An Survey on DNA Based Cryptography

Bonny B Raj
Department of Computer Science and Engineering
HITS,
Chennai, Tamilnadu, India
bonny.renjith@gmail.com

V Ceronmani Sharmila
Department of Information Technology
HITS,
Chennai, Tamilnadu, India
csharmila@hindustanuniv.ac.in

***Abstract* - Information Security play a vital issue in today's life. In this paper DNA is used as an information carrier. DNA cryptography is an upcoming technology, which combines biological information and data security. They provide wide range of extraordinary parallelism, exceptional energy efficiency, storing and computing capabilities. This paper presents a review on DNA Cryptographic approaches, highlights the merits and demerits of each one.**

Keywords - Biological Information, DNA Cryptography, Information security, Information Technology.

I. Introduction

Security threats are increasing day by day due to the information flow in network .The most confidential infor-mation such as business records, health records and private information should be broadcasted via a public transmission media with extreme care. The security of these information involves an enormous threats by an unintentional recipient. Cryptography helps us in ensuring protection of such data. Cryptographic methods provide transmission of data with secure storing of data across channels so that it can be un-derstood only by the intended receiver. Many algorithms have been developed for the purpose of encrypting the in-formation or data.

II. CRYPTOGRAPHIC ASPECTS

Cryptography is concerned with four major aspects. They include:

A. Confidentiality

It deals with protecting data from unauthorized users.

B. Integrity

Receiving data without any kind of alteration.

C. *Non repudiation*

Sender cannot stop sending data.

D. *Authentication*

Confirms that data is coming from right person.

III. DNA

DNA stands for De-oxyribo nucleic acid is a thread like chain of molecules known as nucleic acid. They are used for transmitting genetic instructions which in turn is used in growth, development, functioning and reproduction

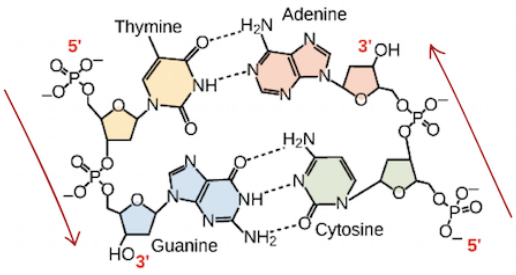

Fig. 1 Structure of DNA

living organism [1]. This intricate design of the organism includes human organ that are the outcome of assigning cer-tain procedure to the virgin message which in turn will be encoded in a DNA sequences which is termed as genes.

Further these complicated mathematical operations are subjected to basic additions and subtractions. One of the major advantage of DNA molecule is that it is a combination of four bases, Adenine (A), Thiamine (T), Guanine (G) and Cytosine (C).These four bases combines in different order to form Purines (Guanine and Adenine combinations) and Py-rimidines (Thymine and Cytosine combinations). These bi-strands of DNA molecules are anti parallel and they can moves in the reverse directions. This complementary adds up DNA with a different data pattern for computing and may be used in different aspects. Fig. 1 shows the structure of DNA.

IV. DNA CRYPTOGRAPHY AND COMPUTING

DNA Computing and Cryptography concept has emerged from computational ability of De-oxyribo Nucleic Acid (DNA). It is a cryptographic branch of biological science which make use of computing more complex to the unauthorized user [2].

A. Operations on DNA

DNA molecules are helpful in solving computational and mathematical problems with the help of different biological operations which can be summarized as :

1) Hybridization: In this method two numbers of single stranded DNA molecules are combined together to form a double stranded DNA molecule.

2) Denaturation: This is the method of separating dou-ble stranded molecule to two separate single stranded DNA.

3) Litigation: This is a method of combing two double stranded DNA sequence to a new double stranded DNA sequence.

4) Polymerase Chain Reaction: A Polymerase chain reaction or simply PCR is a molecular biological application method for amplifying a single copy or a few copies of a piece
of DNA molecule across several orders of degree, generating thousands or millions of copies of a specific DNA combinations.

5) Gel Electrophoresis: A Gel Electrophoresis is a process of separating DNA molecules located in the gel. Here gel is used as a filter. Electrophoresis refers to the push of DNA strands through the gel filter. This can be done with the help of a electrophoresis box by applying electricity, where it moves DNA. Short strands move quickly where as longer strands moves slowly through the holes of gel. Strands with equal length move with uniform speed and ends ups successively and thus the DNA threads gets sort by themselves.

V. BASIC TERMINOLOGIES IN DNA

Some fundamental biological terms which is commonly used in DNA Computation and Cryptography are listed as follows:

A. Codons

Codons are combinations of tri-bases of nucleotide in the form of a triplet.

B. Genes

Genes act as the functioning units of DNA molecules. An individual gene consist of definite sets of data or encoding of a definite protein or particular process. They can be coded or non-coded sequences. The coded sequences are known by the term Exons and non-coded sequences as Introns.

C. Chromosome

Chromosomes are huge DNA structures which is coiled over protein molecules, that includes genes, other regulative elements and sequences of nucleotides. In other words chromosomes may be a long strings of genes.

D. Genome

Genome are specific combinations of being which in-clude genes, chromosome, nucleotides and DNA of a living cell.

The fundamental techniques used in DNA Encryption are DNA digital coding and PCR Amplification [3].

VI. DNA DIGITAL ENCODING

DNA digital encoding is a method of aligning DNA molecules. They act a major role in encryption and decryp-tion of data. We cannot use DNA molecule as alphabets for computation purpose, for this they make use of the approach BCD or binary digital coding, that encode the combinations as 1's and 0's. DNA encoding can be done on the basis of four nucleotide bases A, C, T and G. These four bit subunits of DNA molecule are then converted into two bit binary values as A as 0 [00], T as 1 [01], C as 2 [10], G as 3 [11] as shown in the table 1. DNA digital encoding is a method of aligning DNA molecules. They act a major role in encryption and decryption of data. We cannot use DNA molecule as alphabets for computation purpose, for this they make use of the approach BCD or binary digital coding, that encode the combinations as 1's and 0's. DNA encoding can be done on the basis of four nucleotide bases A, C, T and G. These four bit subunits of DNA molecule are then converted into two bit binary values as A as 0 [00], T as 1 [01], C as 2 [10], G as 3 [11] as shown in the Table I.

TABLE I
DNA DIGITAL ENCODING

Encoding DNA Nucleotide	Decimal Value	Binary Value
A	0	00
C	1	01
T	2	10
G	3	11

Suppose the plain text is "ABC" initial step of this pro-cess is to convert each characteristic to its corresponding ASCII value (A=65, B=66). Then they are converted to its binary value (the sequences of bits 0's and 1's will be 65= 1000001 and 66=1000010, and it will be mapped to ACTG combinations using the table1 as TAAGAAT and the data is ready for encryption. Along with DNA digital encoding fur-ther encryption methods can be included for making it more secure like PCR Amplification process.

VII. PCR AMPLIFICATION

PCR stands for polymer chain reaction which hold the bio-molecular method of DNA Amplification which is placed on the Watson-crick complementary model. Here two pri-mers are coded with a DNA sequences. Primers are small units of DNA. The two primer combinations can be passed on as the key for the PCR Amplification Process [4].PCR Amplification is shown in the Fig. 2.

The original information can be positioned securely be-tween the two primers for encoding them into a new se-quence. The resultant message is very difficult to amplify if the PCR primers are unknown. Different length primers will generate different results, so correctness of primers are very important in this process. Biological PCR operations are as follows:

A. Step 1- Denaturation

PCR Amplification is considered as the first step of denaturation, here a dual stranded molecule can be separated into two single stranded DNA molecule. This fragment can be again heated up to 94 to 95 degree for one to few minutes to denature or to get separated as double stranded molecule to single strand.

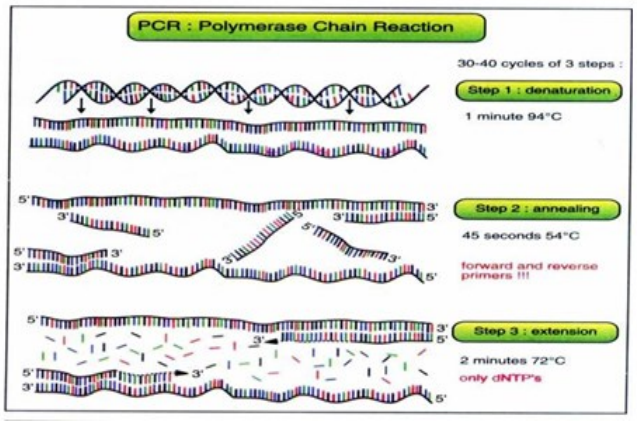

Fig. 2 Polymer Chain Reaction

B. Step 2- Primer Annealing

In Primer Annealing the temperature is reduced to 50 to 65 degree for one to few minutes so that the primer get at-tached to the corresponding complementary pairs. These primers are structured to strengthen the DNA region.

C. Step 3- Primer Extension

Here the temperature can be increased to 72 degree for one to few minutes. Thus polymer enzyme will add nucleo-tides to the strands of shorter primers on original DNA strands, here by amplifying the DNA strands between the primers.

VII. DNA ENCODING METHODS

DNA Digital encoding and PCR methods are the most commonly used encryption techniques in DNA based Cryp-tography. In addition to this the other few techniques, they include:

A. DNA Chip Based

DNA chips and micro-arrays have opened a new way for doing research. DNA chip based technology is best suit-ed for storing and handling biological information. They can support wide range of manipulation of genome sequence. Such DNA Chip consist of numerous ports which is embed-ded in a solid surface, commonly a glass slide. Here each spot consist of a large number of ports, which are nucleo-tides sequence capable of binding complementary nucleotide [5].

B. DNA Steganography

The method of hiding secret message inside an message is known an steganography. Any medium like message, video or image file can be used for hiding message. In DNA based steganography, input DNA strands with the given message could be combined along with a random secret key DNA molecule, which is concealed inside another DNA molecule. An original text can be retrieved using hybridization with the complement of secret key strands. This method is best suited for large-scale encryption.

C. DNA Certification

DNA certification deals with the biological character of DNA. This is widely used in the field of justice, finance etc. They are used to certify biological individual accuracy.

VIII. CONCLUSION

DNA molecule can hold both feasibility and applicability of DNA-based Cryptography. A secured DNA based crypto-graphic algorithms provides multi levels of security. Some DNA algorithms are provided with resisting different types of attacks including statistical, differentia and exhaustive attacks. DNA computation and its applications are still in the world of reverie and is not yet been utilized properly. They can also deal with better performance on data storage and security. Let us hope that DNA computation and its authen-tication processes will extend it in future with better perfor-mance for information storage and security.

REFERENCES

[1] Tushar Mandge, Vijay Choudhary, "A Review on Emerging Cryptography Technique: DNA Cryptography" Proceedings published in International Journal of Computer Applications, pp 9-13, 2012.

[2] Zhihua Chen, Xiutang Geng, Jin Xu, "Efficient DNA sticker algorithms for DES", IEEE 3rd International Conference on Bio-Inspired Computing, BICTA08, pp 15-22, 2008.

[3] L. Adleman, Molecular Computation of Solutions to Combinatorial Problems, Science, vol. 266, pp. 1021-1024, 1994.

[4] R. J. Lipton, Using DNA to solve NP complete problems Science, Vol. 268, pp. 542-545, 1995.

[5] A. Gehani, T. La Bean, and J. Reif, DNA-based Cryptography, Lecture notes in Computer Science, Springer, 2004

.