International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN No.:2349-9745, Date: 2-4 July, 2015

Secure Encryption Techniques Using DNA Computation

Drashti O. Vadaviya[1], Prof. Purvi H. Tandel[2]

[1]*Department of Computer engineering, CGPIT, UkaTarsadia University, Bardoli , India, drashtivadaviya@gmail.com*
[2]*Department of Computer engineering, CGPIT, UkaTarsadia University, Bardoli , India, purvi.tandel@utu.ac.in*

Abstract- The 21st century is a period of information explosion, which leads to increase in importance of information security. As traditional cryptographic methods are vulnerable to attacks, DNA cryptographic concept has been identified as a possible technology in the hope of unbreakable algorithms. DNA cryptography is a new broad scientific branch with combination of classical solutions in cryptography with the strength of the genetic material. In this paper the different methods of DNA Computation are discussed to enhance the security.
Keywords- DNA Cryptography, RSA, AES, Central dogma of molecular biology .

I. INTRODUCTION

Cryptography is the basic need of today's information security. Modern cryptography encryption technology are now a days vulnerable against modern cryptographic attacks. This means that the subject of finding new and powerful ciphers is always of interest and new directions in cryptography which can be explored. DNA cryptographic concept has been identified as a possible technology in the hope of unbreakable algorithms. DNA cryptography uses advantages of DNA structure like randomness and uniqueness in order to achieve strong encryption. There are many advantages of DNA structure like parallel molecular computation and its large storage capabilities, made this research field a very strong and secure one for various applications. To get proper idea of DNA computing we have to know about DNA structure or some basic operations of molecular biology.

1.1 Bio-molecular Technology Background

A DNA molecule has double-stranded structure obtained by two single-stranded DNA chains, containing four different types of nucleotides (A adenine, G guanine, C cytosine, T thymine), bonded together by hydrogen bonds: A = T double bond and C G triple bond [1].Watson-Crick define the complementarity rule which states that A always comes in complimentary to T and C always comes in complimentary to G [7].

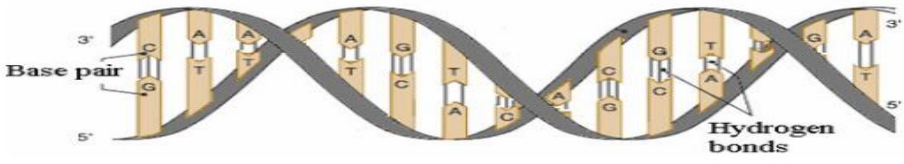

Figure 1.1: DNA Structure [1]

The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information.

Figure 1.2: Central dogma of molecular biology [1]

Polymerase Chain Reaction (PCR) is fast amplification technique. The PCR is a very sensitive method, theoretically single strand DNA molecule can be amplified up to 106 after 20 cycles.

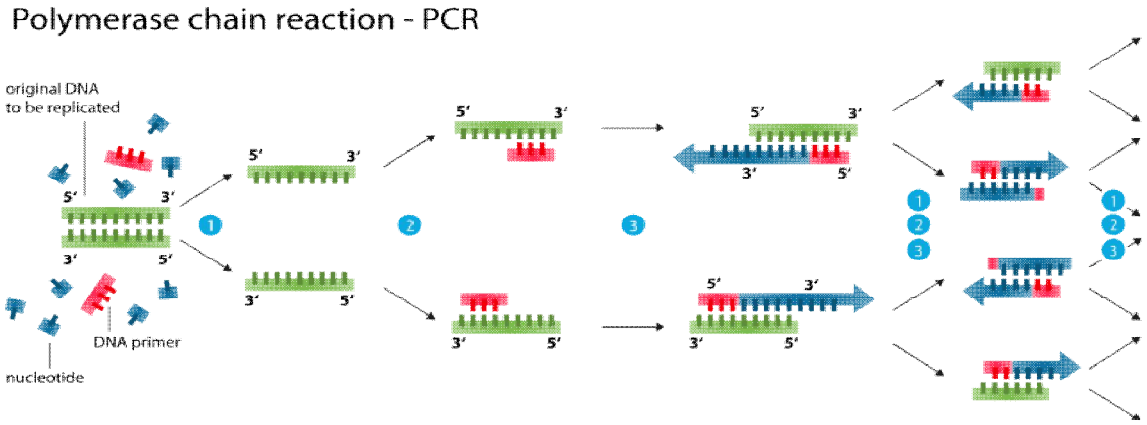

Figure 1.3: Central dogma of molecular biology [1]

The simplest coding pattern to implement is binary, so to encode the 4 nucleotide bases (A, T, C, G) we can use 4 digits: 0(00), 1(01), 2(10), 3(11). According to this concept there are 4! = 24 possible coding patterns by this encoding format, but according to Watson-Crick complementarity rule, that is 0(00) to 3(11) and 1(01) to 2(10). So among these 24 patterns, only 8 kinds of patterns which satisfy complementary rules as shown in Table 1.1 are useful [3].

Table 1: DNA Digital coding [3]

0123 CTAG	0123 CATG	0123 GTAC	0123 GATC
0123 TCGA	0123 TGCA	0123 ACGT	0123 AGCT

II. CRYPTOGRAPHIC METHOD USING DNA

DNA cryptography is new field born from Adleman's research [6]. Adleman describes how DNA computing can be used to solve a directed hamiltonian path problem. After his research work new direction of DNA computation is evolved to solve many complex problems.

2.1 DNA Secrete Writing Technique's

There are different secrete writing technique's discussed in this paper [1]. First is an OTP (one time pad): In this method the ssDNA (Single strand DNA sequence) is used for OTP key generation. Second is a DNA XOR OTP: In this method the DNA tiles are used, each tile contains one bit ether 0 or 1. Which have upper and lower end with DNA strand, used for binding other tiles with complementary DNA strand. Third is a DNA chromosome indexing: The index of the point where the DNA sequence for character is match with FASTA file sequence considered as pointer and stored in ciphertext. So, instead of sending text to receiver the index of DNA FASTA file has been send.

2.2 DNA Computing Based Cryptography

This technique [2] proposes a new encryption schema which was a combination of RSA algorithm with DNA computation concepts. Plain text is first encrypted using RSA algorithm and some complexity is added in process using DNA digital coding process.

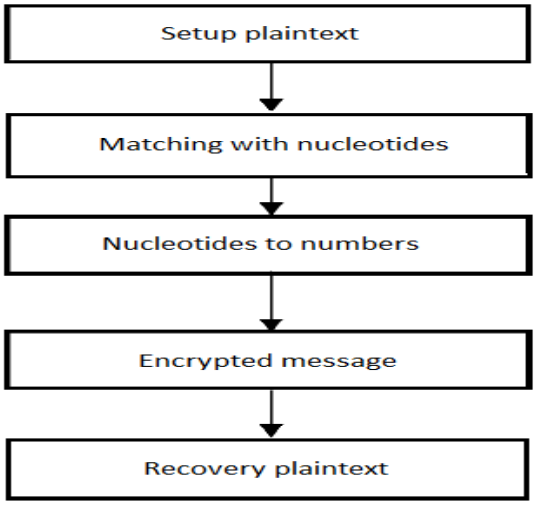

Figure 2.1: DNA Computing Based Cryptography [2]

Step 1: Set positive integer to n and e, where n and e are large random prime numbers.
Step 2: To get the base nucleotides, use the Table 2.1 to turn plaintext into corresponding nucleotides.

Table 2.1: Amino acid table [2]

A-CCA	B-GTT	C-TTG	D-GGT	E-TTT	F-TCG	G-CGC	H-ATG	I-AGT
J-CGA	K-GAA	L-CGT	M-CCT	N-TCT	O-CGG	P-ACA	Q-CAA	R-ACT
S-GCA	T-CTT	U-GTC	V-TCC	W-GCC	X-ATC	Y-AAA	Z-TCA	

Step 3: Turn the base nucleotides chain into numbers with Table 2.2.

Table 2.2: Number Representation [2]

A-01	C-03	G-07	T-20

Step 4: Compute e^{th} power of n. Compute decipher exponential d.
Step 5: Separate the number into three parts. Make sure every part of number is less than n.
Step 6: Through calculation of the three parts numbers, we get the cipher text. In the calculation process, it needs to compute Euler's theorem with.

$$CT=PT^{E} \mod N$$

Step 7: For decryption, calculate the plaintext from the cipher text as this: $PT = CT^{D} \mod N$. Reverse transformation from Table 2.2 and Table 2.1 is calculated and the result of this process is plain text.

2.3 An Encryption Scheme Using DNA Technology

In this technique [3], an encryption scheme is designed by using the technologies of DNA synthesis, PCR amplification and DNA digital coding as well as the theory of traditional cryptography. The step wise process is as below.

Step 1: Key Generation: Two primers pairs are used as the encryption and decryption keys. This process provides more security to this encryption scheme, because even if an adversary

somehow caught one of a primer pair, he is not able to collect the data because by using both correct primer sequences the amplification could be successful.

Step 2: Data Pretreatment: The plain text is converted into hex value and then hex is converted to binary.

Step 3: Encryption: Plain text is encrypted using receiver's secret key e then the ciphertext C is converted into the DNA sequence using DNA digital coding. Then the encrypted message that is nothing but DNA sequence is covered by forward and reverse PCR primers and secrete-message using DNA is prepared.

Step 4: Decryption: After the intended receiver gets the DNA strand, he extracts the secret message from DNA sequence by using the forward and reverse primer pairs. Receiver translates the secret message DNA sequence into the binary ciphertext C. Then, ciphertext C is converted into plain text M using his secret key e.

2.3 YAEADNA Encryption Algorithm

The investigation conducted in this paper [4] is based on a conventional symmetric encryption algorithm called "Yet Another Encryption Algorithm" (YAEA) developed by Saeb and Baith. DNA-based implementation of YAEA encryption algorithm uses a search technique in order to locate and return the position of quadruple DNA nucleotides sequence. This sequence is representing the binary octets of each plain-text characters.

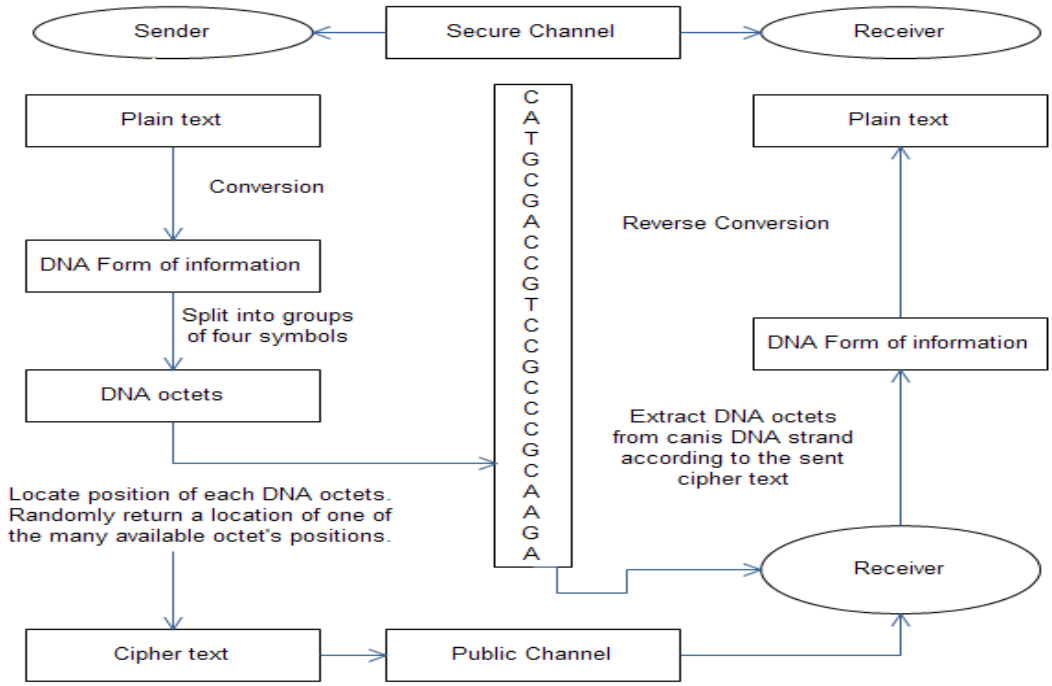

Figure 2.2: The Computational graph of single YAEADNA round [4].

In simple words we can say that each four DNA nucleotides sequence represents a binary octet that in turn represents a plain-text character of the message. The sender will match all the quadruple DNA nucleotides sequence representing all the plain-text characters against the single strand DNA representing the chains familiarize genome in order to locate all of its occurrences and positions or in other words its pointers. This process will generate a file that contains random location pointers for

all plain-text characters; this file is considered as ciphertext. The ciphertext is then transmitted to the receiver.

Upon receiving this file, the receiver uses the ciphered text to recover the sequences of quadruple DNA nucleotides sequence from the Canis DNA strand, which was previously delivered to him through a secured channel. In the reverse order, the recovered DNA form of information will be used to retrieve the binary form of information, which will be then translated into plain text.

2.4 An Encryption Algorithm Inspired from DNA

The proposed encryption algorithm [5] is a symmetric key block cipher that has a fixed block size of 128 bits and a key size of 128, or 256 bits. Like the standard DES and AES, the encryption algorithm consists in three phases: Initial phase, iteration phase and final phase. It uses a sub-keys generator that generates sub-keys of 128 bits from a principal key of 128 or 256 bits. Aim is to apply as well as possible the Shannon's principles of confusion and diffusion via the use of substitutions and permutations [5].

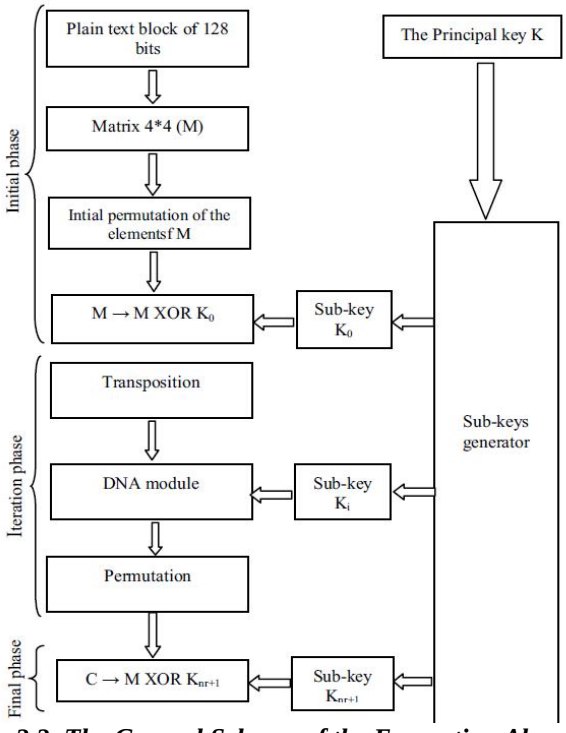

Figure 2.3: The General Schema of the Encryption Algorithm [5]

Encryption Algorithm

Input: A binary form of plain text, the principal Key K.

Output: The ciphered text.

Algorithm body

1. Subdivide the plain text into blocks of 128 bits.
2. Transform a block b into a matrix M of 4*4 bytes.
3. Apply an initial permutation to the matrix M, the result is also named M.
4. Generate the sub-keys from the principal Key K using the sub-key generator of the AES algorithm.
5. Perform XOR operation between M and the first sub-key K_0: M= M XOR K_0.
6. Perform nr rounds, each one composed of three steps,

a. A transposition of the matrix M: M transposition (M).
b. DNA module: it's the only step of the algorithm where the confusion principle is concretized via a set of substitutions inspired from the central dogma of molecular biology.
c. Perform a permutation of the elements of M.

7. Perform a XOR operation between the matrix M and the last sub-key K_{nr+1}: M= M XOR K_{nr+1}.
8. Repeat the same procedure, starting from step 2, for all the blocks.

End Algorithm.

The permutation can be done using Table 2.3. Each permutation is applied to both rows and columns of matrix M. The (i mod 3,M) permutation: apply i mod 3 permutation to row and (i+1) mod 3 permutation to column of matrix M. Through this process, diffusion can be increased in order to achieve more security.

Table 2.3: Permutation Table [5]

Permutation 0	2	0	3	1
Permutation 1	3	2	1	0
Permutation 2	1	3	0	2

DNA module consists of 3 steps. (1) Transcription process: the binary data is converted into DNA string using DNA digital coding process. (2) Bio_XOR: the DNA string is ex-ored with the i^{th} round key. (3) Translation: 16*16 matrix is generated using concept of amino acid table and the DNA strand is substituted using amino acid table [5]. First to base are considered for row index and last two base if for column index

Table 2.4: Bio_XOR Table [5]

BIO_XOR	**A**	**C**	**G**	**T**
A	A	C	G	T
C	C	A	T	G
G	G	T	A	C
T	T	G	C	A

III. COMPARATIVE ANALYSIS

In above section we get brief overview of different techniques as shown in Table 3.1. DNA secret writing techniques are using some biological concepts of DNA computing to encryption process, but they are not applicable in normal environment they need high-tech laboratory for implementation. An encryption scheme using DNA technology uses the primer pair concept. It would still be extremely difficult to amplify the message-encoded sequence without knowing the correct two primers pairs. If an adversary without knowing the correct two primer pairs wants to pick out the message encoded sequence by PCR amplification, he must choose two primer sequences from about 1023 kinds of sequences. YAEADNA encryption algorithm uses pointer concepts. If an adversary get the ciphertext file still he is unable to retrieve plaintext without knowing correct FASTA sequence shared between two parties. An encryption scheme inspired from DNA uses AES algorithm as its base algorithm which itself is strong and by adding DNA computation concepts.

Table 3.1: Comparative Study

DNA Cryptography Algorithm	Technology
DNA Secret writing Techniques[1]	One-Time-Pad (OTP), DNA XOR OTP and DNA chromosomes indexing.
DNA computing based cryptography[2]	Amino acid table - A plain text is encrypted by RSA and the result will replaced by proteins defined in amino acid table.
An Encryption Scheme Using DNA Technology[3]	DNA digital coding PCR primers - A message is converted to DNA template in which primers are used as key to encode and decode the message.
YAEADNA Encryption Algorithm[4]	DNA Sequence Matching - data converted into pointers according to DNA strand taken and key send to the receiver in a secure channel.
An Encryption Algorithm Inspired From DNA[5]	Symmetric key block cipher algorithm Transcription (DNA-RNA) Translation (RNA - Protein) {message converted into matrix with initial permutation and XOR operation is performed with the key which is subjected to DNA module transcription and translation.

CONCLUSION

DNA based cryptographic algorithms have satisfactory results in terms of security and performance. Key features of DNA such as large storage capacity and uniqueness, provides more security to DNA based cryptographic algorithm. Certain DNA algorithms can resist exhaustive attack, statistical attack and differential attack. DNA cryptography is still at its initial stage and it is a very promising direction in cryptographic research.

REFERENCES

[1] Monica BORDA, Olga TORNEA, "DNA Secret Writing Techniques", Communications (COMM), 2010 8th International Conference. pp. 451-456, IEEE 2010.
[2] Xing Wang, Qiang Zhang, "DNA computing-based cryptography", Bio-Inspired Computing, 2009. BICTA '09. Fourth International Conference. PP. 1-3, IEEE 2009.
[3] Guangzhao Cui, Limin Qin, Yanfeng Wang, Xuncai Zhang, "An Encryption Scheme Using DNA Technology", Bio-Inspired Computing: Theories and Applications, 2008. BICTA 2008. 3rd International Conference. pp. 37-42, IEEE2008.
[4] Sherif T. Amin, MagdySaeb, Salah El-Gindi, "A DNA- based Implementation of YAEA Encryption Algorithm", IASTED International Conference on Computational Intelligence, 2006.
[5] SouhilaSadeg, Mohamed Gougache, N. Mansouri, H. Drias, "An Encryption algorithm inspired from DNA", IEEE (Nov 2010) pp 344 349.
[6] Leonard.Adleman, "Molecular computation of solutions to combinatorial Problems", Sience: 266:1021-1024, 1994.
[7] Francis crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid.", April 25, 1953. Nature 171 (April 25,1953): 737-738.

