Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 4, April 2016, pg.100 – 104

A Study on DNA based Information Security Methodologies

Padmapriya M.K
Dept. of Computer Science, NHCE, India
padmapriyabhat@gmail.com

Abstract*— Information security is a process of shielding the information regardless of how it is being formatted, processed, transmitted or stored. Information Security is very much crucial in today's digital world of e-commerce and e-business. Security mechanisms must be good enough to deal with the ever-changing security demands. To provide security to information there are various algorithms of traditional cryptography and steganography. But all these encryption algorithms are not secure enough to provide better security as compared with the today's security requirements. DNA cryptography was a major breakthrough in the field of data security which uses Bio-molecular concepts and gives us a new hope of unbreakable algorithms. This paper discusses various DNA based Cryptographic methods.

Keywords: DNA, DNA Cryptography, DNA chip, PCR, Central dogma of molecular biology

I. INTRODUCTION

Cryptography is a science of securing the information from various types of attacks by converting it into a non readable form. There are various cryptographic techniques like AES, DES, IDEA, RSA etc. The advancement of encryption algorithms depends on human astuteness. Recent researches have shown that DNA based information security as an attractive field.

DNA cryptography is a rapid emerging technology for hiding the data in terms of DNA sequences and it works on the principles of DNA computing. In the cryptographic technique, each letter of the alphabet can be converted into a different combination of the four nitrogen bases of the human deoxyribonucleic acid (DNA). DNA computing was invented in the year 1994 by Leonard Max Adelman for solving the complex problems such as the directed Hamilton path problem and the NP-complete problem. Various researches have shown that DNA can also be used for storing and transmitting the data.

DNA strands are long polymers consisting of millions of linked nucleotides. These nucleotides are made up of one of four nitrogen bases, a five carbon sugar and a phosphate group. The four bases are Adenine (A), Cytosine (C), Guanine (G) and Thymine (T). The DNA strands are named after these four nitrogen bases. DNA strands are complementary in nature. A is complement to T and C to G. Complementary strands bond together to form a double helix structure of DNA as shown in Fig 1. DNA computing uses the double helix structure of DNA. Various researches have shown that the sequence of nucleotides in DNA (A for 00, C for 01, G for 10, T for 11) can be used for cryptographic purpose. DNA stores an enormous amount of information inside the minute nuclei of living cells. It consists of all the instructions required to make every living organism on earth. The main advantages of DNA computation are miniaturization and parallelism. DNA can store very large amounts of data in a compact volume.

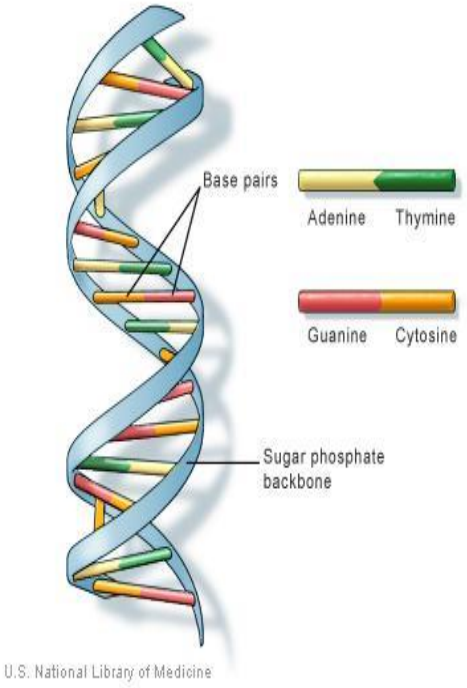

Fig 1 Double helix structure of DNA

II. RELATED WORK

Adelman [1] introduced and demonstrated the first DNA computing, which used DNA to solve a directed Hamiltonian path problem. It initiated a new stage in the era of information with the help of DNA Computing.

KamaljitKainth et .al [4], have build a comparison between old cryptographic algorithms and their issues and it also provides an overview of different approaches used in DNA Cryptography. Also they have discussed about the advantages of DNA cryptography.

Mohammadreza Najaftorkaman et al [5], proposed a novel method to encrypt data by using DNA-based cryptography. In their study, DNA coding technology is used to convert binary data to DNA strings. Finally, they have evaluated the DNA cryptography algorithm's strength based on the properties of DNA strands and probability theories.

L.Jani et al [6], the secret images are encoded into DNA sequences and then shuffled based on the DNA addition rules. The resulting scrambled images are encrypted into n shares using Shamir's polynomial. The scrambled images are retrieved by using the DNA subtraction rules, and the images are recovered using the Lagrange interpolation.

Siddhant Shrivastava et al [7], have described about the various DNA based data storage techniques, their advantages, their flaws, the need for DNA storage, and how it will ultimately become a paradigm shift in computing.

Vijay Dhameliya et al [8], have developed an efficient technique to encode the data into DNA by using non-linear family of ternary codes. Using their method one can store 1.15 ExaBytes (EB) of information in one gram of DNA.

Shalin Shah et al [9], have developed a software called DNACloud which makes it easy to store the data on the DNA. In this work, they have presented detailed description of the software. By considering the current rate of data explosion, demonstrated that DNA storage becomes absolutely indispensable data storage medium because of its low maintenance cost, high data density, ecofriendliness and durability.

Noorul Hussain et al[10], proposed a novel technique for DNA encoding of plain text using string matrix operation and also discussed about the requirements that is to be fulfilled by DNA encryption algorithm.

All of these researches demonstrated that it is possible to construct cryptosystems or perform cryptanalysis based on DNA technology.

III.DNA BASED CRYPTOGRAPHIC TECHNIQUES

DNA cryptography methods use different techniques like Bio-molecular structure, PCR, One Time Pad, Central dogma of molecular biology etc for encrypting the data.

A. Bio-Molecular Structure

All living creatures' information is encoded in the DNA. The complete set of information in an organism's DNA is called genome and it carries the information for all the proteins that the organism will synthesize. The genome can contain incredible amount of information. For example, if we write out the complete Nucleotide sequence of a human genome it would fill more than a thousand books. In addition to that it carries the instruction for about 30,000 distinct proteins. DNA sequence is unique for every living organism. This highly randomness and incredible information storage nature of DNA are the greatest strength for developing a strong encryption algorithm.

B. Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) is a biological process in which certain regions of DNA sequences can be amplified using enzymatic replication. It makes use of Primer pairs where primers are the small DNA fragments. A DNA sequence is encoded using a secret primer pair as a key. The message which should be secured is placed between the two primer pairs to generate a new DNA sequence. By using primers it is possible to generate billions of different DNA sequence. PCR includes the following operations:

1. The double stranded DNA is separated into two single stranded DNA. This process is known as denaturation.
2. For each DNA strand a primer key is selected which sticks to the one end of the DNA strand using annealing process.
3. An enzyme called DNA polymerase is added which can read the bases on template strand and can create its exact compliment.

Fig. 2 represents the PCR amplification process of DNA.

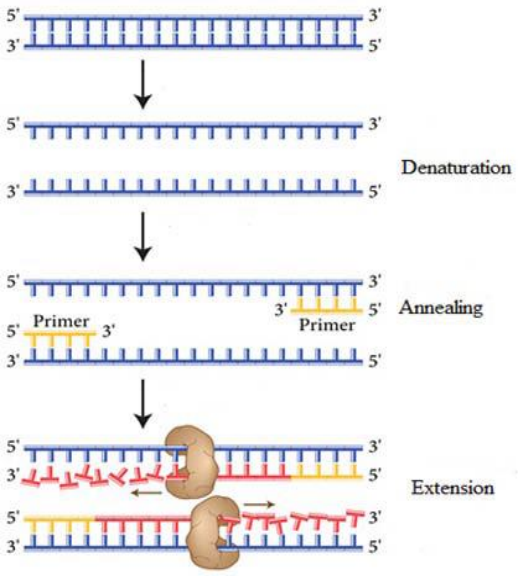

Fig. 2 PCR amplification process [12]

C. Central dogma of molecular biology

Central Dogma of Molecular biology is a process of converting DNA sequence into corresponding protein. It consists of two major operations called as transcription and translation. Transcription is the process in which the DNA strands are converted into RNA strands and in Translation the RNA strands are converted into protein sequence as shown in Fig.3. The plaintext message can be converted into protein by using various complementary rules. Protein form of cipher text can be transferred through public channel and also the size of protein form message will be smaller when compared to original message.

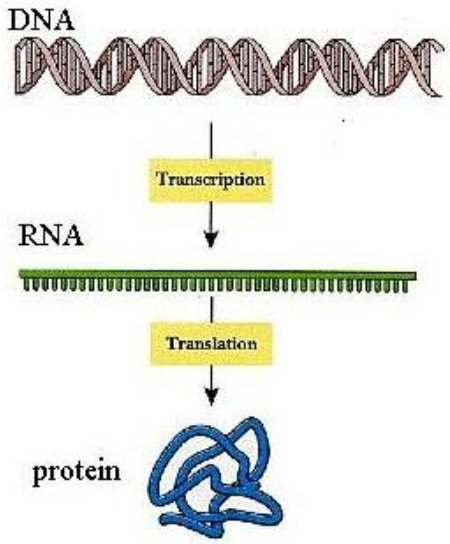

Fig.3 Central dogma of molecular biology [11]

D. One Time Pad (OTP)

In this technique, a plaintext is paired with a random secret key which is called as one time pad using some modular arithmetic operation. If the pad used is long, random and is never used in whole or part, then the resulting cypher text cannot be broken. This technique can be implemented easily using DNA as it provides more compact storage media. Small amount of DNA is enough for huge one-time-pad.

E. DNA chip technology

The DNA chip which is also called as microarray, consists of an ordered srequence of microspots. Each microspot contains thousands of short, synthetic, single-stranded DNA sequences called probes. Using this technology, enourmous parallel operations can be achieved. This advantage can be exploited for developing a crptosystem. While manufacturing microarrays, probes must be selected and it can be directly chosen from available databases like Gene Bank, UniGene etc. Few researches have shown that DNA chip technology is not only limited for encrypting the text data but also useful for encrypting images.

IV. CONCLUSIONS

The main objective of the study on DNA cryptography is to investigate the characteristics of DNA molecule, their reaction and to set up corresponding theories in order to provide the possible development directions for realizing DNA cryptography. DNA Cryptography can have special advantage for secure data storage, authentication, digital signatures, steganography, and so on. With the apt kind of setup, it has the potential to solve huge mathematical problems.

REFERENCES

[1] Adleman L M. *Molecular computation of solutions to combinatorial problems*. Science, 1994, 266: 1021–1024.
[2] Gehani A, LaBean T, Reif J. *DNA-based cryptography*. Lecture Notes in Comput Sci, 2004, 2950: 167–188.
[3] Boneh D, Dunworth C, Lipton R J. *Breaking DES using a molecular computer*. In: Lipton R J, Baum E B, eds. DNA Based Computers: Proceedings of a DIMACS Workshop. American Mathematical Society, 1996. 37–65.
[4] "*A Review to an Invincible Cryptographic Approach: DNA cryptography*", KamaljitKainth, Gurpreet Singh, IJARCCE,Vol. 4, Issue 1, January 2015.
[5] "*A Method to Encrypt Information with DNA-Based Cryptography*", Mohammadreza Najaftorkaman, Nazanin Sadat Kazazi, IJCSDF , 2015 (ISSN: 2305-0012).
[6] "*DNA based Multi-Secret Image sharing*", L.Jani, Anbarasi, G.S.Anandha Mala, Modigari Narendra,, 1877-0509©2015 The Authors. Published by Elsevier.
[7] "*Data Storage in DNA*", Siddhant Shrivastava and Rohan Badlani , BITS Pilani, International Journal of Electrical Energy, Vol. 2, No. 2, June 2014.
[8] "*On Optimal Family of Codes for DNA Storage*", Vijay Dhameliya Dixita Limbachiya , Madhav Khakhar , Manish K Gupta, arXiv:1501.07133v1 [cs.IT] 28 Jan 2015.
[9] "*DNA Cloud: A Tool for Storing Big Data on DNA*", Shalin Shah, Dixita Limbachiya and Manish K. Gupta, arXiv:1310.6992v2 [cs.ET] 16 May 2014.
[10] "*A Novel String Matrix Data Structure for DNA Encoding Algorithm",* Noorul Hussain , Ubaidur Rahmana, Chithralekha Balamurugan, Rajapandian Mariappan, 1877-0509 © 2015 The Authors. Published by Elsevier.
[11] www.design4evolution.net352 .
[12] Anitha M A, Akshatha M A, Vandana B, "*A review on DNA based Cryptographic techniques*", ISSN:2319-7064,Volume 3, Issue 11, IJSR.

