IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

DNA CRYPTOGRAPHY USING TRANSFORMATION ENCRYPTION ALGORITHM

Nikita S. Kolte*, Prof. Dr. K.V. Kulhalli
*ME (E & TC) student, D. Y. Patil college of Engg. & Tech. Kolhapur, Maharashtra, India.
H.O.D. Dept. of Information Tech., D. Y. Patil college of Engg. & Tech. Kolhapur, Maharashtra, India.

ABSTRACT

DNA cryptography is a new and promising field in information security. It combines classical solutions in cryptography with the strength of the genetic material. By introducing DNA into the common symmetric key cryptography, it is possible to benefit from the advantages of the classical cryptosystems and solve some of its limitations. In this paper we focus on Transformation Encryption algorithm based on confusion principle, number conversion, DNA digital coding which can effectively prevent attack. Data treatment is used to transform the plain text into cipher text which provides excellent security.

KEYWORDS: DNA, cryptography, confusion principle, transformation encryption algorithm, symmetric key.

INTRODUCTION

Cryptography provides range of features for information security. The main aspects treated by cryptography are: confidentiality, data integrity, authentication, and non repudiation. DNA cryptography consists in the use of genetics and bimolecular computation and it is one of the newest directions in cryptography. Genetic material such as DNA can be used as a vast storage space. This idea is inspired from the fact that DNA is a natural carrier of information which is encoded by a 4-letter alphabet: A, C, G, and T.

Considering the fact that DNA cryptography is a novel domain, one of the objectives of this paper is to find the way in which it can be applied in information security. Three main directions of using DNA in cryptography are found: storage space, computational power, generation of cryptographic keys from its long sequences. There can be two working environments with DNA: at molecular level, in a laboratory, with biological DNA and with digital DNA using available genetic databases.

DNA STRUCTURE

A cell is the fundamental functional unit in biological organisms. Most of the cells contain a nucleus and chromosomes inside of it. Genetic information (DNA) that controls cell functionality is divided into chromosome. Each chromosome is composed of a single DNA molecule which carries genes. The genome is the amount of all the genetic, hereditary information of an organism. It contains information from all the chromosomes. Complex organisms contain billions of cells. Each cell holds in its nucleus the same copy of chromosomes, but depending on the cell type it activates only a specific part of the whole genetic material (gene expression). The cell has the ability to store, retrieve and translate genetic instructions providing life to the organism.

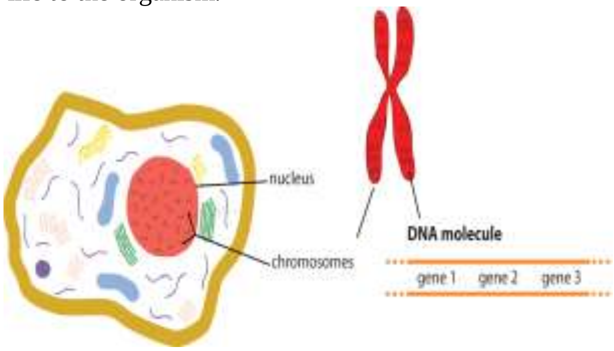


Fig. 1 Cell and its Genetic Material

Deoxyribose Nucleic Acid (DNA) has a helical shape, comprised of two long strands of nucleotides. A nucleotide has one of 4 bases: A – adenine, G – guanine, C – cytosine, or T – thymine, a deoxyribose sugar and a phosphate group. The sugars and phosphates make nucleotides to bind in a single DNA strand. The hydrogen bonds hold 2 strands together and create a double-stranded DNA. Hydrogen bonds last only between complementary pairs: A-T and C-G.

The three-dimensional structure of DNA was discovered in 1953 by Watson and Crick. DNA strands twist around each other forming a helix.

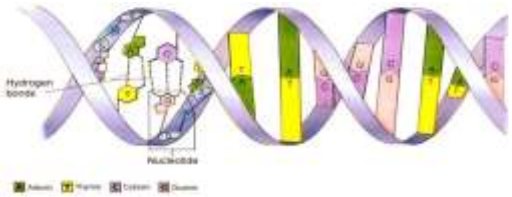


Fig.2 DNA Structure

Central dogma of molecular biology explains how genetic information is expressed in a protein. Genetic code is a correlation between the codons (3 nucleotide bases of DNA) and amino acids. Amino acids are structural units of proteins, the same way as nucleotides are in DNA. DNA genes "tell" the cell in what order to assemble the sequence of amino acids in protein.

DNA MOLECULAR STORAGE AND COMPUTATIONS

DNA has the capacity of storing all the large and complex data of any living being in the form of four bases which are A,C,T and G. Once it was revealed that the DNA has the ability of computing while performing the seven vertices problem of Hamilton path, the computers started dealing with the language of DNA containing the letters of the four bases A,C,T and G. and then the computational capability in DNA has been also taken in the field of cryptography. This means that a large amount of information can be stored in a compact, invisible for us. Different concealing techniques have been proposed to explore a hardly noticeable molecular medium.

DNA computing can be used to implement existing cryptographic algorithms at molecular level. Vernam cipher, based on binary XOR between plaintext and cipher text bit streams can be performed using DNA tiles by encoding them in binary ones and zeros.

DNA DIGITAL CODING SCHEME

In order to store data in DNA molecules it must be encoded in DNA bases. In a mapping method each letter from the English alphabet (A - Y), numbers (0 - 9), and some punctuation marks where encoded in 3 DNA letters, like: Q – AAC, or 9 – GCG. Considering that all of the digital information is encoded in binary, a straightforward method is to make a mapping table between DNA and binary alphabets. Using this mapping method any digital information can be transformed easily in DNA sequence.

Binary Value	DNA Base
00	A
01	T
10	C
11	G

Table 1. DNA digital coding Table

In DNA Digital coding the most fundamental coding
Method that is Binary Digital Coding, which is anything, can be encoded by two state 0 or 1 and a combination of o and 1. In DNA there are four kind of bases, which are Adenine (A) and Thymine (T) or Cytosine(C) and Guanine (G.) The simplest coding pattern to encode nucleotide bases (A,T,C,G) is by means four digits: 0(00),1(01),2(10),3(11),there are possibly 4!=24 possible pattern by encoding format like (0123/CTAG).

PROPOSED ALGORITHM

In our proposed algorithm, transformation encryption algorithm we are using DNA digital encoding scheme and confusion principle. It decreases the information redundancy and increases the efficiency. The digital coding of DNA sequence convenient for further mathematical operation and logical operation.

In the Proposed algorithm the text from user first converted into the ASCII form and then into the binary form. Then the DNA digital coding scheme is applied over the binary stream of data. In short we are transforming our plain text into the cipher text as DNA sequence form. That is why the algorithm named as Transformation Encryption Algorithm.

ENCRYPTION

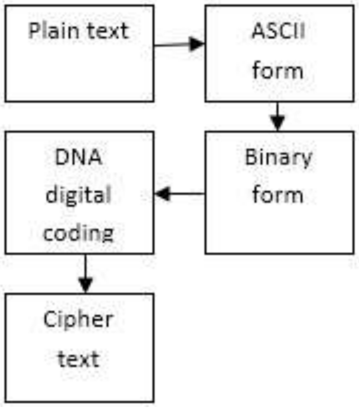


Fig 3. Encryption for transformation Algorithm

A plain text is converted into the ASCII and binary form. Then the binary steam of data grouped into two bits. DNA digital coding scheme is applied and at the output we get the cipher text in the form of ATCG.

DECRYPTION

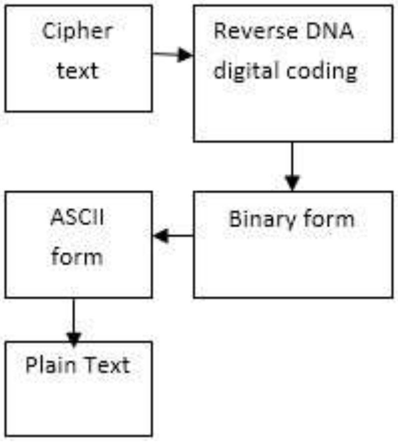


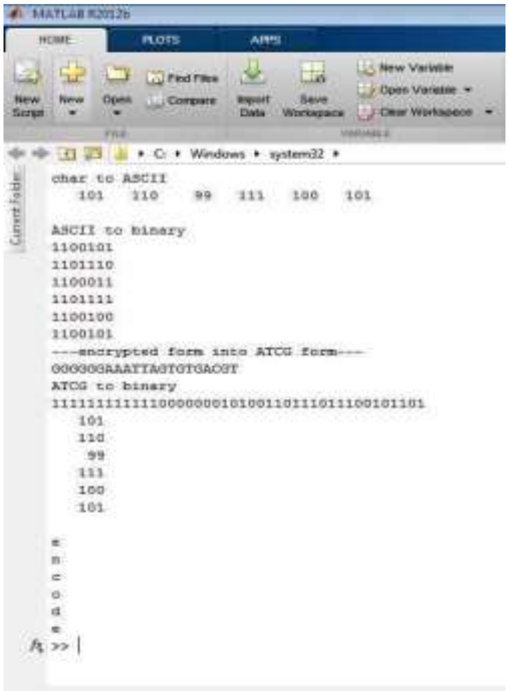
Fig 4. Decryption for Transformation Algorithm

At the receiver side, the encrypted text is received which is in the form of DNA sequence. The reverse DNA digital coding scheme is applied over it. We will get binary stream of data. Then it will be converted back into the Plain text.

In our proposed scheme the encryption and decryption key is nothing but the DNA digital coding table. Both the sender and receiver have to know the exact order of the table to retrieve the data.

In case the key is wrong, then there possibility of missing data or improper form of data. So, there is more secrecy of data is maintained.

MATLAB SIMULATION OUTPUT

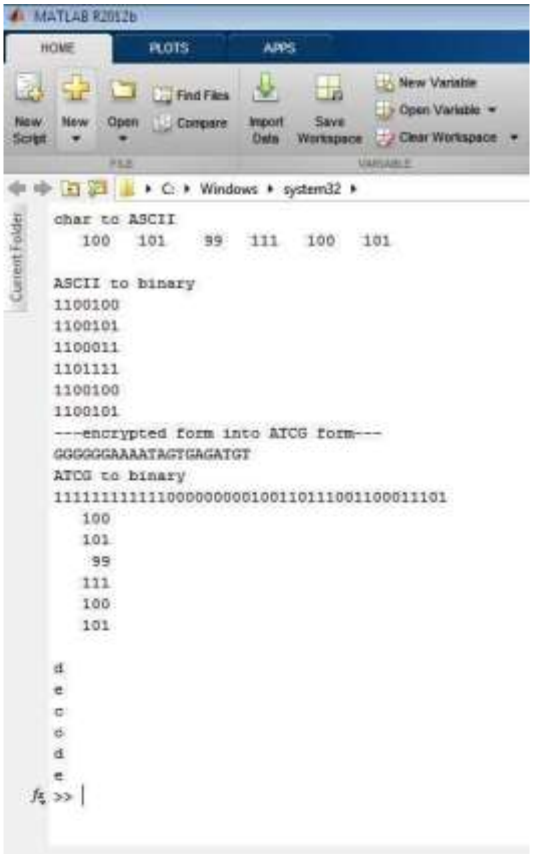


```
char to ASCII
   101   110    99   111   100   101

ASCII to binary
1100101
1101110
1100011
1101111
1100100
1100101
---encrypted form into ATCG form---
GGGGGGAAATTAGTGTGACGT
ATCG to binary
1111111111111000000010100110111011100101101
   101
   110
    99
   111
   100
   101

e
n
c
o
d
e
fx >> |
```

Fig   5. The encryption and decryption output for the Word ‘encode’



```
char to ASCII
   100   101    99   111   100   101

ASCII to binary
1100100
1100101
1100011
1101111
1100100
1100101
---encrypted form into ATCG form---
GGGGGGAAAATAGTGAGATGT
ATCG to binary
1111111111111000000000100110111001100011101
   100
   101
    99
   111
   100
   101

d
e
c
o
d
e
fx >> |
```

Fig 6. The encryption and decryption output for the Word 'decode'

The simulation of proposed algorithm is done with the help of MATLAB.

The table given below shows time measurement to transform the DNA sequence into binary for different plain text files sizes:

Plain Text(in KB)	DNA to binary transformation time(in ms)
43.9	10
105	23.4
199	38.6
475	91
768	152
1510	320
3110	637

Table 2. Time measurements of the DNA to binary transformation

FUTURE WORK

The complementary pair encryption algorithm and the DNA indexing encryption algorithm using DNA sequences will be implemented and will get evaluate on the basis of encryption and decryption time.

CONCLUSION

In this paper we implement the Transformation Encryption Algorithm for DNA cryptography. We use number conversion and DNA digital coding scheme for implementation of algorithm. The use of DNA digital coding scheme decreases redundancy of information coding and increases the efficiency compared to traditional encoding methods.
By using the technology of DNA digital coding, the traditional encryption method such as DES or RSA used to preprocess the plaintext. The digital coding of DNA sequence convenient for mathematical operation an logical operation.

REFERENCES

[1] L. M. Adleman, "**Molecular computation of solutions to combinational problems**," *Science*, vol. 266, pp. 1021–1024, 1994.

[2] A. Gehani, T. H. LaBean and J. H. Reif, "**DNA-based cryptography**," *DNA Based Computers V. Providence: American Mathematical society***,** vol. 54, pp. 233–249, 2000.

[3] Souhila Sadeg ,"**An Encryption algorithm inspired from DNA**", *IEEE*, pp 344 - 349 November 2010.

[4] Xing wang, Qiang Zhang ,"**DNA computing based cryptography**", *IEEE*, pp 37-42, 2009.

[5] Qinghai Gao, "**A Few DNA-based Security Techniques**", *IEEE*, 2011

[6] Deepak Singh Chouhan, R.P. Mahajan, "**An Architectural Framework for Encryption & Generation of Digital Signature Using DNA Cryptography**", *IEEE*, pp 743-748, 2014

[7] Ashish Kumar Kaundal, A.K Verma**, "DNA Based Cryptography: A Review",** *IJICT*, Volume 4, pp. 693-698, 2014

[8] B A Mitrans, A Kh Aboo, **"Proposed Steganography Approch Using DNA Properties",** vol.14, 2013

[9] William Stallings, *"***Cryptography and Network Security***"*, Third Edition, Prentice Hall International - 2003.