A Dynamic DNA for Key-based Cryptography

Bahubali Akiwate
Department of Computer Science & Engineering
K.L.E College of Engineering and Technology
Chikodi,India
bahubalimakiwate@gmail.com

Dr. Latha Parthiban
Department of Computer Science & Engineering
Pondicherry University
Pondicherry,India
lathaparthiban@yahoo.com

***Abstract*— A dynamic DNA for key-based Cryptography that encrypt and decrypt plain text characters, text file, image file and audio file using DNA sequences. Cryptography is always taken as the secure way while transforming the confidential information over the network such as LAN, Internet. But over the time, the traditional cryptographic approaches are been replaced with more effective cryptographic systems such as Quantum Cryptography, Biometric Cryptography, Geographical Cryptography and DNA Cryptography. This approach accepts the DNA sequences as the input to generate the key that going to provide two stages of data security.**

Keywords— ***DNA Cryptography, Dynamic DNA, DNA Digital Coding, Encryption, Decryption***

I. INTRODUCTION

Cryptography is the field in which information can be encrypted and transmitted over the network then it can be decrypted back to its original form. Here the encryption can be done at the sender before transferring it into network and after receiving encrypted message is decrypted at the receiver using a key provided to it. Before the communication between sender and the receiver both of these entities must agree upon some protocols by how they are exchanging their information, it is some kind of handshake procedure. While information is in transit in network, even the intruder or attacker gets the information is not in understable manner. There is a need to adopt more secure and reliable encryption algorithm. The DNA Cryptography is one such secure and reliable data security approach. Here we proposed a technique that is able to accept various data forms as input such as text, image, audio and video with Unicode approach to reach more users worldwide. However existing approaches in the field of cryptography focuses only based on ASCII character set, ignoring non English users [1].

As we know the use of internet is increasing day by day not only through computers but also through smart phones, a lot of new technologies and optimized use of resources are expected by an individual or by any organization. Data load will be more on networks and at the same time there are higher chances of capturing, stealing, modifying or cracking of data can be done by an attacker or intruder, which leads to a need of a new technology which can fulfill high storage, randomness nature in generating keys used in encryption as well as decryption by providing more secure and reliable communication [6].

II. DNA CRYPTOGRAPHY

DNA Cryptography is a field in which lots of researches are happened and happening and it is still expected to come up with better solution, meeting modern era problems and issues. The technologies under DNA Cryptography which are already accepted are PCR (Polymerize Chain Reaction), DNA synthesis and DNA Digital Coding [13]. Here we used DNA Digital Coding technique in which encoding and decoding can be done with the use of binary values such as 0 and 1.DNA Digital Coding is based on biological structure such as DNA (Deoxyribo Nucleic Acid) which is composed of four basic nucleotides such as Adenine (A), Cytosine (C), Guanine (G) and Thymine (T). The proposed system combines the traditional, currently available cryptosystems, uses DNA Digital Coding and maps digital data into biological DNA sequences and vice versa.

The proposed system may be applicable to the areas of digital transactions such as credit card/debit card payments, e-mail, SMS (Short Message Service) encryption where users wants to have more secure communication.

III. SYSTEM DESIGN

The following figure 1 shows the system design in which sender is sending the information to the receiver, the sender encodes the data by using the DNA digital coding and key combination to produce the cipher text and generate the random key used for decryption process by the receiver.

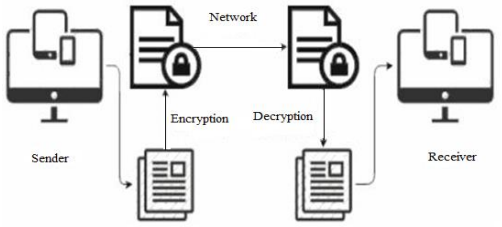

Fig. 1. System Design

The receiver is receiving the key from sender; the receiver uses this key to get the original message. The cipher text is unreadable after the encryption; we can read the text by using the key to convert cipher text to original message. There are two types of cryptography one is symmetric cryptography in which sender and receiver share a single secret key between them (Such as in AES algorithm). Another one is asymmetric cryptography in which pair of keys are used; one is public key at the sender and the private key at the receiver. Public key

may be known to anyone (Such as in RSA algorithm).The results proved that symmetric cryptography is faster than asymmetric one.

A. Activity Diagram

The following figure 2 shows the activity diagram of a dynamic DNA for key-based Cryptography.

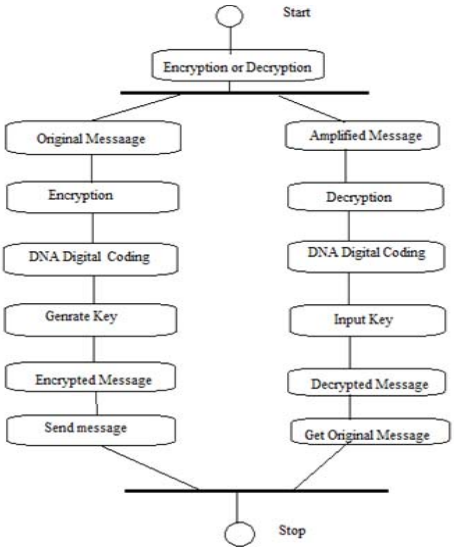

Fig. 2. Activity Diagram

The above figure 2 shows the workflow of the software on the original message selected to encode i.e start of encryption process, DNA digital coding is used to encode the message, after the encryption the random key is generated, and cipher text is made ready to send. In decryption side the amplified message is received, started to decrypt by using DNA digital coding and finally entering the key to get the original message.

B. Sequence Diagram

The figure 3 shows a sequence diagram of dynamic DNA for key-based Cryptography approach.It shows an interaction between components such as Sender,Encryption Module, Decryption Module, Receiver. Initially a sender sends an original message and converts it into ciphertext by using DNA Cryptographic algorithm (Encryption Module) which is also able to generate a random key communicated with Receiver.When the receiver receives a ciphertext it applies the random key communicated by sender over ciphertext to collect back the original message by using DNA Cryptographic algorithm (Decryption Module).The arrow labelled with 'communication' indicates the handshaking where parties will be agreed upon some protocols such as in case of Advanced Encryption standard (AES),Diffie-Hellman(DH).The decryption process is reversed process of encryption using DNA Digital Coding.

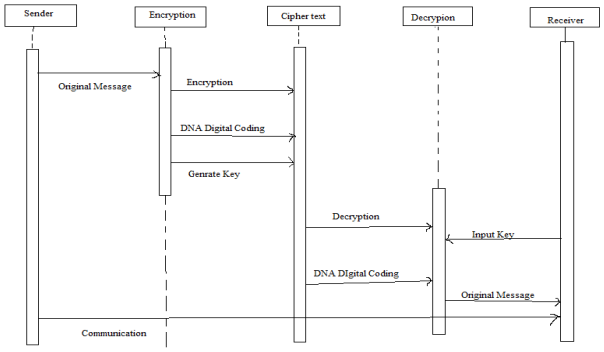

Fig. 3. Sequence Diagram

IV. IMPLEMENTATION

Here we implemented the code for DNA algorithm in NetBeans IDE environment, which encrypts and decrypts the characters, text file, image file and audio file by using Java language through following steps:

1. Select the text/file which contain data
2. Convert data into ASCII equivalent and then into Unicode characters
3. Convert Unicode to hexadecimal
4. Convert hexadecimal to binary
5. Convert binary to DNA digital code
6. Check the file content which contain DNA amplified data (Generated Message)
7. Get DNA amplified data to original text (Decryption)

A. Encryption Module

The following figure 4 shows the actual encryption process in which the original message is converted to the ASCII later into Unicode characters then converted to hexadecimal. After the hexadecimal it is converted into binary. After the binary conversion, the message is divided into four parts of message then using DNA digital coding and key combination the message was generated and transferred over the network.

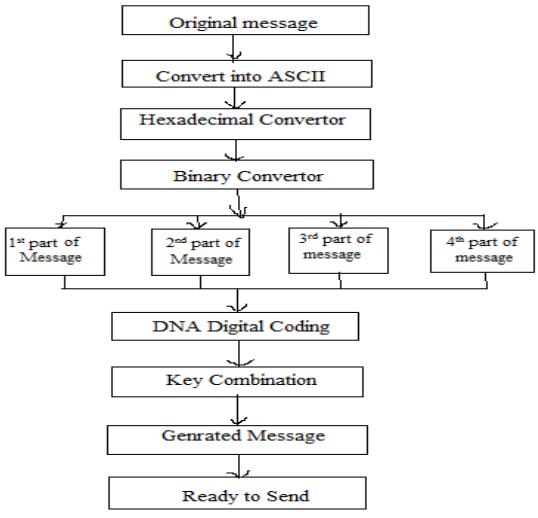

Fig. 4. Encryption Process

B. Decryption Module

The following Figure 5 shows the decryption process in which the receiver uses the key combination and DNA digital coding along with key to decode amplified message into original message.

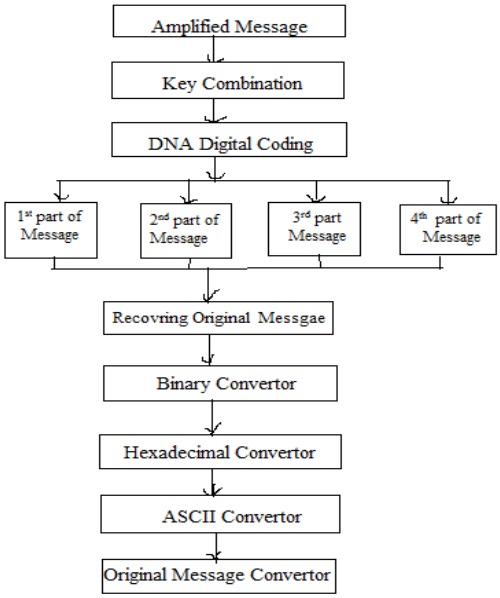

Fig. 5. Decryption Process

After the message is divided into four parts the message is converted to the binary form to hexadecimal form and then ASCII to original Message.

C. DNA Digital Coding

DNA digital coding is required in which we are considering DNA basic nucleotides assigned with binary values. The binary values use two state levels such as combinations of 0 and 1.As the DNA digital coding uses four nucleotides (A, T, G, and C) can be initialized and assigned with binary values as shown in table I [1].

TABLE I. DNA DIGITAL CODING

Binary Value	DNA Digital Coding
00	A
01	T
10	G
11	C

D. Key Combination

Using ATGC as an initial key, every base have 2 bits like A=00, T=01, G=10, and C=11.We are going to combine one base with all other bases i.e key combination and later assigning of random values can be made respectively with their equivalent pattern values in the form of binary is as shown in table II. By using this key combination table we can convert the hexadecimal value to binary form and later into DNA digital code and finally into the amplified message. From the table II, Here we are able to generate total of 72 bit keys that is 64 bits key value from key combination adding along with 8 bits of ATGC. The initial key in the form of ATGC will be used to produce a random key at the sender will be submitted to receiver. In this system, every time we can generate key at sender with particular value will be randomly changed for different communication or transaction [1].

TABLE II. KEY COMBINATION

Key Combination	Patterns	Values
AA	0101	5
AT	0011	3
AG	0001	1
AC	0010	2
TA	0110	6
TT	1111	15
TG	0111	7
TC	1001	9
GA	1010	10
GT	0100	4
GG	1000	8
GC	1100	12
CA	1110	14
CT	1011	11
CG	0000	0
CC	1101	13

To understand the scenario of encryption process flow chart we consider one example. In this example the plaintext is '**cryptography**' and performing encryption operation.

Encryption Process:

Plaintext: cryptography

We can convert the plaintext to Unicode format to get the cipher text.

Unicode:

\u0e0\u0aa\u02020\u0e0\u0aa\u0b6\u0e0\u0aa\u0bf\u0e0\u0a\u0b7

After Unicode is converted into hexadecimal value.

Hexadecimal value:

5c753065305c753061615c7530323032305c753065305c753061615c753062365c753065305c753061615c753062665c753065305c753061615c75306237

The hexadecimal value is then converted into the binary form by using the key combination. Every bit has the 2 bits. Suppose if the hexadecimal value is 5 then its binary value will be 0101.

Binary value:

01011100011101010011000001100101001100000101110001
11010100110000011000010110000101011100011101010011
00000011001000110000001100100011000001011100011101
01001100000110010100110000010111000111010100110000
01100001011000010101110001110101001100000110001000
11011001011100011101010011000001100101001100000101
10001110101001100000110000101100001010111000111010
10011000001100010011001100101110001110101001100000
11001010011000001011100011101010011000001100000101
10000101011100011101010011000001100010001101 11

After the binary value it can be now converted to DNA digital coding format.

DNA Digital coding:

From table I, we can write

TTCATCTTACAATGTTACAATTCATCTTACAATGATT
GATTTCATCTTACAAACAGACAAACAGACAATTCAT
CTTACAATGTTACAATTCATCTTACAATGATTGATTT
CATCTTACAATGAGACTGTTCATCTTACAATGTTACA
ATTCATCTTACAATGATTGATTTCATCTTACAATGAG
TGTGTTCATCTTACAATGTTACAATTCATCTTACAAT
GATTGATTTCATCTTACAATGAGACTC

Now from table II, by using the DNA digital coding and the key combination we can generate amplified message that can be transferred over the network as shown below.

Amplified Message:

11111110100111110010010101111111001001011111111010
01111100100101011100110111001111111110100111110010 0
10100110001001001010010000100100101111111110100111
11001001010111111100100101111111101001111100100101
01110011011100111111111010011111001001010111000100
10011111111110100111110010010101111111001001011111
11101001111100100101011100110111001111111110100111
10010010101110001011101111111111010011111001001010
11111110010010111111110100111110010010101110011 01
11001111111110100111110010010101110001001010 01

Decryption Process:

Now at receiver side, the receiver receives the amplified message and uses ATGC key for decryption purpose.

Amplified Message:

11111110100111110010010101111111001001011111111010
01111100100101011100110111001111111110100111110010 0
10100110001001001010010000100100101111111110100111
11001001010111111100100101111111101001111100100101
01110011011100111111111010011111001001010111000100
10011111111110100111110010010101111111001001011111
11101001111100100101011100110111001111111110100111
11001001010111000101110111111111101001111100100101
01111111001001011111111010011111001001010111001101
11001111111110100111110010010101110001001010 01

Now after receiving the amplified message it is converted to the DNA digital coding using ATGC key and key combination, to retrieve original message.

DNA Digital coding:

TTCATCTTACAATGTTACAATTCATCTTACAATGATT
GATTTCATCTTACAAACAGACAAACAGACAATTCAT
CTTACAATGTTACAATTCATCTTACAATGATTGATTT
CATCTTACAATGAGACTGTTCATCTTACAATGTTACA
ATTCATCTTACAATGATTGATTTCATCTTACAATGAG
TGTGTTCATCTTACAATGTTACAATTCATCTTACAAT
GATTGATTTCATCTTACAATGAGACTC

From the table I of DNA digital coding now we can generate the binary form. Suppose if TT is DNA digital code we can convert to binary form 0101.

Binary value:

01011100011101010011000001100101001100000101110001
11010100110000011000010110000101011100011101010011
00000011001000110000001100100011000001011100011101
01001100000110010100110000010111000111010100110000
01100001011000010101110001110101001100000110001000
11011001011100011101010011000001100101001100000101
11000111010100110000011000010110000101011100011101
01001100000110001001100110010111000111010100110000
01100101001100000101110001110101001100000110000101
10000101011100011101010011000001100010001101 11

Now we can convert binary form to hexadecimal value. Suppose if binary value is 0101 then we get the hexadecimal value as 5.

Hexadecimal value:

5c753065305c753061615c7530323032305c753065305c7530
61615c753062365c753065305c753061615c753062665c7530
65305c753061615c75306237

The hexadecimal value is then converted to the Unicode value to get original message.

Unicode:

\u0e0\u0aa\u02020\u0e0\u0aa\u0b6\u0e0\u0aa\u0bf\u0e0\u0a
a\u0b7

Finally the Unicode value is converted to the original message as below.

Plaintext: cryptography

The table III shows the time and space analysis in milliseconds (ms) and in Kilo Bytes (KB) respectively, for encryption and decryption processes along with input size on disk.

TABLE III. TIME AND SPACE ANALYSIS

Input	Size on Disk	Time Taken for Encryption	Time Taken for Decryption
Plain Text	5 letter	8325.816ms	5.346728ms
Text File	10KB	5529.8784ms	4.54851ms
Image	90KB	7397.661ms	5223.019ms
Audio	490KB	14580.631ms	2243.8176ms

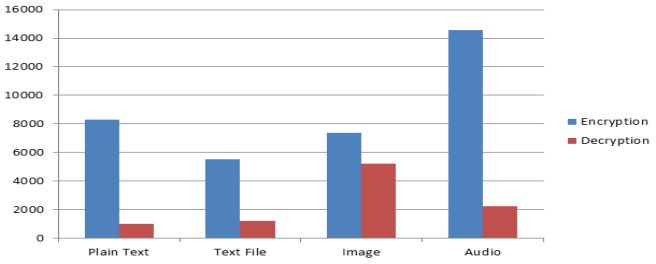

Fig. 6. Performance Analysis

The above figure 6 shows the performance analysis in terms of time taken for the encryption process and decryption process in milliseconds for different inputs such as plain text, text file, image, audio. Note that the decryption time taken will be less compared to encryption time taken with various data. Hence proposed approach is energy efficient.

CONCLUSION

The proposed approach a Dynamic DNA for key based approach is able to accept various forms of data such as characters, text file, image and audio. Random key will be generated at the sender every time will be used for decrypting the ciphertext at the receiver makes the approach very strong against various attacks. The proposed approach is difficult to break by a common cryptanalysis techniques.This approach provides two-stage security, improved reliability and better time and computational complexities.

ACKNOWLEDGMENT

Gratitude is the noblest gesture of one towards another. Any endeavor undertaken is incomplete without expressing gratitude to the people who made it possible. I take this opportunity to express my profound gratefulness and regards to my guide and mentor Dr. Latha Parthiban for her exemplary and constant encouragement. I am greatly indebted to her for valuable timely inputs and rigorous reviews and suggestions.

I extend thankfulness to Dr. Veena Desai, Prof., Dept. of E&C, GIT, Belagavi for her timely suggestions and motivation.

REFERENCES

[1] Prajapati Ashishkumar B and Prajapati Barkha, "Implementation Of DNA Cryptography In Cloud Computing and Using Socket Programming", IEEE, January 2016.

[2] Naveen Jarold, P Karthigaikumar, N M Sivamangai, Sandhya R, Sruthi B Ashok, "Hardware Implementation of DNA based Cryptography", Conference on Information and Communication Technologies, IEEE, pp. 696-700,2013.

[3] M R Saranya,Arun K Mohan and K.Anusudha, "Algorithm for Enhanced Image Security Using DNA and Genetic Algorithm", IEEE, April 2015.

[4] Ajit Singh and Reena Singh, "Information Hiding Techniques Based on DNA Inconsistency: An Overview", IEEE, May 2015.

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

[6] Tushar Mandge and Vijay Choudhary, "A DNA Encryption Technique Based on Matrix Manipulation and Secure key Generation Scheme",IEEE, 2013.

[7] Anchal Jain, "Adaptive Key Length Based Encryption Algorithm using DNA Approach", IEEE ,2013.

[8] Zhang Yunpeng, Zhu Yu, Wang Zhong and Richard O.Sinnott, "Index-Based Symmetric DNA Encryption Algorithm",IEEE,pp. 2290-2294,2011.

[9] G. Cui, L. Qin , Y. Wang , X. Zhang, "An Encryption Scheme Using DNA Technology", IEEE, 2008.

[10] Hamza Hammami, Hanen Brahmi, Sadok Ben Yahia "Secured Outsourcing Towards a Cloud Computing Environment Based on DNA Cryptography",IEEE page no. 31-36,2018

[11] S.V.Keerthana Priya, S.J.Saritha , "A Robust Technique to Generate Unique Code DNA Sequence",IEEE, page no.3815-3820,2017

[12] Mona Sabry, Mohamed Hashem, Taymoor Nazmy, Mohamed Essam Khalifa , "Design of DNA-based Advanced Encryption Standard (AES)",IEEE, pp. 390-397,2015.

[13] Deepak Kumar and Shailendra Singh ,"Secret Data Writing Using DNA Sequences",IEEE,pp. 402-405,2011.

[14] Guangzhao Cui , Limin Qin, Yanfeng Wang and Xuncai Zhang, "An Encryption Scheme Using DNA Technology",IEEE,pp. 37-41,2008.

[15] Vikas Sagar and Krishan Kumar, "A Symmetric Key Cryptography using Genetic Algorithm and Error Back Propagation Neural Network",IEEE, pp. 1396-1391,2015.

[16] Nisha Kumari, Akhil Kaushik, "A Three Dimensional Model for Image Based Information Security", International Conference on Computational Intelligence and Communication Networks, 2014.

[17] Sudesh, Akhil Kaushik, Satvika Kaushik. "A two stage hybrid model for image encryption and compression to enhance security and efficiency", International Conference on Advances in Engineering & Technology Research (ICAETR – 2014).