DNA Computing As A Boon In The Field of Information Security

Sangita Vishwakarma[1]
[1] Assistant Professor, Department of Computer Science
[1] Maharaja Agrasen International College,Raipur, (C.G.)

Abstract*- Silicon microprocessors have been the heart of computing for more than four decades. The manufacturers of computer chips are continuously making betterment in the speed and performance of microprocessor chips by integrating more and more devices onto the microprocessor and thus miniaturizing the chip. But there is a limit on this miniaturization and it has been predicted that Moore's law will cease to be obeyed and in near future they would need a new material that could complement the current computing speed and performance of the silicon chips along with equal or fewer complexities. Scientists have found the material that might become the foundation of next era of computing. And the material is DNA, the basic element of our genes. There are numerous benefits that DNA computing offers over conventional silicon based computing. They have enormous storage capacity which is much larger than that of the conventional computers and the enzymes and biological catalysts act as software for executing the required tasks. They exhibit massive parallelism that makes the computation of complex problems much faster than that can be done on conventional silicon based computers. DNA computing has achieved great success in almost every field it has been applied like biomedical, pharmaceutical, information security, cracking secret codes, etc. One such field is Information Security because security of data has always been the matter of great concern. The data being transmitted is under great threat of being attacked and the network carrying data does not have inherent security. Most of the modern cryptographic algorithms are broken, so the DNA computing has brought a new hope in the direction of development of unbreakable algorithms. In this paper the principles of DNA computing and use of DNA computing in the areas of secure data transmission has been outlined. Though DNA computing has almost been successful, the constraints on its implementation are very much demanding like high tech laboratories, labor intensive extrapolation, computational limitations, etc., that moves it far away from being efficiently implemented in today's security world. The aim of this paper is to give a detailed view of DNA computing that could provide a better environment for secure data transmission across networks and the challenges that this technology is facing will be discussed.*

Keywords- DNA, DNA Computing, Cyptography, Algorithms, Molecules.

I. INTRODUCTION

DNA computing is a novel interdisciplinary research area that simulates bio-molecular structure of DNA & computes by means of molecular biological technology. It combines the techniques of biology, chemistry, mathematics and computer science. One of the main goals of this research area is to develop computers which will be biologically inspired & based on DNA molecules which might replace silicon based computers or at least complement them. In DNA computing, strands of DNA are used to represent data. When we compare the execution time of a DNA reaction to the speed of silicon based processor, we'll find that it is much slower but the massive parallelism feature present in it can be used to solve NP-complete and NP-hard problems. DNA computing was first demonstrated by Adleman in his study as a proof of concept that solved Hamiltonian Path problem. Since then many advancements have been made in this field. Molecules of DNA try different possibilities of a problem at once which is the main reason behind its parallelism. Computation with the assistance of DNA introduces a completely new paradigm in the field of computing. In the recent years it has become an exciting area of research but still there is a long way to implement DNA computing in real life. The scientists and researchers are continuously devoting their efforts in developing models and algorithms for DNA computers.

II. DNA & ITS MOLECULAR COMPONENTS

Before understanding the application of DNA in data security, its basic structure should be understood. Each organism on this planet is made up of same type of blueprint. The way in which this blueprint is coded differentiates one organism from the other. DNA (Deoxyribonucleic Acid) is a nucleic acid found in the cell of every living organism that contains all the information and instructions for the growth of any organism and is passed from generation to generation. Its main role is to provide the storage medium for all genetic information that acts as the building block and major source of information for growth and development of any living organism.

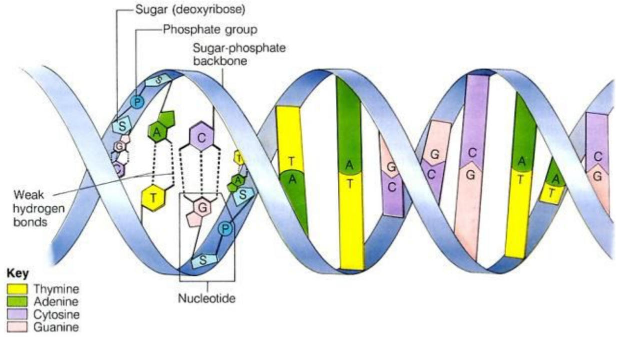

Fig.2. Basic DNA Structure

DNA is basically a polymer which is a collection of various monomers, each monomer is called nucleotide and each nucleotide contains a base. It is double stranded helix of these nucleotides. Each strand of DNA is a long polymer linking millions of nucleotides. A nucleotide consists of one of four nitrogen bases, a five carbon sugar and a phosphate group. There are four different nitrogen bases: Adenine, Guanine, Cytosine and Thymine abbreviated A, G, C and T, respectively. While modeling DNA mathematically, it is represented as X = {A, G, C, T}. All nucleotides differ from each other in terms of their bases. These nucleotides combine in such a way that Adenine is paired with Guanine resulting in Purines and Cytosine is paired with Thymine resulting in Pyrimidines. These combination of nucleotides in the extensively long polymer results in billions of combinations in DNA structure because of which there exists an extensively large variety of living things on this planet ranging from small (mammals as well as plants) to large.

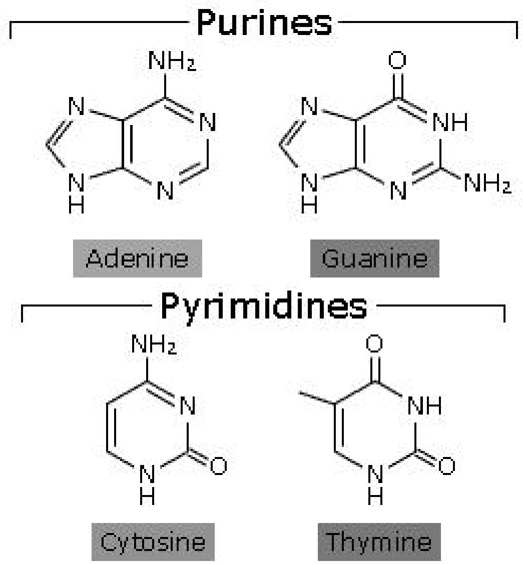

Fig.2. Combinations of Bases Forming Purines & Pyrimidines

The two strands of DNA run anti-parallel to each other. This ability of DNA to bind its pair of strands together forms the basis of its exploitation in various application and is known as Watson-Crick complementarily.

III. DNA COMPUTING

The field of DNA Computing has risen in the past decade. The double helix structure of DNA molecule and Watson-Crick rule form the main principle of DNA Computing. DNA computing or molecular computing are terms used to describe utilizing the inherent combinational properties of DNA for massively parallel computation. The idea is that with an appropriate setup and enough DNA, one can potentially solve huge mathematical problems by parallel search. Basically this means that you can attempt every solution to a given problem until you came across the right one through random calculation. Utilizing DNA for this type of computation can be much faster than utilizing a conventional computer, for which massive parallelism would require large amounts of hardware, not simply more DNA. Leonard Adleman, a computer scientist at the University of Southern California was the first to pose the theory that the makeup of DNA and it’s multitude of possible combining nucleotides could have application in brute force computational search techniques. In early 1994, Adleman put his theory of DNA computing to the test on a problem called the Hamiltonian Path problem or sometimes referred to as the Traveling Salesman Problem. The ‘salesman’ in this problem has a map of several cities that he must visit to sell his wares where these cities have only one-way streets between some but not all of them. The crux of the problem is that the salesman must find a route to travel that passes through each city (A through G) exactly once, with a designated beginning and end. The salesman wants to make efficient use of his time and does not want to backtrack or double back on a path he has already taken previously.

IV. CHALLENGES POSED BY DNA COMPUTING TO TRADITIONAL CRYPTOGRAPHY

The cryptographic algorithm is usually based on complex mathematical problems such as RSA algorithm. Once these mathematical formulae are broken, it gets easier to break the algorithms. But DNA computing provides a parallel processing at molecular level by introducing new data structures. It poses new challenges to the traditional cryptographic field. A number of algorithms have been proposed to attack a number of problems.

A. Challenges to DES

DES is a cipher which based on a Symmetric-key algorithm that uses a 56-bit key. DES is now considered to be insecure for many applications. This is chiefly due to the 56-bit key size being too small. Dan Boneh constructed DES liquid that can break DES within a day. It has been claimed that any symmetric system under 64 bits can be broken with this method. The process to solve this kind of problem is listed as follows: Firstly, encode appropriate binary codes, create initial DNA liquid which contains all possible keys; Secondly, carry out 16 wheels of encryption after pasted known plaintext strands respectively. Lastly, find the solution by searching. Though this idea was simple in theory, the practical operation and execution is not that easy because binary system is completely abstract.

B. Challenges to RSA

RSA is a public-key cryptographic algorithm. The security of the RSA cryptosystem is based on two mathematical problems: the problem of factoring large numbers and the RSA problem Weng-Long Chang have designed integer factorization way for utilizing DNA computing, which can break RSA. Beaver analysed 1000 bits RSA and concluded that to solve HPP problem required the acme number to be 106 at least, namely 10200000L liquid to be needed on the grounds of conservative estimation but it is infeasible. For this, Winfree came up with the idea of computation by self-assembled tiles since DNA tiles can be more easily programmed to incorporate the constraints of a given problem.

V. DNA CRYPTOGRAPHY

In this paper, the research conducted by a number of authors related to the discipline of DNA Cryptography has been studied and has tried to find out the basics of DNA Cryptography that how DNA cryptography field emerged and how DNA computation can be used in cryptography for encrypting, storing and transmitting the information. It has been shown that how DNA cryptography uses DNA as the computational tool with molecular techniques to manipulate it with various algorithms for encryption.

A. Advantages of DNA Cryptography

The biggest advantage of cryptography is its secure nature although; it never needs to be transmitted to anyone.

1. Moreover, encrypting along with the DNA sequence makes data more secure. One gram of DNA contains 10^21 DNA bases = 10^8 tera-bytes of data. A few grams of DNA can hold all the data stored in world.

2. Since DNA is used for encryption, Signature authorization is not needed. DNA replaces the cause of Digital signatures.

3. Works in a massively parallel fashion: DNA is modified biochemically by a variety of enzymes, which are minute protein machines that read and process DNA according to nature's design. There is a wide variety and number of these "operational" proteins, which manipulate DNA on the molecular level. Just like a CPU has a basic set of operations like addition, bit-shifting, logical operators (AND, OR, NOT NOR), etc. that allow it to execute even the most complex calculations, DNA has cutting, copying, pasting, repairing, and many other capabilities.

4. Large storage: A gram of DNA contains about 10^21 DNA bases, or about 10^8 tera-bytes of data. Hence, a few grams of DNA have the capability of storing all the data stored in the world.

5. Input and output of the DNA data can be moved to conventional binary storage media by DNA chip arrays.

6. The main goal of the research of DNA cryptography is exploring characteristics of DNA molecule and reaction, establishing corresponding theories, discovering possible development directions, searching for simple methods of realizing DNA cryptography, and laying the basis for future development.

B. Limitations of DNA Cryptography

Apart from advantages, DNA cryptography has few disadvantages. They are:

1. Lack of the related theoretical basis.

2. Difficult to realize and very expensive to apply.

VI. CONCLUSION

DNA cryptography is basically hiding of data in terms of DNA sequences. This is done by using various DNA technologies with the biological tools. In this paper we summarized basics of DNA and basics of where DNA is found are discussed. We have also discussed the factors on which DNA cryptograph differs from traditional cryptography. Few of the advantages and disadvantages have been summarized. Later on the techniques used by the DNA computing to break the existing cryptographic algorithms have been studied. We want to conclude that although DNA computing has broadened the view of people towards natural phenomena of computing but it is still in its theoretical stage and moving it towards the practical stage will require huge time, enormous computation and large number of expertise. The future work in this field will consist of analyzing deeply the performance of all the DNA cryptographic techniques based on secure data transmission processes.

REFERENCES

[1] Donald Nixon, “DNA and DNA Computing in Practices – Is the Future in our Genes?”, Global Information Assurance Certification Paper

[2] J. Clerk Maxwell, “Integrating DNA Computing in International Data Encyption Algorithm (IDEA)”, International Journal of Computer Applications (0975 – 8887), Volume 26– No.3, July 2011

[3] Sanjeev Dhawan, Alisha Saini, “Secure Data Transmission Techniques Based on DNA Cryptography”, International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

[4] Harneet Singh, Karan Chugh, Harsh Dhaka, A. K. Verma,,“DNA based Cryptography: An Approach to Secure Mobile Networks”, International Journal of Computer Applications (0975 - 8887), Volume 1 – No.19.

[5] Guangzhao Cui, Cuiling Li, Haobin Li, Xiaoguang Li, “DNA Computing and Its Application to Information Security Field”, 2009 Fifth International Conference on Natural Computation.

[6] Junzo Watada, “DNA Computing and its Application”

[7] Rohani binti abu Bakar, Junzo Watada, “DNA COMPUTING AND ITS APPLICATIONS: SURVEY”, ICIC Express Letters, Volume 2, Number 1, March 2008