International Journal For Technological Research In Engineering
Volume 2, Issue 9, May-2015 *ISSN (Online): 2347 - 4718*

DNA BASED CRYPTOGRPHY FOR SECURE DATA COMMUNICATION

Sarvdeep Kushwaha[1], Ravi Kant[2], Chandresh Pandey[3], Dr. Vijay Kumar Chaurasiya[4]
Department of Master of Science in Cyber Laws & Information Security
Indian Institute of Information Technology, Allahabad (U.P) 211012, India.

Abstract: In the ever-changing and developing world of communication systems we need stronger and efficient security mechanism. Cryptography is a technique of storing or transmitting data in a secured manner so that the legitimate users can have access to it. Cryptography includes the process of scrambling the readable text into unreadable form and vice-versa to make it secure from unauthorized user. The process of converting the plain text into cipher text is called encryption and from cipher text to plain text is called decryption. DNA is the fundamental molecules which is present in every living organism and carries genetic information from one generation to other. We are using the bio-chemical properties of DNA for enhancing security of encryption and decryption. We are proposing a cryptographic technique using the protein synthesis table of DNA.

Keywords: DNA, DNA Cryptography, Encryption, Decryption, mRNA, Nucleotides, Protein, Codons, Binary Number Crunching, Genetic Code Table.

Abbreviations: DNA (Deoxyribonucleic acid), RNA (Ribonucleic acid), mRNA (Messenger Ribonucleic acid), Guanine (G), Adenine (A), Thymine (T), Cytosine(C), Uracil (U)

I. INTRODUCTION

In the era of modern information Technology the domain of information security [3] is prevailing everywhere and the process of providing security to the information system is getting more and more complex. In order to do that the DNA cryptography is one of the leading approaches to provide the secure mechanism to the information systems. The advantage of using DNA cryptography is its energy efficiency and storage capacity. One gram of DNA can store about 108 Tera bytes. [4]The DNA cryptography required huge computing time high-tech bimolecular laboratory and huge computational complexity. [5]The DNA cryptography used in this paper includes a conversion of the bio-molecules of DNA into a binary format and then reshuffles the sequence of the binary into a secret format. The DNA cryptography is the domain in which there is a great scope of research to be done. The major thing with the DNA cryptography is its storage capacity which is relatively high then the other biological existing things. Conventional cryptography involves mathematical problems that are difficult in nature. The theory and the realization are based on the more advance mode. The public-key and the secret-key in modern cryptography have specific flaws. [9]Now a days it is practically impossible for billions of processors processing in parallel to crack the immensely complex and big cryptographic keys being used. [10]In our cryptographic technique every character is assigned with four distinct arrangements of DNA bases which make Deoxyribonucleic Acid (DNA). The DNA based cryptography in this uses conversion of the DNA bio-molecule into RNA and RNA into protein strands and then conversion of protein triplet codons into the binary number and then further scrambles it using Doubly Even Order Magic Square. Adleman[11] started the DNA computing for trying to solve the Hamiltonian path problem through parallel computing.DNA can be consider as a storage medium of information in a biological medium extending the traditional electronic mediums of storing the information. To improve the security and encryption and to mitigate the current flaws we use the unique properties of DNA biomolecules.

II. RELATED WORK

Various research works have already been done in the field of DNA cryptography but there is still a constant need of improving the technique used in DNA cryptography most of the important thing is that the practical implementation of the DNA cryptography is till need highly sophisticated computing equipments which needed to be explored. [1]Behnam Bazli,Mustafa Anil Tuncel and David Llewellyn-Jones has done its work on Data Encryption Using Bio Molecular Information in this work it takes inspiration from DNA encryption scheme and use of biological alphabets to manipulate information by DNA sequence reaction to autonomously make a copy of its threads as extended encryption key. In another research work [2] Guangzhao Cui,Limin Qin,Yanfeng Wang,Xuncai Zhang have done another research work which is based on DNA synthesis and Digital Coding encryption. In the other research work [6] Xing wang, Qiang Zhang proposed a methodology which is based on DNA Computing. They included RSA based asymmetric key cryptography for encrypting and decrypting the message. In other research paper of [7]Souhila sadeg, Mohamed Gougache, Nabil Mansouri, Habiba Drias proposed an encryption algorithm based on symmetric key block cipher. The proposed algorithm models the process based on transcription and translation, with a high efficiency in the computation and security. The basic feature of Shannon's confusion and diffusion is essence of this algorithm. In another research paper [8]Bihash Roy, Gautam Rakshit, Pratim Singha, Atanu Majumdar, Debabrata Datta proposed a concept of strong DNA cryptography based on the symmetric cipher by using the concept of DNA. The

efficiency in computation, security while storing and transmitting the message has been improved.

III. PROPOSED MODEL

Various types of works has already been done by various researchers in the field of DNA cryptography but in this related work we have tried to implement the DNA cryptography using the different approach in which we used a special approach.
Process of encryption and decryption is as follows.

ALICE SIDE:

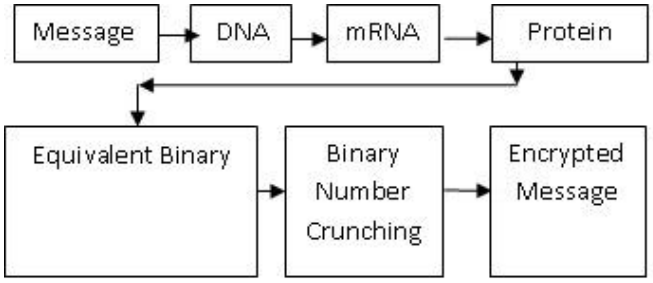

Figure.1

BOB SIDE:

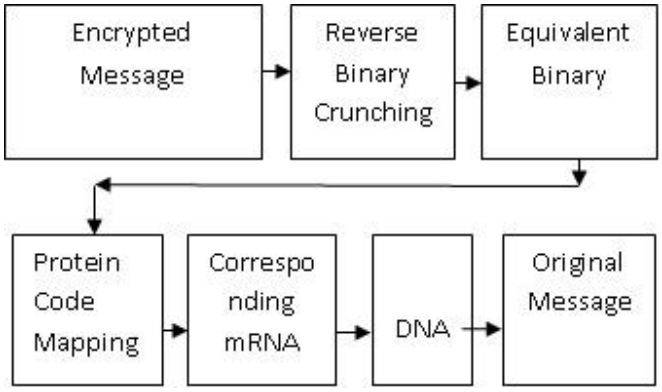

Figure.2

ALICE SIDE:
In the Alice side [figure.1] Alice takes a message say hello as a message and converted it into DNA sequences and further these DNA is converted into the mRNA sequences and then this mRNA is converted into the triplets of codons and the sequences of triplets of codons form the specific protein structure then this protein is converted into the equivalent binary then we shuffle this binary number using Doubly Even Order Magic Square and we get the encrypted message.

BOB SIDE:
The bob [figure.2] takes the encrypted message and then start decrypting it using the reverse process and the process is it takes the encrypted message and start doing reverse binary crunching and it get the equivalent binary and then he did the protein code mapping using the provided protein table and then it get the corresponding mRNA and that is converted into DNA and then the required message is generated at BOB side.

Protein Synthesis Table:

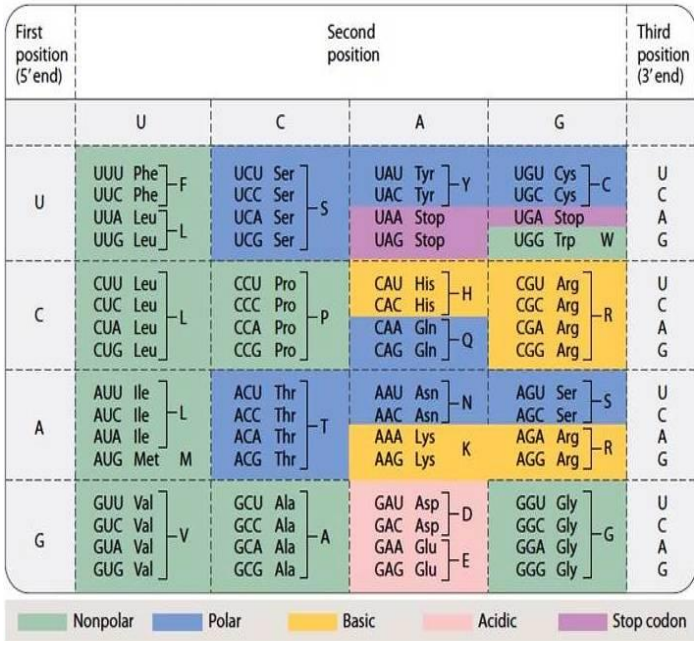

Figure.3

The above table [figure.3] is Protein Synthesis table for the conversion of mRNA into the protein sequence.

Binary table of Protein Synthesis Table:

	U(00)	C(01)	A(10)	G(11)	
U(00)	000000 000001 000010 000011	000100 000101 000110 000111	001000 001001 001010 001011	001100 001101 001110 001111	U (00) C (01) A (10) G (11)
C(01)	010000 010001 010010 010011	010100 010101 010110 010111	011000 011001 011010 011011	011100 011101 011110 011111	U (00) C (01) A (10) G (11)
A(10)	100000 100001 100010 100011	100100 100101 100110 100111	101000 101001 101010 101011	101100 101101 101110 101111	U (00) C (01) A (10) G (11)
G(11)	110000 110001 110010 110011	110100 110101 110110 110111	111000 111001 111010 111011	111100 111101 111110 111111	U (00) C (01) A (10) G (11)

Figure.4

The above table is the binary equivalent table of the genetic code table provided above.
In this table
U is considered as 00
C is considered as 01
A is considered as 10
G is considered as 11
And then find the possible permutation and combinations of the triplets of the codons to the corresponding equivalent binary.

Process of Encryption:-
The process of encryption is done using the doubly even order magic square method.
Let's take an example of digits.

Original-1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16
crunched no:-1,15,14,4,12,6,7,9,8,1,11,5,13,3,2,16

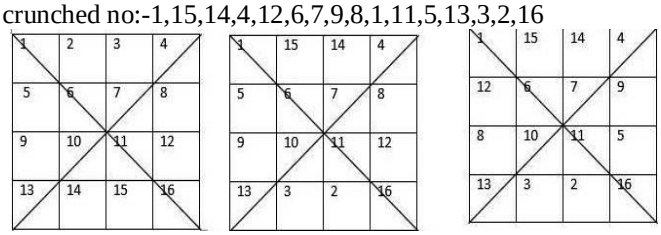

Figure.5

In the above example [figure.5] 2 and 15 position has been switched within itself and similarly the position of the 5 and 12 has been switched and then similarly the position of 9 and 8 and 14 and 3 has been replaced. And then read the numbers generated in the sequence from 1 to 16 from left to right.

Magic Square Number Crunching Using Binary Number [4]:
Original:-100011-00-1000111-01
Crunched:-1010110100011001

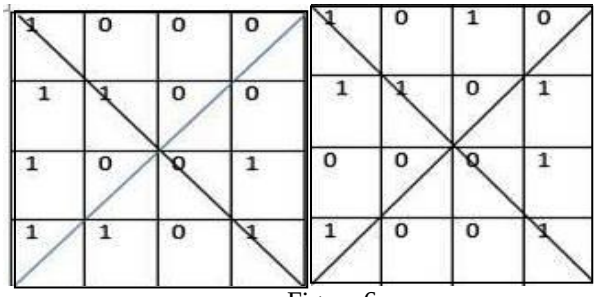

Figure.6

In this example [figure.6] we take the example of binary number which we crunched using the Doubly Even Order Magic Square we take the binary number 1000-1100-1001-1101 and after crunching we get the number 1010-1101-0001-1001.

Magic Square Number Crunching Using Binary Number:
Crunched:-1010110100011001
Rev. Crunched:-100011-00-100111-01

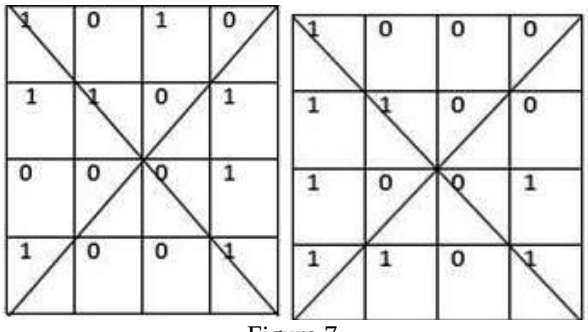

Figure.7

In this example [figure.7] we take the crunched binary number and then apply the doubly order magic square and then generate the original binary number. i.e;
The crunched number is 1010-1101-0001-1001

Original number after the reverse crunching is: - 1000-1100-1001-1101

Example:

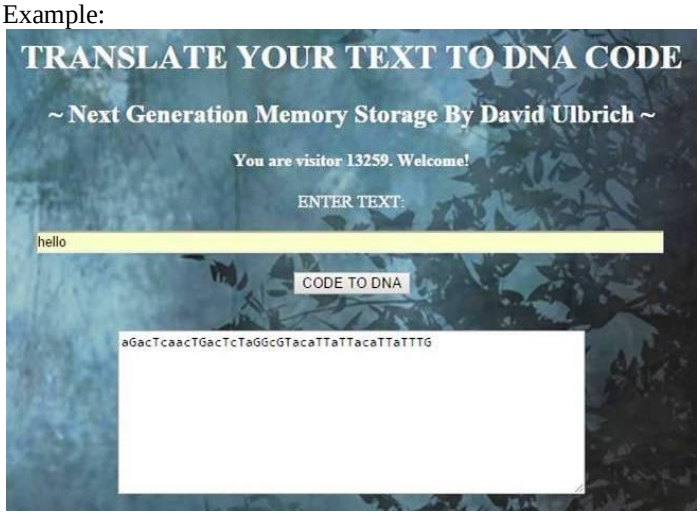

Figure.8

In this example we use the open algorithm for conversion of message into the DNA sequence [figure.8].

- HELLO is converted into DNA sequence.
- DNA CONVERSION→AGAC TCAA CTGA CTCT AGGC GTAC ATTA TTAC ATTA TTTG
- RNA CONVERSION →AGAC UCAA CUGA CUAU AGGC GUAC AUUA UUAC AUUA UUUG

mRNA CONVERSION→ UCU GAG UUG ACU GAG AUC CGC AUG UAA U AA UGU AAU AAA CCC

PROTIEN SYNTHESIS → Ser-Glu-Leu-Thr-Glu-lle-Arg-Met-Stop-Stop-Cys-Asn-Lys-Pro [figure.3]

Binary Equivalent→000100-00-111011-01-000011-10-100100-11-111011-00-100001-01-011101-10-100011-11-001010-00-001010-01-001010-10-001100-11-101000-00-101010-01-010101-10 [figure.4]

Crunching of Bits using Doubly Even Order Magic Square:-

0 0 0 1	0 0 1 1
0 0 0 0	0 0 0 1
1 1 1 0	0 1 1 0
1 1 0 1	1 0 0 1
000100-00-111011-01	0011000101101001

Reverse-Crunching of Bits using Doubly Even Order Magic Square

0 0 1 1	0 0 0 1
0 0 0 1	0 0 0 0
0 1 1 0	1 1 1 0
1 0 0 1	1 1 0 1
0011000101101001	000100-00-111011-01 (Original Binary)

Mapping Of Original Binary with the Encoded Protein Table:

000100-00-111011-01→Map with Protein Table→Ser-Glu
Protein Conversion to mRNA
Ser-Glu → UCU GAG
mRNA to DNA Conversion
UCU GAG→AGA CUC →AGA CTC→ Original Message

Concept of Bit-Guard:
In order for the convenience of the decryption side says BOB side we use the concept of bit-guard. This bit guard is like a safety control and help in to maintain the integrity of the message. It works like that we put sequence of 00,01,10,11 after every six bits so that if the sniffer sniff the packets and change it then the bob need not to check the entire bits of the message it just check the sequence of this bit guard and if he founds that the sequence has been altered then it comes to the conclusion that the data has been tampered and he discard the packets or tell him to resend it.

Example of Bit-Guard:
000100-00-111011-01-000011-10-100100-11-111011-00-100001-01-011101-10-100011-11-001010-00-001010-01-001010-10-001100-11-101000-00-101010-01-010101-10

Result Analysis:
Message Hello converted into DNA sequence AGA-CTC
And finally we get AGA-CTC into original message Hello.
NB: we are taking a part of the DNA sequence for the sake of the convenience in the computing of the result.

IV. CONCLUSION

As there are many models has been proposed around the globe on the field of DNA cryptography by following the various approaches to secure the confidentiality, Integrity and Availability of the information system in the same path we mention in our research a different approach to maintain the three traits of information security using the DNA cryptography following the different approach and methodology. This model increases the complexity but at that instant only so that the receiver can able to decrypt the message within the stipulated period of time and also the CIA of the message remain intact.

REFERENCES

[1] Data Encryption Using Bio molecule Information Behnam Bazli, Mustafa Anil Tuncel and David Llewellyn-Jones.
[2] Information Security Technology Based on DNA Computing. Guangzhao Cui, Limin Qin, Yanfeng Wang, Xuncai Zhang
[3] An Encryption Scheme Using DNA Technology Guangzhao Cui, Limin Qin, Yanfeng Wang, Xuncai Zhang
[4] On the Construction of Doubly Even Order Magic Squares Grasha Jacob1, Dr. A. Murugan
[5] Code for Encryption Hiding Data Intro Genomic DNA of Living OrganismsShuhong Jiao1 Robert Goutte2
[6] DNA computing-based cryptography Xing Wang, Qiang Zhang
[7] An encryption algorithm inspired from DNA Souhila Sadeg, Mohamed Gougache, NabilMansouri Habiba Drias.
[8] An improved Symmetric key cryptography with DNA Based strong cipherBibhash Roy, Gautam Rakshit,
[9] DNA Cryptography Sabari Pramanik1,*, Sanjit Kumar Setua
[10] A Novel DNA Sequence Dictionary method/or Securing Data in DNA using Spiral Approach and Framework 0/ DNA CryptographyShipraJain, Dr. Vishal Bhatnagar.
[11] An Architectural Framework for Encryption & Generation of Digital Signature Using DNA Cryptography. Deepak Singh Chouhan R.P. Mahajan
[12] http://dulbrich.is2.byuh.edu/dna/

