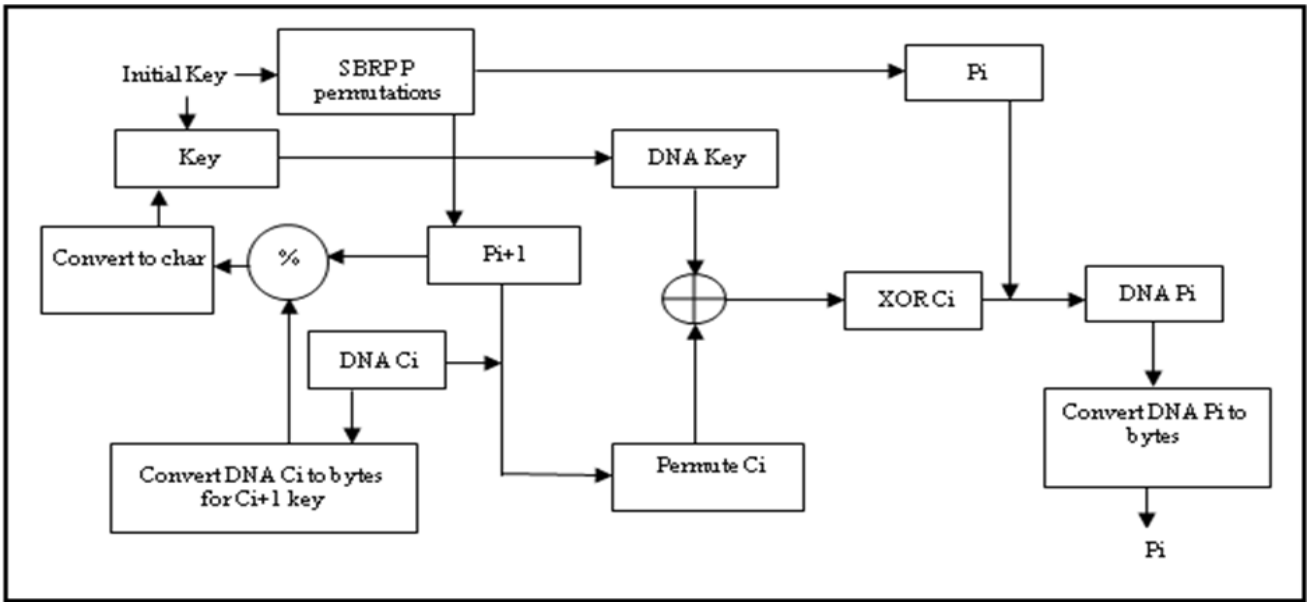

Figure 3: Flowchart of the decryption method

IV. ILLUSTRATED EXAMPLE AND RESULTS

To demonstrate the encryption process, let take a plaintext of size 14 characters and a block size of 7. Also, we need an initial key in order to generate the first DNA key and two random permutations for each plaintext block to be encrypted.

Initial Key: computers

Another two keys are extracted from the key

Plaintext block size: 7

Permutation size: 28

DNA key:
GACACCCTACACGGGACACAGGACACTG
Plaintext: dna encryption
DNA plaintext
Block1 (b1): dna enc:
CGCACGTGCGACAGAACGCCCGTGCGAT
Block2 (b2): ryption:
CTAGCTGCCTAACTCACGGCCGTTCGTG
Permutation
SBRP gives three permutations, the first two will be used to encrypt the first plaintext block.

P1: 25 1 14 7 9 21 12 19 13 27 24 26 23 5 18 8 2 11 6 20 3 22 28 17 4 10 16 15
P2: 21 19 7 13 18 24 17 1 27 11 4 15 5 8 12 23 28 14 22 6 10 9 2 20 26 25 16 3

The result of encryption shown in Figure 4. The matching between the generated DNA key and the DNA ciphertext. The correlation between DNA key and DNA ciphertext gives a low value (0.062) that refers to a good confusion. Permutations that are used to permute plaintext, ciphertext, and used for generating all keys (permute and adds with modulo operation) provide full diffusion.

Table 2, shows the plaintext and the ciphertext as a character, hexadecimal, and decimal values. Comparing the corresponding elements of the original plaintext and ciphertext shows that there is no relation between them.

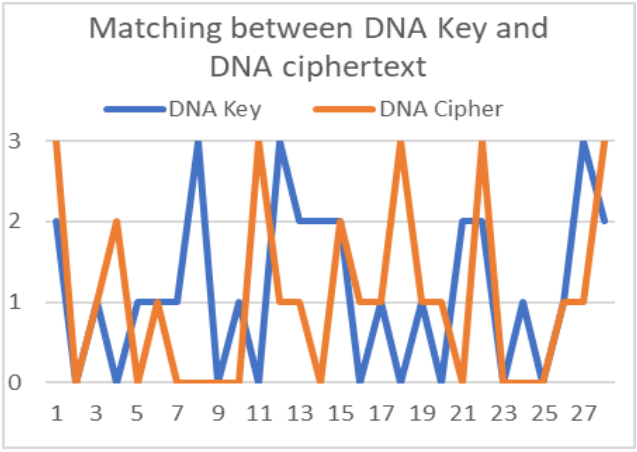

Figure 4. Matching between DNA key and DNA ciphertext

P1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
25	1	14	7	9	21	12	19	13	27	24	26	23	5	18	8	2	11	6	20	3	22	28	17	4	10	16	5

b1

C	G	C	A	C	G	T	G	C	G	A	C	A	G	A	A	C	G	C	C	C	G	T	G	C	G	A	T

DNA key

G	A	C	A	C	C	C	T	A	C	A	C	G	G	G	A	C	A	C	A	G	G	A	C	A	C	T	G

Permuted b1: Rearrange b1 according to the random permutation P1, here first element of b1 'C' goes to cell 25 according to first element in P1

G	C	C	C	G	C	A	A	C	G	G	T	C	C	T	A	G	A	G	C	G	G	A	A	C	C	G	T

Xb1 = DNA key DXOR Permuted b1

A	C	A	C	T	A	C	T	C	T	G	G	A	A	C	A	T	A	T	C	A	A	A	C	C	A	C	C

P2

21	19	7	13	18	24	17	1	27	11	4	15	5	8	12	23	28	14	22	6	10	9	2	20	26	25	16	3

Ciphertext block1: Permuted Xb1 by P2

T	A	C	G	A	C	A	A	A	A	T	C	C	A	G	C	C	T	C	C	A	T	A	A	A	C	C	T

DNA plaintext:
CGCACGTGCGACAGAACGCCCGTGCGAT
DNA ciphertext:
TACGACAAAATCCAGCCTCCATAAACCT

Table 2: Result of encryption

	char	Decimal value	Hex	correlation
Plaintext	dna enc	100 110 97 32 101 110 99	64 6E 61 20 65 6E 63	0
Ciphertext	╞Iu0	198 16 13 73 117 48 23	C6 10 0D 49 75 30 17	

V. CONCLUSION

This paper introduced a new DNA cryptographic method for symmetric block cipher system in which permutations play main role in the process. The hardness of this process is the generation of permutations that is built dynamically rather than uses a predefined permutation. In addition, the key is not used directly; but it used to generate one DNA key in order to apply it for encryption/decryption data. Other DNA keys are generated during the encryption/decryption the remaining plaintext data. The preprocessing of the plaintext is used to convert it to DNA base elements. Different plaintext block size can be used. The best plaintext block size we suggest to be used is 64 in order to use permutation size and DNA key size 256; this gives more secured ciphertext with attacking complexity 256!

REFERENCES

1. M. Karimi and W. Haider, "Cryptography using DNA Nucleotides", International Journal of Computer Applications (0975 – 8887) Volume 168 – No.7, June 2017
2. S. Dhawan and A. Saini, "Integration of DNA Cryptography for Complex Biological Interactions", International Journal of Engineering, Business and Enterprise Applications (IJEBEA), pp. 31-36, 2 (1), Aug-Nov, 2012.
3. Rachael Rettner, "DNA: Definition, Structure & Discovery",2017-12-08T02:30:00Z, https://www.livescience.com/37247-dna.html
4. L. M. Adleman, "Molecular computation of solution to combinatorial problems", Science, vol. 266, pp. 1021-1024, November 1994.
5. Dhawan et al., "Integration of DNA Cryptography for Complex Biological Interactions", International Journal of Engineering, Business and Enterprise Applications, 2 (1), Aug-Nov, 2012, pp. 31-36
6. R. Biswas et al., A technique for DNA cryptography based on dynamic mechanisms, Information Security and Applications, Vol. 48, October 2019, 102363
7. O. Tornea and M. Borda, "DNA Cryptographic Algorithms" International Conference on Advancements of Medicine and Health Care through Technology pp 223-226, 23–26 September, 2009.
8. Chen, J.2003. A DNA-based, bio molecular cryptography design. Circuits and Systems", IEEE International Symposium on Circuits and Systems, vol. 3, pp. 822-825, June 2003
9. Kang Ning. A pseudo DNA cryptography method, http://arxiv.org/abs/0903.2693v1

10. Li et al., “A novel generation key scheme based on DNA”, 2008 International Conference on Computational Intelligence and Security, pp. 264 – 266, 2008.
11. T. Anwar et al., “DNA Cryptography Based on Symmetric Key Exchange”, International Journal of Engineering and Technology, vol. 7, No. 3 June-Jul 2015.
12. Ochani A, Jadhav D, Gulwani R (2017) DNA Image encryption using modified symmetric key (MSK). In: International conference on inventive computation technologies, IEEE, Coimbatore, pp 1–4
13. T. Mandge and V. Choudhary, “,A DNA encryption technique based on matrix manipulation and secure key generation scheme”, 2013 International Conference on Information Communication and Embedded Systems (ICICES), 21-22 Feb. 2013
14. Shakir M. H. Al-Farraji, USPTO patent No. 10.496.377
15. Shakir M. Hussain and Hussein Al-Bahadili. A DNA-Based Cryptographic Key Generation Algorithm. Proceedings of the International Conference on Security and Management (SAM'16), Las Vegas, USA, 25-28 July 2016.

AUTHOR PROFILE

Shakir M. Al-Farraji (shussain@uop.edu.jo) He received his B.A. degree in statistics from University of Al-Mustansiriyah, Iraq, in 1976 and M.Sc. degree in Computing and Information Science from Oklahoma State University, USA, in 1984. In 1997 he received his Ph.D. degree in Computer Science from University of Technology, Iraq. From 1997 to 2008 he was a faculty member at Applied Science University, Jordan. Currently, he is an associate professor at Petra University, department of computer Science, Faculty of Information Technology, Jordan. His research interest covers encryption, key generation, authentication, and data compression. He received a USPTO patent for generating random permutation and currently working on DNA based cryptography in encryption, authentication, and digital signature. He is a member of ACM.