Volume 4, Issue 5, May 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Comparative Study on DNA Cryptography

Yesha Pruthi* PDMCEW,MDU
Sunita Dixit PDMCEW,MDU (Assistant Prof.)

Abstract— ***The new promising field in cryptography which emerged with the process of DNA computing is DNA cryptography .For cryptographic purposes the concept of large information density in DNA molecule and massive parallelism is exploited. D NA computing is a very hopeful direction and its theories can be applied in cryptography. It is applied in many fields and solves some hard problems successfully. This paper uses a new way to show how DNA computing works with cryptography, can transmit message effectively and securely. A new parallel cryptographic technique is proposed which certainly minimizes time complexity using DNA molecular structure, DNA hybridization and one time pad scheme. The computational complexity and requirement of high tech biomolecular laboratory are currently the main difficulties of DNA cryptography. An asymmetric key algorithm is used in this paper connecting with DNA computing technique to encrypt message.***

Keywords— DNA cryptography, single stranded DNA, double stranded DNA, hybridization

I. INTRODUCTION

DNA computers all together come in a different dimension than the silicon computers. They have a totally new method to perform calculations which involves test tubes; enzymes etc. and incorporate biology, chemistry, biotechnology and finally computer science. New directions of information security are being required to protect the data as some of the modern cryptographic algorithms such as DES, MD5 are broken. To bring forward a new hope for powerful or even unbreakable cryptography algorithms the concept of using DNA computing in the field of computer security is a possible technology.

In 1994 Adleman, proposed DNA for computations. After that many approaches have been investigated. Theoretical consideration that it may simulate turing machines and cryptography. It has been shown that DNA computing was suitable for some problems currently hard to resolve and could work more faster than electronic computer. A good background between the DNA molecule and computer engineering are required to develop the efficient algorithms of DNA computing. After Adleman solved the Hamilton Path Problem using a combinatorial molecular method, many hard computational problems were calculated by DNA computer. In his pioneering work, the base for the new field of biomolecular research was build by Adleman.

DNA computing and parallel computing are fundamentally similar. To try many different possibilities at once it takes advantage of the many different molecules of DNA. For some specialized problems, DNA computers are the smaller and the faster than any other computer built so far. Its vast parallelism exceptional energy efficiency and extraordinary storage capacity is the main advantage behind the DNA molecular structure. About 10^8 bytes can be stored in one terabyte of DNA. Inspite of having such bright future towards cryptography it is confronted with some drawbacks which includes high computational complexity, huge computing time and high tech bimolecular laboratory. In general modern biological techniques are used as implementation tools and DNA works as information carrier. Hybridization, synthesizing message DNA and PCR amplification are some of the biological techniques used but these methods are complex, costly and need enormous amount of time.

Its just the initial stage of DNA cryptography and many challenges are yet to be faced. The reason as to why only few algorithms on DNA cryptography are processed till now as many researchers are working on this field it is far from maturity both in theory and realization. DNA technology is based upon modern biological technologies and is much laboratory dependent.

In this paper a new parallel DNA cryptography technique is investigated using DNA molecular structure which certainly minimizes the time requirement. The cipher text is generated in encryption method from plain text using OTP (one time pod) scheme in a severe way and the modified version of plain text by the position of binary 1's. Hybridization technique is used for decryption process which makes use of randomly received individual packets. In this paper the following contents, section II presents the DNA design rationale, section III contains a general overview of cryptography followed by DNA cryptography in section IV and a formal description of the proposed method is described in section V.

II. DESIGN RATIONALE

The hereditary material of almost all living organisms ranging from a complex human being to a very small virus is deoxyribonucleic acid. The kind of molecule which can form the genetic instructions is DNA. It is a long polymer of small units called nucleotides. Genes are the fragment of DNA which has genetic message. For other DNA sequence,

some still take roles by their structural directly, another will involve in and regulated the genetic information's performance.

Each nucleotide has three basic components:

1. a phosphate group
2. a nitrogenous base
3. a five carbon sugar

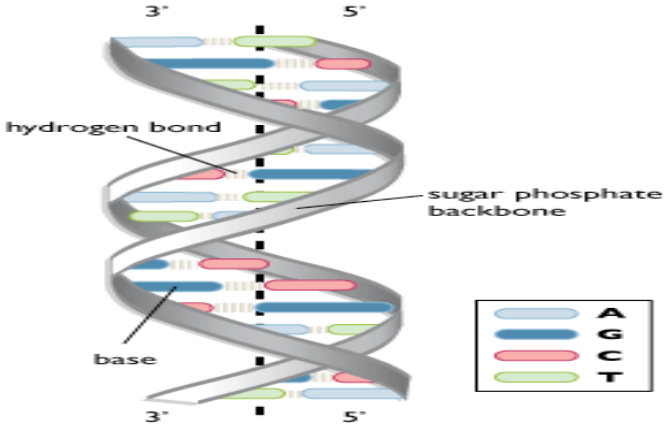

Fig 1. Double helical DNA structure showing its basic components

Depending upon the type of nitrogenous base they have got, there are four different nucleotides. In addition the two pureness and pyramiding are A ,T ,G and C i.e Adenine, Thiamine, Guanine and Cytosine respectively. There is a principle called Watson Crick base complementarily principle between the connection of these bases: A and T matches, C and G matches. DNA has a double helical structure with two strands running anti parallel. With the combination of only these four letters A, T, G and C DNA stores all the complex and huge information about an organism. The structure of DNA strands are formed by these bases by forming hydrogen bonds with each other to keep the two strands intact. A forms hydrogen bond with T whereas C and G forms with one another.

Using the following phenomena DNA sequences are responsible for transfer of complex information through the process of transcription DNA provide information to message RNA then message RNA transfers the information protein by the process called translation. From information transfer from one age group to another age group, these processes play an important role.

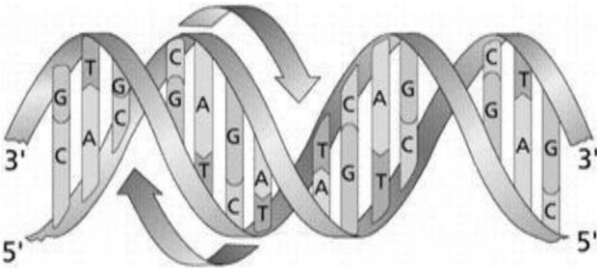

Fig 2. Double helical structure of DNA with four nucleotides

In this paper we denote individual stranded as single stranded DNA and a double helix as double stranded DNA. Under certain conditions a single stranded DNA can form double stranded DNA with other single stranded DNA which are complimentary.

III. CRYPTOGRAPHY

The security of information and communication becomes more and more important with the rapid development of information technology. The kernel of whole information security technology is cryptographic technique. In earlier times cryptography used to be executed by using manual techniques but now the basic frame of performing cryptography has remained just about the same as past. Most importantly, these cryptographic algorithms and functions are accomplished thus making the ways a cot secure and faster.

Cryptography is necessary as a means for communication over any untrusted medium for telecommunications and data, which includes any network or particularly any systems. The main goal of cryptography is not only to protect data from being attacked or stolen rather it can be used for authentication for users. There are three types of cryptographic schemes which symbolisticaly achieve these goals.

1. Public key cryptography (asymmetric)
2. Private key cryptography (symmetric)
3. Hash functions

The unencrypted data is referred to as plain text which is then encrypted into cipher text. In recent years public key cryptography is considered the most significant development. It is also called asymmetric key cryptography as two keys are used differently. The one key is used for encryption which is referred to as public key and the other key is used for decryption which is referred to as private key.

The most important and proven asymmetric key cryptographic algorithm is RSA algorithm. The RSA algorithm is based on mathematical fact which is easy to get and multiply large prime numbers together. The keys in RSA i.e public and private are based on extremely large prime numbers. This is quite easy algorithm. The real major challenge for RSA is the generation and selection of public keys and private key or else it can easily be cracked by an attacker.

The following description of the figure shows the working of RSA algorithm:

1. P and Q are two random large prime numbers.
2. E is a public key and D is a private key. When A sends the message to B's public key, then afterwards B can decrypt cipher text to plain text using B's private key.
3. CT is cipher text and PT is plain text.

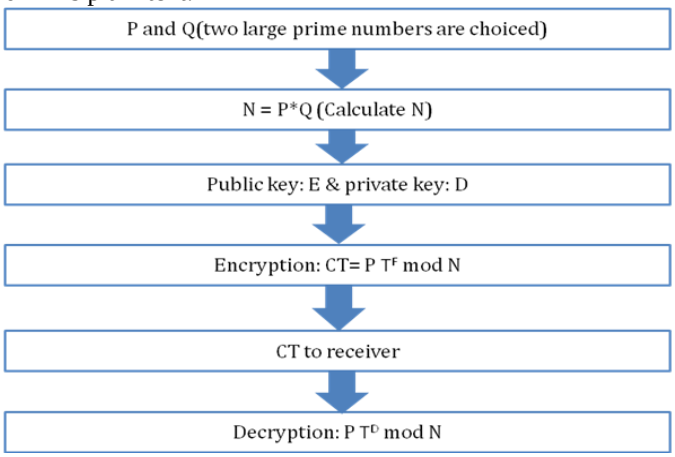

Fig 3. RSA Algorithm

IV. DNA CRYPTOGRAPHY

A new way to cryptography is given by DNA computing. The powerful parallelism capability in DNA computation takes new challenges for modern cryptography. Cryptography is not only a critical aspect of traditional computing but also important to DNA database application. DNA cryptography which is a new born cryptographic technique is the one in which DNA is used as carrier of information and the modern biological technology is used as implementation tools for all sorts of cryptographic technique the vast parallelism and extraordinary information density are inherent in DNA molecule.

In today's modern time, the field of biology and that of security, cryptography have come to combine. DNA cryptographic systems are based on DNA and one time pads and if they are implemented correctly it becomes virtually impossible to crack the system. The size of one time pad depends on the cryptographic systems. Various procedures are there for DNA one time pad encryption scheme. DNA is very powerful from the cryptographic point of view.

A great opportunity is offered by binding capabilities of nucleotide bases (AT and GC) for creating self assembly structures that are excellent means of executing computations. The major advantage of using DNA is that it has huge storing capacity but its practical implementation requires a lot of resources and time.

DNA also has several computational limitations therefore efficient use of DNA cryptography is still difficult if seen from practical point of view. This paper provides overview of basic technology used cryptography merged with DNA computing. Let us assume that a sender sends a message to a receiver, in the sending process no one could read the message. Initially the message is written in plain text with a list of alphabets. Encryption is the art of confusing the plain text message and transforming it into encrypted message known as cipher.

Combining the asymmetric key cryptography of RSA algorithm with DNA computing technology to show how to encrypt a message and how the processing of an algorithm works. The figure demonstrate the algorithm.

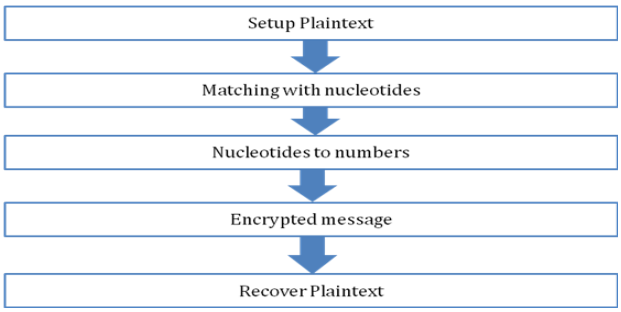

Fig 4. Algorithm

V. PROPOSED WORK

We now describe the proposed method where difficult biological problem such as DNA, hybridization and one time pad scheme are used whose security is major concern. To decrypt the message which certainly minimizes the time taken for decryption this method uses parallel technique a safe way of exchanging messages between sender and receiver will be shown. Lets say the sender is Alice and the receiver is Bob the definition of encryption and decryption can be extended as follows:

Let us suppose that Alice (sender) owns an encryption key K_A and sends it to Bob (receiver). Alice uses K_A to translate a plain text M into cipher text C by a translation E. Bob uses the same encryption key K_A to translate the cipher text C into plain text M by the translation D. Then the encryption process is given by:

$$C = E\ K_A\ (M)$$

And the decryption process is given by:

$$D\ K_A\ (C) = D\ K_A\ (E\ K_A\ (M)\ ) = M$$

It is difficult to obtain M from C until one has K_A .He call the translation E as encryption process, translation D as decryption process and C as cipher process. Here K_A can be any method, material, data, DNA sequence etc. and not just limited to digital data. Similarly E and D can be any physical or chemical or biological or mathematical processes and not just limited to mathematical calculations. We now describe the general process of the proposed cryptographic technique.

A) Key generation :-

The most simplest and secured type cipher is one time pad for each encipherment. A one time pad cipher uses a pad or key that is chosen randomly. Hence there is no way out that any adversary can guess/detect the key. Through secure channel prior to encryption process Alice sends the key to the receiver Bob, Alice (sender) first encrypts the message using the key and then destroys the key. Similarly Bob uses this key for decryption and destroys the same key. Next time a new key is generated for encryption and decryption -> One time pad key are used as a randomly generated single stranded DNA string is generated by the message sender Alice and transmit it to Bob using a secure channel. The key length depends on the length of the plain text which will be transmitted.

B) Encryption:-

In the encryption method the randomly generated single stranded DNA is used as key K_A. In reverse order the single stranded DNA key is scanned from sequence end towards the beginning of the key.Firstly plain text M is translated to the corresponding ASCII code by the sender. Then this code is converted into binary plain text M. finally the binary plain text M is translated into several packets of DNA cipher text C_i using the algorithm.

C) Decryption:-

The intended receiver Bob takes the Waston-crick complementary of the packet content after getting the DNA cipher text packets from Alice. The probable starting position for matching the single stranded DNA substring with the single stranded DNA key is determined by the packet number which is attached with the packet. So by this way the decryption algorithm will be able to minimize the searching time. The whole decryption process is done in parallel to parallel computing environment. When Bob receiver a cipher text packet from Alice, to evaluate each packet and to find out the substring's exact matching position in the single stranded DNA key it is assigned to the processor. In this way the Bob gets the binary plain text message M which he converts into ASCII code to get the actual plain text M.

VI. CONCLUSION

In this paper by using DNA digital coding technique and DNA hybridization we presented an original DNA cryptography technique. One time pad was used as an encryption key. From the analysis it can be concluded that the proposed DNA cryptography method promises to be a good solution for implementation in secure network. Further, in multicore environment this method can be implemented. The issue that still arises is its increasing computational complexity, which can be worked upon in future.

REFERENCES

[1] L. Adleman, Molecular Computation of Solutions to Combinatorial Problems", Science266:1021-1024 (Nov. 11) 1994.

[2] D. Boneh, R. Lipton, \Making DNA Computers Error Resistant", Princeton CS Tech-Report CS-TR-491-95 in proceedings of second annual conference on DNA based computers, Princeton, 1996.

[3] G. Paun, G. Rozenberg and A. Salomaa. DNA Computing: New Computing Paradigms,Springer-Verlag, Berlin, 1998.

[4] R. 1. Lipton. DNA Solution of Hard Computational Problems, Science, 268, 1995, pp. 542-545.

[5] D. Boneh, C. Dunworth, R. Lipton, 1. Sgall, \On Computational Power.

[6] G. Z. Cui, L. Qin, Y. Wang, X. Zhang, "An Encryption Scheme using DNA Technology", *BICTA 2008.*

[7] X. Guozhen, L.U. Mingxin, Q. Lei, L. Xuejia, "New field of Cryptography: DNA Cryptography".

[8] A. Gehani, T. Labean, J. Reif, "DNA-based cryptography", *DNA based computers V. Providence: American Mathematical Society*, vol. 54, pp. 233-249, 2000

