ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

Vol. 4, Issue 9, September 2016

# A Survey on Diverse DNA Cryptographic Techniques

Nileena Ouseph

M.Tech Student, Dept. of C.S.E., Mar Baselios College of Engineering and Technology, Trivandrum, India

**ABSTRACT**: Cryptography is a science of applying complex logics and mathematical procedures to develop efficient encryption techniques. It can also be defined as an art that allows people to keep their data confidentially in the world of electronics. DNA cryptography is a new and upcoming domain in the field of security, utilized by the researchers to have better and secure communication over a network. It emerged with the advancement of DNA computing. Cryptography exploited certain properties intrinsic in the DNA molecule such as the extensive parallelism and massive information density.

The main objective of DNA cryptography is to acquire higher confidentiality and integrity while sending data or any other information over a network, thereby protecting them from brute force attack. At present, the limitations faced by DNA cryptography is the need of bio-molecular laboratories and difficulties in computations. This paper proffer an analysis of various schemes used in DNA Cryptography.

**KEYWORDS**: Encryption, Decryption,Pain text, Cipher text, Integrity, Confidentiality, Authenticity, DNA.

## I. INTRODUCTION

In this era of rapid advancement and information explosion, data has become a very dominant strategic resource. Data plays a major role in many fields, especially in confidential business and military affairs. Some of these data may be highly sensitive and thus will be prone to unauthorized access.

Automated tools are necessary for protecting files and other data stored on the computer. With the advent of time-sharing systems, the need for these tools became mandatory. The introduction of distributed systems and network facilities for data transfer between the end user and the system affected the security of data. At the same time the systems that could be accessed over the Internet or a public telephone network needed security as well. Thus during a transaction, confidentiality becomes a requisite factor and security, increasingly significant.

A variety of techniques have been used to keep the unauthorized user away, such as cryptography and data hiding. Cryptography is essential for both network and computer security. It provides authenticity and confidentiality to the data thereby protecting it from being eavesdropped. Cryptography is identified as a procedure of secret writing to protect the information shared between two communicating parties from attacks of the intruder. It can be done by using a cryptographic algorithm for transmuting a message called plain text into cipher text. But since today achieving total security to the data is a challenging issue. Data confidentiality, Integrity and data availability are the three main goals of data security. This iscommonly referred to as the CIA triad [2] shown in fig. 1.

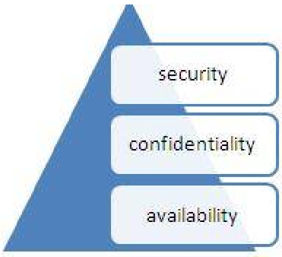


Fig. 1. CIA security triad

Information security experts have focused on nano-cryptography to create more secure systems. When scientists

DOI: 10.15680/IJIRCCE.2016. 0409024

found that binary computers (digital computers) have various physical constraints, especially in data storage and computation processes, they concentrated on DNA computers (bimolecular computers) and quantum computers. DNA cryptography is a new science, of which deoxyribonucleic acid (DNA) is the most important feature. Cryptographers have been working on DNA cryptography to solve the limitations so as to make a system, which is immune to popular attacks like brute-force attacks, dictionary attacks, etc.

Over the past 50 years, DNA, is found to be the basic building block of living organisms. By using DNA, Adleman proposed the solution for Hamiltonian path problem in 1994 [1]. This study resulted in a progress to the field of Bio-computing. Certain properties like density of information and vast parallelism of DNA encouraged many researchers to exploit the area of bio-computing. This approach was later extended by Lipton to solve NP complete problem [1]. Thus DNA computing is providing a brand new data structure and evaluating techniques for the parallel processing capabilities of molecules. DNA cryptography uses DNA as the information carrier and modern biological technology as the implementation tool.

Even though DNA shows a brilliant fortune to the field of cryptography, it is restricted with some drawbacks such as requirement of huge computing time, high computational complexity and high tech bio molecular laboratory.

## II. BIOLOGICAL BACKGROUND

DNA (Deoxy-ribo Nucleic Acid) is the basic building block of the human body which supports the proper functioning and development of all the living organisms. The DNA is made up of genes. As the DNA holds the necessary genetic information, it is used to build molecules like RNA (Ribo Nucleic Acid) and proteins. There are 4 nucleotides in DNA namely Adenine (A), Thymine (T), Guanine (G) and Cytosine(C).The structure of DNA is shown in the figure 2 below.

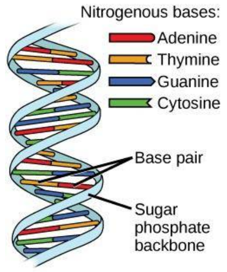


Fig. 2. DNA structure

DNA computing has many advantages over the silicon machines. These advantages mainly include the size, high parallelism and speed of computation. One gm of DNA can store about $10^8$Tera bytes of data. Every molecule of DNA can act as a single processor thus providing parallelism. Operations can thus be done in parallel, which increases the speed of computations. At the same time, DNA is highly energy efficient i.e., $10^{19}$ operations per Joule.

DNA is not only used to transmit and store the information, but also to conduct computations as well. The absence of theoretical proofs and practical procedures are found to be the difficulties faced in this field which causes a hindrance for it to be readily and reliably be implemented in the field of security.

DNA [3] [11] along with the RNA makes new proteins and guides the cell to identify all the biological attributes so as to pass it down to the next. Thus it is responsible for carrying information between generations. This results in inheritance. DNA includes all the necessary information that an organism needs which helps the living growth and development.

Hybridization is a process by which single stranded DNA combines to form double stranded DNA molecules. During this process Adenine pairs with Thymine and Guanine always with Cytosine [2]. The amplification of single or multiple strands of DNA to millions of copies of a certain DNA sequence is called as Polymerase Chain Reaction (PCR). Primer that functions as the beginning point of DNA synthesis, is a strand of nucleic acid. Transcription and Splicing is a process in which a portion of the DNA is copied to the RNA (mRNA). This mRNA is then converted to another set of amino acid called the proteins. This process is known as the translation.

DNA cryptography can be explained as the process of securely hiding data in the form of DNA sequences. This technique make use of biological processes, arithmetic operations or both. Polymerase Chain Reaction, DNA Hybridization, DNA

Fabrication, DNA Fragment Assembly, Transcription, Splicing and Translation are the frequently used biological processes. At the same time it uses arithmetic operations to manipulate the data.

In order to perform arithmetic operations over the data, they are initially converted to its corresponding binary. The binary data is encrypted by replacing them with the associated DNA sequences. The most commonly used DNA digital coding is shown in Table 1. Any number system can be used for encryption like hexadecimal, decimal, octal etc. Addition, subtraction, complement, XOR, substitution, insertion, random number etc. are the most commonly used arithmetic operations.

| DNA nucleotide | Binary |
|---|---|
| A | 00 |
| C | 01 |
| G | 10 |
| T | 11 |

TABLE I. DNA digital coding

## III. SURVEY ON DNA CRYPTOGRAPHY

The area of DNA cryptography is rather an untouched one. Over the years, many initiatives have been proposed to explore the process of DNA cryptography. But few have been implemented. The major development in the area of cryptography is the establishment of DNA computing to the traditional cryptographic techniques. This paper gives an outline of the different DNA cryptographic techniques.

Cui Guangzhao, Limin Qin, Yanfeng Wang, and Xuncai Zhang [4], proposes a new scheme for data encryption that uses PCR primers as the encryption key KA and 'e' used as receivers public key, also a decryption key named KB that includes receivers secret key 'd' and a pair of PCR primers. Encryption includes a process of data pre-treatment. By doing this pre-process operation, one would get an entirely different cipher text from the same-to-same plain text.
DNA digital coding technology is used to get the cipher text by scrambling the plain text. After the coding is done, DNA sequence is synthesized by the sender .This sequence is appended by reverse and forward PCR primers, which is 20-mer oligo-nucleotide in length. During decryption, when the correct receiver receives the DNA assemblage, the person is able to find the sensitive data in the form of DNA nucleotides.Because the apt receiver knows the exact two PCR primer pairs via a protected channel, the receiver on getting the message amplifies the secret message DNA sequence by performing PCR on the DNA assemblage. This method has both mathematical and technical drawbacks which causes an adversary to recognize the original information.

Prabhu D and M. Adimoolam [5] uses Diffie-Hellman key exchange (DH) protocol for Key Initiation and Transmission to jointly establish a shared secret key for communication over insecure communications channel. The plain text is manipulated using logical operation XOR. DNA digital coding is used to pre-process the XOR-ed output into corresponding binary number system. On this data PCR amplification is performed using the primer pairs. When the correct receiver receives the DNA assemblage, the person is able to find the sensitive data in the form of DNA nucleotides. One of the drawback of Diffie-hellman key exchange is that it does not provide authentication to the communicating parties. Thus Diffie-hellman is vulnerable to the man in-the-middle attack.
Majumder, Atanu, AngshulMajumdar, TanusreePodder, NirmalyaKar, and Mukesh Sharma proposed [7] a round key selection encryption method. Plain text is divided into 256-bit blocks after converting it to its corresponding binary values. Each of these blocks will pass through an encryption procedure. Round Key operation includes 256-bit block of plain text being divided into four 64-bit blocks. Ex-OR operation is performed between fourth block of plain text and the round 1 key, K1. This result is then XOR-ed with thethird block of the plain text. Then with the second block and so on. Then these four 64-bit results will go through a straight D-Box. The output of the D-Box will be used as input for the second round. Randomly selected 256 bit is used as key. Decryption is done in a reverse process to obtain the plain text.

Mandge, Tushar, and VikasChoudhary [6] presents an algorithm which uses an initial key, the generated key and a pair of PCR primers. Secure key generation scheme converts Initial key to ASCII Code. The algorithm also includes the random number generation (Prime Random number <255).The remainder and quotient required for the key generation is obtained by:

Rem=Random number / ASCII value [1 to end]. Quo= Random number/ASCII value [1 to end].

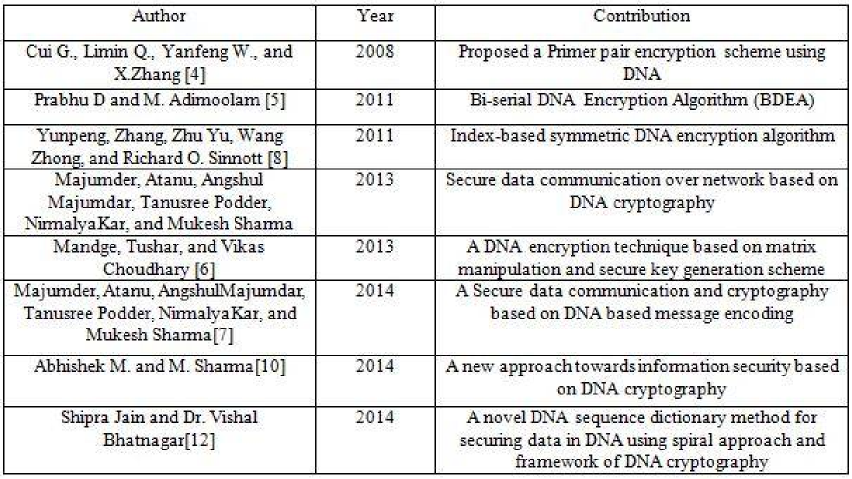
The Generated key is given as Random number, Rem, Quo. Only the generated key is shared with the receiver and not the initial key. Initial key is given by the formula (Random number Remainder [1 to end]) / Quotient [1 to end]. Original plain text undergoes matrix manipulation and logical operations. This results in a mini cipher on which DNA digital coding is performed. Security of this scheme strictly depends on the key.

Yunpeng, Zhang, Zhu Yu, Wang Zhong, and Richard O. Sinnott [8] uses an index based symmetric algorithm which encodes each character into ASCII codes. The ASCII codes will then be converted to the DNA sequences. Besides, a special DNA sequence is selected by the cryptographer as the index for encryption, and likewise, the pre-treated plain text will be divided into different groups. The key used in this index based algorithm is created by the Chaos Key Generator based on the Logistic Then, the result of these processes will be translated to the corresponding DNA sequence. The Decryption proceeds in the reverse way of encryption after the keys are securely transformed to the receiver through a secure/insecure channel.

Majumder, Atanu, AngshulMajumdar, TanusreePodder, NirmalyaKar, and Mukesh Sharma introduces a novel method to provide [6] security on the basis of an algorithm where a long and strong 256- bit key is generated and use of the round encryption.This technique has been carried out using, four round operations between the plain text and the generated key. Later on the resultant cipher text is transformed to a DNA sequence and appends some extra information bits within, to provide more security in the message. A random key is chosen to do the encryption with the DNA base, namely A, T, C, and G. This key acts as the round 1 key. Then the key order of DNA bases is right shifted to 1 block. This process continues 4 times until 4 keys are generated by shifting of 1 block (64 bit) in each round. The overall encryption is done in 4 consecutive rounds. For each round previously generated keys i.e., Key 1, Key 2, Key 3 and Key 4 are used to encrypt the plain text. This CT is mapped into a randomly selected array of 16 characters out of 256 ASCII symbols using a hash mapping technique. At the receiver's end, randomly selected hash function and the DNA sequences are available. Receiver can compute the cipher text using hash mapping function and binary coding scheme.

Shipra Jain and Dr. Vishal Bhatnagar [11] provides security at two levels by using spiral transposition and DNA sequence dictionary table. The spiral motion of the binary sequences interchanges the positions of the data thereby resulting in binary encryption. The spiral motion can be done row-wise or column-wise. For the DNA encryption, the DNA sequence table is used as a key. Each sequence corresponds to the decimal value. The dictionary is then shared among the communicating parties via a secure channel. By secretly sharing the static Dictionary through a secure channel along with the spiral motion, the decryption can be done in the reverse order. The Drawbacks of this scheme is that Static table could be easily compromised which poses a threat to the security.

The survey can be summarized as follows:



| Author | Year | Contribution |
|---|---|---|
| Cui G., Limin Q., Yanfeng W., and X.Zhang [4] | 2008 | Proposed a Primer pair encryption scheme using DNA |
| Prabhu D and M. Adimoolam [5] | 2011 | Bi-serial DNA Encryption Algorithm (BDEA) |
| Yunpeng, Zhang, Zhu Yu, Wang Zhong, and Richard O. Sinnott [8] | 2011 | Index-based symmetric DNA encryption algorithm |
| Majumder, Atanu, Angshul Majumdar, Tanusree Podder, NirmalyaKar, and Mukesh Sharma | 2013 | Secure data communication over network based on DNA cryptography |
| Mandge, Tushar, and Vikas Choudhary [6] | 2013 | A DNA encryption technique based on matrix manipulation and secure key generation scheme |
| Majumder, Atanu, AngshulMajumdar, Tanusree Podder, NirmalyaKar, and Mukesh Sharma[7] | 2014 | A Secure data communication and cryptography based on DNA based message encoding |
| Abhishek M. and M. Sharma[10] | 2014 | A new approach towards information security based on DNA cryptography |
| Shipra Jain and Dr. Vishal Bhatnagar[12] | 2014 | A novel DNA sequence dictionary method for securing data in DNA using spiral approach and framework of DNA cryptography |

DOI: 10.15680/IJIRCCE.2016. 0409024

## IV. CONCLUSION

Since the demand for storage has increased, the need for reliable and safe storage of information also increased as well. This gave a way to the fire current and upcoming technologies for the efficient storage of data. One among them is the DNA which provides efficient data storage capacity. Thus DNA can possibly be used in the field of cryptography and steganography. The paper identifies the different algorithms outlined by the researchers till today. The survey also explains the various arithmetic and biological operations used in the DNA cryptographic algorithms. DNA cryptography aims at achieving the maximum security to the data stored.

## ACKNOWLEDGMENT

I would like to thank God for all the help rendered in choosing and grasping information. I would like to thank our college for helping me with all the resources and my family members who have been very patient during this work.

## REFERENCES

1. Rakheja,"Integrating DNA Computing in International Data Encryption Algorithm (IDEA)", International Journal of Computer Applications, 26(3), pp.1-6, 2011.
2. Cherian, Raj, Abraham, "A Survey on different DNA cryptographic methods", International Journal of Science and Research (IJSR), ISSN, pp.2319-7064, 2013.
3. Cui Guangzhao, Limin Qin, Yanfeng Wang, and Xuncai Zhang, "An encryption scheme using DNA technology", 3rd International Conference on Bio-Inspired Computing: Theories and Applications, BICTA, pp. 3742. IEEE, 2008.
4. Prabhu D and M. Adimoolam,"Bi-serial DNA Encryption Algorithm (BDEA)", IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2011.
5. Mandge, Tushar, and VikasChoudhary, "A DNA encryption technique based on matrix manipulation and secure key generation scheme," IEEE International Conference on Information Communication and Embedded Systems (ICICES), pp. 47-52, 2013.
6. Majumder, Atanu, AngshulMajumdar, TanusreePodder, NirmalyaKar, and Mukesh Sharma, "Secure data communication and cryptography based on DNA based message encoding", IEEE International Conferenceon Advanced Communication Control and Computing Technologies (ICACCCT), pp. 360-363,, 2014.
7. Yunpeng, Zhang, Zhu Yu, Wang Zhong, and Richard O. Sinnott. "Index-based symmetric DNA encryption algorithm", 4th International Congresson Image and Signal Processing (CISP), Vol. 5, pp. 2290-2294. IEEE, 2011.
8. Majumder, Atanu, AngshulMajumdar, TanusreePodder, NirmalyaKar, and Mukesh Sharma, "Secure data communication over network based on DNA cryptography", International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), Vol.3, Issue 4, pp. 193-201, IEEE Oct 2013.
9. Abhishek M. and M. Sharma , " A new approach towards information security based on DNA cryptography" , International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), Vol.4, Issue 4, pp. 59-68Aug 2014.
10. Shipra Jain and Dr. Vishal Bhatnagar, "A novel DNA sequence dictionary method for securing data in DNA using spiral approach and framework of DNA cryptography", IEEEInternational Conference on Advances in Engineering and Technology Research (ICAETR), pp. 1-5, 2014.
11. J. Chen, "A DNA-based, bio molecular cryptography design", IEEE International Symposium on Circuits and Systems (ISCAS), 2003.
12. Javheri, Snehal, and Rahul Kulkarni, "Secure Data communication and Cryptography based on DNA based Message Encoding." International Journal of Computer Applications 98, pp.no. 16, 2014.