ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 7, Issue 6, June 2020

# A DNA Sequence Dictionary Method for Securing Data Using Spiral Approach

Dr.F.R.Shiny Malar[1], Ms. S.Geethu[2]

Associate Professor, Department of Computer Science and Engineering, Stella Mary's College of Engineering, Nagercoil, Tamil Nadu, India, shybertrijo@gmail.com [1]

Assistant Professor, Department of Computer Science and Engineering, Stella Mary's College of Engineering, Nagercoil, Tamil Nadu, India,geethus138@gmail.com[2]

**Abstract**: Need for information security and privacy is increasing in recent times. To ensure the secure transmission of data, cryptography is treated as the most effective solution. Numbers of cryptographic methods are now available. Traditional cryptography faces many problems such as difficulty in memorizing random keys and key storage in network. DNA cryptography method is one of the important branch of information security and it also the combination of both computer science and biological domain. This DNA computing technology introduces unbreakable algorithms and this methods converts text, image, audio, video in to an encrypted form that is nuleotides form. The DNA nucleotides are adenine, guanine, cytosine, and thymine. This type of encryption utilise the large storage capacity of DNA and also the features of DNA, the features means the properties that is high stability and durability of molecule, and less cost and easily available. In this paper, proposed a new method for providing security to the data in the form of DNA sequence. The proposed method provides security at two level using spiral transposition and DNA sequence dictionary table.

**Keywords**: DNA cryptography; Encryption; Decryption; DNA nucleotides

## I. Introduction

This DNA cryptography is a new field in the area of cryptography [3]. It is a combination of two domain i.e. computer science and biological science. DNA is the basic unit of all living organism. For understanding DNA cryptography, the researcher must have the knowledge of DNA. DNA is an abbreviation of Deoxyribo nucleic acid. DNA was discovered by Friedrich Miescher in 1869 [1]. DNA consist of four nucleotides namely Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). The double stranded complementary structure was discovered by Watson and Crick. In this complementary structure adenine pair with thymine and guanine pair with cytosine [2].

In DNA based cryptography the information are stored in the form of DNA nucleotides. This cryptographic method utilize the vast storage capacity of DNA, one gram of DNA is known to store about $10^8$ tera bytes of information. Like all other storage device DNA also requires security, the security for data in DNA cryptography is given in the same way as security given to other storage devices. The DNA cryptography provides security to the data in the form of DNA nucleotide form. The information passed through the network is encrypted by using DNA cryptography. Traditional cryptographic methods can not fulfilled the security requirements, but the DNA cryptographic method fulfills the storage and security requirements. In this paper, proposed a method that provide security at two level using spiral transposition and DNA sequence dictionary table and also give a brief literature survey about DNA cryptographic method.

## II. DNA Cryptography

DNA cryptography is a new branch of information security [3], and it is the combination of computer science and biological domain. In DNA cryptography any data that is text, image, video and audio information are converted into DNA nucleotides that is adenine, guanine, thymine, and cytosine. The DNA cryptography utilizes the features of DNA. DNA cryptography is the most effective cryptographic method compared to other traditional cryptography and stenographic method. Paper study about one of the important DNA crypto- graphic method, and analyze more about DNA cryptography.

### *A. Advantages*

Silicon chip are used for making of microprocessors and this silicon chip has reach their limits of speed so the chip makers uses the new material increasing the computing speed. DNA might one day be integrated into a computer chip to create a so called biochip that will push computers even faster the other computing devices [5].

Advantages are:

- Parallel Computing.
- High storage capacity and light weight.
- Less power consumption.
- Fast computations.

### *B. Applications*

The important applications are:

- String searching and comparison.
- Aggregate data release.
- Alignment of raw genomic data.
- Clinical use.
- Construction of one time password.

DNA based algorithms can be used in various fields such as job scheduling for clusters, GPU applications etc

### *C. Basic comparison between traditional and DNA cryptography*

The table.1 shows the basic comparison between traditional cryptography and DNA based cryptography.

**TABLE I**
BASIC COMPARISON BETWEEN TRADITIONAL AND DNA CRYPTOGRAPHY

| **Parameters** | **Traditional Cryptography** | **DNA Based Cryptography** |
|---|---|---|
| Ideal System | Silicon Chip Based | DNA Chip Based |
| Information Storage | Silicon Computer Chip | DNA Strands |
| Storage Capacity | 1 gram Silicon Chip Carries 16 Mega Bytes | 1 Gram DNA Carries $10^8$ Tera Bytes |
| Processing Time | Less | High |
| Performance Dependency | Implementation and System Configuration | Environmental Conditions |

## III. PROBLEM STATEMENTS

To ensure the secure transmission of data, cryptography is treated as the most effective solution. Numbers of cryptographic methods are now available. Traditional cryptography provides security through passwords, PINs, token etc. But it faces many problems such as difficulty in memorizing random keys and key storage. To overcome these difficulties DNA cryptographic methods are used and it has several advantages over traditional cryptographic methods. DNA cryptographic method utilizes the vast storage capacity of DNA and also the biological features of DNA. Here, a method is proposed to implement a DNA based cryptography that provides security at two level using spiral transposition and DNA sequence dictionary table.

## IV. RELATED WORKS

To study and analyze more about the DNA encryption techniques, the following literature survey has been done. In [9] developed a one of the DNA encryption method for encrypting data, and it consist of different steps, they are follows:

- The first step is to convert the data into binary form; this binary conversion is done by using taking the ASCII value of the data.
- The whole binary information is group into two bits. These two bit are converted into DNA nucleotides.
- In the next step group whole alphabets into group of three letters (cordons) and convert these cordons into amino acids.
- And also store ambiguity numbers in a file, this is important for decrypting original message.
- The important step is to encrypt the amino acid into secret form by using playfair encryption method.
- The output of playfair cipher method to convert into binary form by taking the ASCII value.
- The whole binary information are again group into two bits. These two bits are converted into DNA nucleotide. Finally store the resultant DNA sequence in a text file.

In [10] authors proposed a new encryption method based on DNA nucleotide. This method consists of four phases. In the first phase information is first converted into binary value, and then these binary forms are converted into matrix format. In the second phase the matrix information are converted into DNA nucleotide form that is adenine, guanine, thymine, and cytosine form. In the third phase the

DNA sequence addition operation is performed, added matrix is developed by using two Logistic maps. After the addition operation matrix is complemented. The result of complemented matrix is decoded; the decoded matrix is encrypted by using Data encryption standard (DES) algorithm and finally obtained the encrypted data.

In [11] authors developed another method for encrypting data by using DNA nucleotides. In this technique the security is obtained by using complementary rules and randomly generated keys. There are sixteen complementary rules. The keys and complementary rules are generated by sender and it is send to the receiver. The key is ten bit long and it is the combination of DNA nucleotides. For every information used the separate random key, the attacker cannot easily identify the keys so security is obtained. In this method the information is first converted into binary form, and then checks each binary bit. If the binary bit is one, the complementary rules are applied from left to right and if the bits are zero the complementary rules are applied from right to left. After this operation, obtained the encrypted text.

$\forall \in \{ \quad \}$

## V. Proposed Method

This proposed DNA based encryption method encrypts information such as image, audio, video and text file into DNA nucleotide forms and it provide security at two levels first level at binary and second at DNA conversion. The proposed method consists of 8 steps as shown in Fig.1.

$\oplus$ $\in \{ \quad \}$

### A. *Data*

The data used for encryption is any form that is text, audio, video, and image. These data can be converted DNA nucliotide form, adenine, guanine, thymine and cytosine.

### B. *Binary Conversion*

The data can be converted into binary form; the data can be audio, video, image, and text. The all binary data are group into block of 64 bits. These 64 bits of information are converted into the 8*8 matrix form [6].

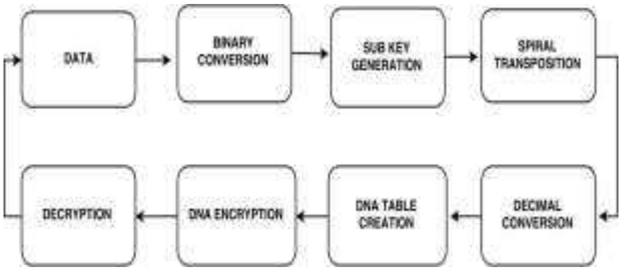

**Fig.1 Process Diagram of Securing Data Using Spiral Approach**

### C. *Sub Key Generation*

The sub key generation is an important part of DNA encryption. The sub key generated from an encryption password, which is super key. The super key is a 128 bit key. Two sub keys are generated. These two sub keys are also a 128 bit, from that key, derive 8 numbers. The number is the combination of 1 to 8 and it is used for spiral transposition. Sub Key Generation is accomplished by using an algorithm [7].

```
input := $Ink, r$
output :    =   $KEY$,   I 0, $r$-1
initialise := $Stk$,  $s_0$, $s_1.....s_7$, $S_0,S_1 .......S_7$
 $K := Ink \oplus Stk$
   for $i$ := 0 to r -1 do
      $x := K <<< 8$
       $K$  0 := $x$  $s(i \, mod \, 8)$
      $K_0 := b_0$ ——$b_1$—— .... $b_{15}$
       where each   $b_i$,  $i$   0,15 represents a byte
      for $j$:= 0 to 15 do
         $B0 := S(i \, mod \, 8)(bj)$

      end for
      $KEYi := B_0$——$B_1$——. ... —— $B_{15}$
       $K = KEYi$
end for
   return $KEY_i$. $\forall i \in \{0, r\text{-}1. \}$
```

TABLE II
SAMPLE BIT STRING

| Identifier | Sample bit string each of 128 bits |
|---|---|
| $s_0$ | B9C6 A9EA B7E2 5FD6 9E86 369A 1856 EC4A |
| $s_1$ | AE98 5DFF 2661 9FC5 8623 DC8A AF46 D590 |
| $s_2$ | CB5E 129F AD4F 7E66 780C AA2E C8C9 CEDB |
| $s_3$ | 2102 F996 BAF0 8F39 EFB5 5A6E 3900 02C6 |
| $s_4$ | 3DD4 254E EBDC FE2B FF0C D7F7 F7BE 2DC2 |
| $s_5$ | 6BFA BF9F FFFE EFFD FECD 9DEF ED3D EDFA |
| $s_6$ | BDFF 77FF FCFF BDF7 EFDF DFAF FF53 D9FF |
| $s_7$ | FDFE BBFC FAFF FFDF 47D6 D7FF 7FBF E76E |

The sub keys are generated from super key (*Ink*). This algorithm used super key (*Ink*) and number of rounds *r* is used as input parameters. The super key is a 128 bit number, and number of rounds is two, because two sub keys are generated. The generated sub keys are represented by *KEYi*. For generating sub keys these algorithm uses some parameters they are denoted by *Stk*, $s_0$, $s_1$.....$s_7$ and $S_0$,$S_1$ $S_7$. The *Stk* is a 128 bit string; this is generated by using pseudorandom number generation. The $s_0$, $s_1$.....$s_7$ are sample bit strings each bit string is 128 bit and it is shown in Table 2.4. These bit string are based on prime numbers. The other variables used is eight *S* boxes they are represented as $S_0$,$S_1$ $S_7$ and these *S* boxes are 16x16 arrays.

The algorithm operation is as follows [6]. First the super key is bitwise xored with $S_{tk}$, the output of this operation is *K*, which is also a 128 bit string. The algorithm performs circular left shift operation of *K by* eight bits. The output of this generated is represented as a variable denoted by *x*. Next step is the bit wise operation of *x* and the corresponding bit string. The result of this is stored in variable $k_0$. The $k_0$ is then split into sixteen bytes and it is represented as $b_0$ , $b_1$,. . . , $b_{15}$. These are used as an input for producing output bytes B0, $B_1$ $B_{15}$. This output bytes are combined together to form sub keys. This same operation is performed twice and to generate two sub key.

### D. *Spiral Transposition*

The spiral transposition is done by using the generated sub keys. The sub keys are generated from the super key. The matrix information is used for spiral transposition; the matrix information is the binary value of the original message. Spiral transposition changes the position of bits in the matrix. Two types of transposition done are row wise transposition and column transposition [6]. Spiral transposition changes the position of bits so it becomes difficult for the attacker to guess the actual data.

### E.*Decimal conversion*

The decimal conversion [6] is another step. This step reduces the information size. In this step the result of spiral transposed information are converted into decimal form. This decimal information is used for DNA encryption. The decimal information is replaced by DNA nucleotide format from the DNA table. The decimal value can be generated by taking the 8 bits of row or by taking the 8 bits of column. The eight bit binary data gives the decimal value ranging from 0 to 256.

### F. *DNA table creation*

The DNA table creation [6] is an important portion of this proposed method. The DNA table consists of the combination of DNA nucleotides. The DNA table is created by using super key, each time super key is different, the DNA table elements are different, so the attacker cannot easily identified the DNA table elements. The DNA table consists of 256 combinations of DNA sequence information of length four.

### G. *DNA Encryption*

This is the important step of this method. In this step the decimal information are converted into DNA nucleotide format, which is adenine, guanine, thymine, and cytosine. This is the last phase of the proposed encryption method. In this phase decimal data are converting into the DNA sequence [6]. The DNA sequence are obtained from the DNA table, the DNA table are already created using the sub key. The DNA table consists of 256 decimal numbers and their corresponding DNA sequence of length 4. These 256 DNA sequence is different depends on super key, each time super key is different the DNA table information are different. This step hides the encrypted binary information into the DNA sequence.

### H. *Network creation*

Network creation is one of the important parts of the project; Number of nodes is created so a network is created. Each time node is created the public key of the node is broadcasted and all other nodes in the network know the presence of new node and also know the public key of the the new node created. If a node communicates with other node in the network first encrypt the super key by using the public key of the receiver with the help of RSA algorithm. Finally send the encrypted message and encrypted super key to the receiver.

### I. *Super key encryption*

Super key encryption is the important part of the proposed method. The super key is encrypted by the sender using the public key of the actual recipient with the help of RSA algorithm. At the receiver side the receiver decrypt the super by using the private key of the receiver. The attacking possibility is reduced here because the super key encryption.

### J. *Decryption*

Decryption phase is the important portion of this proposed method, in this phase the original message is recovered from encrypted message. In this phase the receiver receives the encrypted message and encrypted super key. The receiver first decrypts the super key by using the private key of the receiver with the help of RSA. The super key is retrieved back and from the super key creates a sub keys using an algorithm already used. This sub keys are used for reverse operation of transposition. The DNA table is also created by the receiver from the sub keys. In the decryption phase operation the encrypted form of the messages that is DNA nulceotide form are converted back into decimal form using the DNA table already created. The decimal information is then converted into binary form and this binary information are transformed into a 16*16 matrix format. The reverse operation of spiral transposition are done on this binary information, first perform the column reverse operation is perform the row operation and finally obtained the binary information are converted into original message. The attacker cannot decrypt the original message because the super key is decrypted only by the original receiver by using their private key. All other decryption operation is performed with the help of super key without super key decryption operation cannot perform.

## VI. Experimentation

Tools used to implement the proposed method are JAVA. The messages are encoded in the form of DNA nucleotides and exchange between sender and receiver using TCP, which is implemented using two laptops connected to same network. In this proposed method, the sub keys are generated by using an algorithm and these sub keys are used for spiral transposition and DNA table creation. The sub keys are generated from super key and each time super key is different the sub keys are different and DNA table are different, this is one of the important part of the proposed method. The super key is encrypted by using RSA algorithm with the help of public key of the receiver and sends the encrypted super key to the receiver. The receiver decrypts the super key and decrypts the encrypted message

## VII. Conclusion And Future Scope

The proposed DNA cryptographic method encrypts data in the form of DNA nucleotide, adenine, guanine, thymine, and cytosine format. The DNA encryption does not use the complex biological functions of DNA for encrypting data. The one of the important disadvantages of traditional cryptographic method are storage, DNA cryptographic method can overcome these difficulties and store vast amount of data. In this method keys are dynamically generated from super key likewise DNA table is also created dynamically, this is one of the important advantages, so the attacker cannot easily identify the keys and DNA nucleotides. The difference between the traditional and DNA cryptography clears the importance of the DNA cryptography. In future an algorithm can be designed for DNA based cryptography, and also in future, will study the attacks possible on the DNA cryptographic algorithms and on the proposed method also.

## References

[1]. M.E. Jones, Albrecht Kossel, ”A Biographical Sketch”, Yale Journal of Biology and Medicine, Vol. 26, Issue No. 1, pp. 80-97, 1953.

[2]. J.D.Watson, F.H.Crick,“Molecular Structure of Nucleic Acids”, Nature, Vol. 171, Issue No. 4356, pp. 737-738, 1953.

[3]. Anam B., Sakib K., Hossain M.A., Dahal K.: Review on the Advance-ments of DNA Cryptography. ar Xiv:1010.0186v1 [cs.CR] 1 Oct 2010.

[4]. Tushar Mandge, Vijay Choudhary A DNA Encryption Technique Based on Matrix Manipulation and Secure key Generation SchemeIEEE.

[5]. Swarnendu Mukherjee, Debashis Ganguly, Swarnendu Bhattacharya, Partha MukherjeeA Cognitive Study on DNA Based Computation In-ternational Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009.

[6]. ShipraJain, Dr. Vis hal Batnagar, ”A Novel DNA Sequence Dictionary

method for Securing Data in DNA using Spiral Approach and Framework of DNA Cryptography, IEEE conference 2014.

[7]. Oppel-1:A New Block Cipher Arshad Ali CESAT, Islamabad, Pakistan Ar- shad.Ali.2008@live.rhul.ac.uk.

[8]. Shyamasree C M, Sheena Anees, ”Highly Secure DNA-based Audio Steganog- raphy”, 2013 International Conference on Recent Trends in Information Technology (ICRTIT) 2014.

[9]. Ranu Soni, Arun Johar, and Vishakha Soni, ”An Encryption and Decryp- tion Algorithm for Image Based on DNA”, 2013 International Conference on Communi- cation Systems and Network Technologies.

[10]. Shipra Jain, Vishal Bhatnagar, “Bit Based Symmetric Encryption Method Using DNA Sequence,” 2013 International Conference on Communication Systems and Network Technologies.

[11]. P.Surendra Varma, K.Govinda Raju, (2014), “Cryptography based on DNA using random key generation scheme”, International Journal of Science Engineering and Advance Technology, IJSEAT, Vol 2, Issue 7, ISSN 2321-6905, pp.168-175.

[12]. Mohammad Reza Abbasy, Pourya Nikfard, Ali Ordi, Mohammad Reza Najaf Torkaman, (2012), “DNA base data hiding algorithm”, in international Journal on New Computer Architectures and Their Applications., ISSN: 2220-9085, pp.183-192.

[13]. Kumar, S., Wollinger, T. (2006), “Fundamentals of Symmetric Cryptography”. Embedded Security in Cars, 125143.

## BIOGRAPHY

**Dr. F.R. Shiny Malar was** born in Nagercoil, Tamil Nadu State, India in 1986. She studied Information Technology in St.Xavier’s Catholic college of Engineering, Chunkankadai, Kanyakumari District, Tamil Nadu State, India from 2003 to 2007. She received Bachelor’s degree from Anna University, Chennai 2007. She received the Master degree from Manonmaniam Sundaranar University Tirunelveli. And also received Doctorate degree in the Department of Computer Science and Engineering, in Noorul Islam Centre for Higher Education, Noorul Islam University, Kumaracoil, Tamilnadu, India; Currently she is working as a Associate Professor in Stella Mary’s College of Engineering, Nagercoil, Tamil Nadu State, India. She has published more than 8 international journals and presented more than 10 papers in international and national conferences and their research interest include image security, networking and image processing.

Ms. Geethu S was born in Thiruvananthapuram, Kerala State, India in 1993. Studied Computer science and engineering in IHRD College of engineering karunagapally, velluthamanal, Kollam District, Kerala, India from 2011 to 2015. Received the Bachelor's degree from Cochin University, Cochin 2015. Received the master degree from Government engineering College, Wayanad, Kerala, India, Abdul Kalam Technical University, Kerala. Currently she is working as assistant professor I Stella Mary's College of engineering, Aruthenganvilai, Nagercoil, Tamil Nadu State, India. She has published one journal paper in cloud computing.