Multilevel Security Using DNA Cryptography

Sini Thankachan[1], Jeena P Abraham[2], Smita C Thomas[3]
[1, 2, 3] Dept of Computer Science and Engineering
[1, 2, 3] Mount Zion College of Engineering, Kadammanitta, Pathanamthitta, Kerala

Abstract- *DNA cryptography is an emerging field in cryptography which makes use of biological DNA sequences for encryption of data. DNA sequences are unique in nature and this makes the algorithm highly resistant to attacks. The proposed method has multiple levels of security for encryption to enhance security. Higher level of security is provided by using random DNA conversion tables and transposition techniques. Partitioning technique is used to provide parallelism. The proposed DNA cryptographic technique provides faster and better security compared to other cryptographic techniques.*

Keywords- DNA cryptography, Partitioning, Transposition, Random DNA conversion table

I. INTRODUCTION

Security is one of the fundamental and significant concerns in data transmission. Cryptography is a process through which we protect our information from the unauthorised person. Multiple cryptographic techniques have been used for securing the data over network. Cryptography plays a vital role in data integrity in the three components of the CIA triad (Confidentiality, Integrity, and Availability). DNA cryptography is an emerging field of cryptography and it uses human DNA sequences for encryption of data. DNA cryptography plays a vital role in the data transmission and communication areas. DNA computing was developed by Leonard Adleman in 1994. DNA stores a large amount of information inside the tiny nuclei of living cells. It includes all the instructions needed to make every living creature on earth. Primary advantages of DNA computation are miniaturization and parallelism of conventional silicon-based machines. For example, a square centimeter of silicon can support around a million transistors, whereas current manipulation techniques can handle it in 1020 strands of DNA. DNA, with its unique data structure and parallel operations, allows one to solve a computational problem from a different point of view. In DNA, information is stored as a code made up of four nitrogenous bases: Adenine (A), Guanine (G), Cytosine(C), and Thymine (T). These chemical bases are combined to other by hydrogen bond. Adenine always combined with Thymine and Cytosine combined with Guanine, and this is called complementary rule. DNA cryptography provides solution to the weakness or limitations of existing systems by developing algorithms which can only be decrypted if the key and the correct DNA sequence can be generated, and this is possibly difficult as DNA sequences are unique in nature. DNA is very powerful from cryptographic point of view. The binding capabilities of nucleotide bases (A-T, C-G) enable creation of self-assembly structures that are an excellent means of executing computations. Huge storage capacity of DNA is another advantage. DNA has a random nature and hence the cryptography which is based upon its principle is unbreakable. But on the other hand practically using the implementations requires a lot of time and resources. Simple and effective algorithms are necessary.

II. LITERATURE REVIEW

DNA cryptography is a one of the new technology used for providing higher security of data.

Ning,Kang., proposed the pseudo DNA cryptography method which is not limited to encryption and decryption, but also for message authentication. The limitation here was the length of the resultant MAC which was tough to manage and the solutions to this problem was to do several rounds of the procedure, along with padding to get the MAC with fixed length. [1]

Zhihua Chen, proposed a recursive sticker molecular algorithm to the DES. The molecular sticker algorithm consists of three parts: key space initialization with all potential keys, encryption and detecting the equivalent key. The essential operations required in DES are implemented by the molecular sticker functions. This work indicated that the DES are perhaps insecure[2].

Borda et al., in their study offered the principles of biomolecular computation (BMC) and several algorithms for DNA steganography and cryptography like; One Time-Pad (OTP), DNA XOR OTP and DNA chromosomes indexing[3]. The technique proposed in [4] uses efficient DNA operations, but it depends on static hash mapping which can be vulnerable to attacks. The technique proposed in [6] has three levels of security, but it makes use of fixed DNA lookup tables.

III. EXISTING SYSTEM

The existing paper implemented 3-Levels of security.

i. First level uses a key of any desired length and maps binary value of input text.
ii. Second level uses LBP operations on output of first level.
iii. In the third level, DNA sequence values are obtained using lookup table values.

3.1. LIMITATIONS

- Existing system make use of fixed lookup table values and this may weaken the security of the technique.
- Each DNA sequence of length four in cipher text can be mapped to a particular character in plain text. More than one plaintext character is not affecting a single sequence. If an attacker can get hints on some statistical characteristics of cipher text or algorithm or key, possibilities of extracting plain text from the same cannot be ignored.
- Each character in plain text is processed individually which consumes more time.

IV. PROPOSED SYSTEM

The proposed system in this paper overcomes the limitations of existing system. Proposed technique has three levels of DNA security. In the first level shift the text by using the shift key and convert into to the binary equivalent. In the second level, binary values are converted to DNA sequences using DNA sequence table. In the third level, partitioning is done to provide parallel processing and higher security.

The proposed technique does not use fixed DNA base sequence values corresponding to each binary value; rather it makes use of dynamic or random values. At each iteration, the binary values corresponding to DNA bases (A, T, C, and G) will be changed. Value of A in current iteration will be equal to the value of its complementing pair (T) in the previous iteration and vice versa (T gets previous value of A). Similarly, C gets previous value if G and G uses previous value of C. For example, initially "A" corresponds to binary value 01. In the next iteration, "T" corresponds to 01. Below shown table is the values corresponding DNA bases at a particular iteration. During transmission, sender notifies the iteration number to receiver so that receiver can create the conversion table to be used for decryption. The advantage of using Random DNA conversion table is the higher level of security it promise, since it is easy to breach the conversions using fixed DNA conversion tables.

Table 1: Random DNA Conversion Table

DNA base	Binary value
A	00
G	01
C	10
T	11

4.1 ENCRYPTION

- first level of security

For example, take the string "UNIVERSITY OF KERALA". The encryption steps are as shown below:

Step 1: Shift the characters of plain text by
a key of fixed length.
Let the length of the shift key be 3. Each character of the plain text is shifted forward by 3. Shifted string is:

XQLYHUVLWBRINHUDOD

Step 2: Find out the binary value corresponding to each character in plain text.

X: 01011000
Q: 01010001
L: 01001100
Y: 01011001
H: 01001000
U: 01010101
V: 01010110
L: 01001100
W: 01010111
B: 01000010
R: 01010010
I: 01001001
N: 01001110
H: 01001000
U: 01010101
D: 01000100
O: 01001111
D: 01000100
"Space": 00100000

Converted binary string is
01011000010100010100110001011001010010000101010101
01011001001100010101110100001000100000010100100100

1001001000000100111001001000010101010100010001001111101000100

- Second level of security

Step 3: Map the resultant binary values to DNA sequence using Random DNA Conversion table.

Mapped DNA sequence string is:
GGCAGGAGGATAGGCGGACAGGGGGGGCGATAGGG
TGAACACAAGGACGACGACAAGATCGACAGGGGGA
GAGATTGAGA

- Third level of security

Step 3: Partitioning - The unique DNA sequences are split into smaller n*m sequences.

Taking n=10 and m=8, Partitioned strings are:
GGCAGGAGGA
TAGGCGGACA
GGGGGGGCGA
TAGGGTGAAC
ACAAGGACGA
CGACAAGATC
GACAGGGGGA
GAGATTGAGA

Step 4: Transposition technique: Partitioned strings are read vertically(column wise) to obtain the cipher text.

GTGTACGGGAGACGAACGGGAACGAGGGACAAGCG
GGAGTGGGTGAGTAGGGAGGGGACACAGGGCGAGT
GGAAACACAA

Security of this step can be increased by performing multiple stages of transposition. The output will be a complex permutation which cannot be reconstructed easily. Here in this example, single stage transposition is shown.

Flow chart of the above discussed steps is given below:

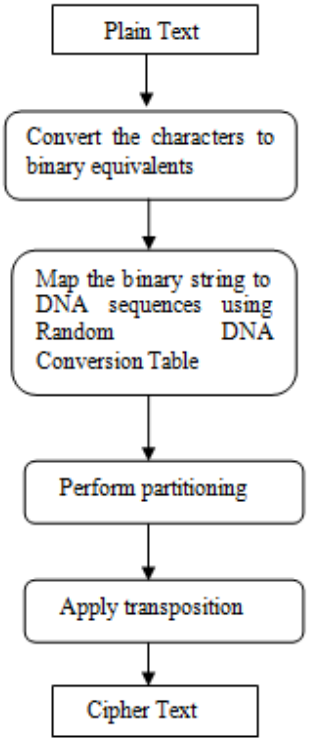

Fig. 1 Proposed Encryption Process

4.2. DECRYPTION

Received cipher text is

GTGTACGGGAGACGAACGGGAACGAGGGACAAGCG
GGAGTGGGTGAGTAGGGAGGGGACACAGGGCGAGT
GGAAACACAA

Steps for decryption are given below:

Step 1: Partitioning - The unique DNA sequences are split into smaller m*n sequences.
Using n=10 and m=8,

Partitioned string is,
GTGTACGG
GAGACGAA
CGGGAACG
AGGGACAA
GCGGGAGT
GGGTGAGT
AGGGAGGG
GACACAGG
GCGAGTGG
AAACACAA

Step 2: Transposition: Partitioned strings are read vertically (column wise).

Resultant string is:
GGCAGGAGGATAGGCGGACAGGGGGGGCGATAGGG
TGAACACAAGGACGACGACAAGATCGACAGGGGGA
GAGATTGAGA

Step 3: Map the DNA sequence to corresponding binary values using Random DNA Conversion table for the particular iteration.

Mapped string is
0101100001010001010011000101100101001000010101010101
0101100100110001010111010000100010000001010010010010
0100100100000010011100100100001010101010001000100111
1101000100

Step 4: Map each character corresponding to binary value in the obtained text.

String is :
XQLYHUVLWBRINHUDOD

Step5 : Shift back the characters of the resultant text by a key of fixed length 3 to obtain the plain text.

Plain text is:
UNIVERSITY OF KERALA

V. CONCLUSIONS

DNA cryptography is hiding data in DNA bases. In this proposed technique, a random DNA conversion table is used with which it is difficult for attacker to guess the DNA sequence substitution. Also, multiple characters of plain text impact on a particular sequence in cipher text which makes the technique more resistant to attacks. Processing takes place on the entire sequence, not on separate sequences corresponding to individual characters which increase the speed. Level of security can be increased further by using multiple stages of transposition. It is anticipated that the proposed solution will provide promising results.

VI. ACKNOWLEDGEMENTS

We would like to thank, Almighty God at first, without his grace and blessings this work would not have been possible. We would also like to thank the faculty members of Mount Zion College of Engineering, for their great support towards this work.

REFERENCES

[1] Ning, K., 2009. A pseudo DNA cryptography method ar Xivpreprint arXiv:0903.2693
[2] Chen, Z., Geng, X. and Xu, J., Efficient DNA sticker algorithms for DES. In Bio-Inspired Computing:Theories and Applications, 2008. BICTA 2008. 3rd IEEE International Conference pp. 15-22, Sept. 2008
[3] Borda, M. and Tornea, O., DNA secret writing Techniques. In IEEE conferences , pp. 451-456,June 2010
[4] Dr. B. Lavanya, Ms. Aafreen Nawresh. A, Encapsulating Security Mechanisms On Data, IEEE International Conference on Computational Intelligence and Computing Research,2016
[5] D. Prabhu and M. Adimoolam, "Bi-serial DNA encryption algorithm (BDEA)", [online]. Available <arXiv: II 01.2577vl> [Jul 12, 2012]
[6] N. Srilatha and G.Murali, Fast Three Level DNA Cryptographic Technique to Provide Better Security, 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)