Implementation of DNA Cryptography based on Dynamic DNA Sequence Table using Cloud Computing

Vinay S
Computer Science and Engineering
Alva's Institute of Engineering and Technology , Mijar,Karnataka

Adarsh Pujar
Computer Science and Engineering
Alva's Institute of Engineering and Technology, Mijar,Karnataka

Ankith
Computer Science and Engineering
Alva's Institute of Engineering and Technology, Mijar,Karnataka

H.Akshay Kedlaya
Computer Science and Engineering
Alva's Institute of Engineering and Technology, Mijar,Karnataka

Vasudev S Shahapur
Associate Professor
Computer Science and Engineering
Alva's Institute of Engineering and Technology, Mijar,Karnataka

***Abstract*—Securing data in cloud has major issues such as processing, compression and speedup computation. To address this issue, there are many approaches applied by various researchers worldwide to strengthen security of the stored data on cloud computing. This paper is one of such cryptographic technique to overcome the security problem. The proposed technique considers dynamic sequence table to assign ASCII characters initially and then, a finite number of iterations are applied and Later On One-Time-Pad (OTP) is applied on the modified binary value. Obtained OTP Cipher text is again processed through genomic conversion. And finally, it is converted into compressed cipher text using amino acid table which consists of protein sequences that increases the confusion of the cipher text. The cipher text obtained at the last will contain the information that will provide enhanced security against the intruder's attack.**

Keywords— Cloud computing, DNA cryptography, dynamic DNA sequence table, Iteration process, Amino acid table;

I. INTRODUCTION

Cryptography is a process of achieving confidentiality in data or message transfer. It is also a process of transforming the sender's message to a secret format called the cipher text that only the intended receiver will get know the secret message. Now a day's achieving complete security is an issue of the data/message transfer. In that regard, DNA Cryptographic technique has been introduced which includes DNA based encryptions and decryptions. DNA has huge computing power and enormous parallelism. DNA also has massive storage capacity. Several algorithms have also been introduced in DNA cryptography such as encryption to overcome the security issues.

Cloud computing makes computer system resources, especially storage and computing power, available on demand without direct active management by the user. Large clouds, predominant today, often have functions distributed over multiple locations from central servers. If the connection to the user is relatively close, it may be designated an Edge server. Cloud-computing providers offer their services according to different models, of which the three standard models are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). These models offer increasing abstraction; they are thus often referred as a layers in a stack: infrastructure-, platform- and software-as-a-service, but these need not be related. For example, one can provide SaaS implemented on physical machines (bare metal), without using underlying PaaS or IaaS layers, and conversely one can run a program on IaaS and access it directly, without wrapping it as SaaS. A. Infrastructure-as-a-service: "Infrastructure as a service" (IaaS) refers to online services that provide high-level APIs used to deference various low-level details of underlying network infrastructure like physical computing resources, location, data partitioning, scaling, security, backup etc. A hypervisor runs the virtual machines as guests. Pools of hypervisors within the cloud operational system can support large numbers of virtual machines and the ability to scale services up and down according to customers varying requirements.

B. Software-as-a-service: The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

C. Platform-as-a-service: The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

The existing DNA cryptographic techniques such as static DNA sequence table focuses on fixed DNA sequence for each ASCII character. DNA cryptography requires high computational complexity which are complex and costly.

Hence, encryption becomes more difficult. To overcome these issues, this paper proposes DNA cryptography using dynamic DNA sequence table in Cloud computing.

II. RELATED WORKS

As per Atanu Majumder, Abhishek Majumder, Tanusree Podder, Nirmala Kar and Meenakshi Sharma [1]: Secure data communication is the main issue in the field of message transmission. Hence, they have proposed an encryption technique called DNA based Message Encoding which will enhance the message security. The technique which they proposed is applicable to all types of files. The algorithm includes two phases: round key selection and message encryption. The advantage of their work is that it provides stronger protection against the various intruder's attacks.

As per Prajapati Ashish Kumar B and Prajapati Barkha [2]: They have introduced a new technique called Bi-directional DNA Encryption Algorithm (BDEA) to overcome the data security issues. They have implemented a 2-layer security for ASCII character sets. Their proposed work focuses on BDEA algorithm to use with the Unicode characters. But they are unable to measure the possible attacks.

As per Mehdi Hojabri and Mona Heidari [3]: They first perform the concept of Kerberos authentication services. In the next step Authenticate Server (AS) of Kerberos verifies users and creates the ticket granting ticket and session key and sends it to the users. The next step users send the ticket granting ticket and session key to Ticket Granting Server (TGS) for getting the service. Then TGS send ticket and session key to the user. In final step the user sends the request service to cloud service provider for using the cloud service and also cloud service, provide service to users. After that user can use the cloud service provider's service. But for more security they performed RSA algorithm for encryption & decryption and then they use Digital Signature for Authentication.

As per Mr. Prashanth Rewagad and Ms. Yogita Pawar [4]: They have proposed a technique that make use of digital signature, Diffie Hellman key exchange and Advanced Encryption Standard (AES) algorithm for authentication, data security and verification. They have used AES algorithm to protect sensitive data in the cloud. Their proposed architecture makes it tough for the intruders to hack the security system.

As per Uma Somani, Kanika Lakhani and Manish Mundra [5]: They have proposed a technique which includes digital signature for authentication and RSA Encryption algorithm to enhance the security of data in the cloud. Here, In digital signature, software will crunch down the data into few lines or message digest using hashing algorithm. Then encrypts the message digest with its private key. Then it will produce the digital signature. Then software decrypts the digital signature into message digest with public key of sender's and its own private key. This proposed technique helps to detect forgery and tampering.

III. PROBLEM STATEMENT

In cloud computing, there are some security problems such as confidentiality, integrity, etc. The existing traditional encryptions such as Polymerase Chain Reaction (PCR) amplification technology of encrypted DNA does not have enough key space which unconditionally hampers the security issues. And encryption increasingly becomes tougher if too many keys and more complex keys are used. There are many attacks in cryptography that effects the security in cloud such as Brute-force attack, Symmetric block cipher attacks, Stream cipher attacks, Hash function attacks, Message Authentication Code (MAC) attacks, Birthday attack, Man-in-the-middle attack.

IV. PROPOSED METHOD

This paper proposes a technique called DNA cryptography using dynamic DNA sequence table. Multiple steps of conversions and sequences in this technique increases the secrecy of ciphertext.

- **DNA Digital Coding**

The binary digital coding encoded by two state 0 or 1 and a combination of 0 and 1. But DNA digital coding can be encoded by four kind of base as shown in table 1. That is ADENINE (A) and THYMINE (T) or CYTOSINE (C) and GUANINE (G).

Table 4.1. DNA Digital Coding

Binary value	DNA digital coding
00	A
01	T
10	G
11	C

Here in this work, we are using ATGC as a key. Every bit has 2 bits like A=00, T=01, G=10, and C=11 and by using ATGC, key combinations is generated and give numbering respectively that is given into table.

Table 4.2. Dynamic DNA sequence table

DNA Base sequence	DNA Base sequence	DNA Base sequence	DNA Base sequence
AAAA –y	TCCG –g	GCAA –m	TCTC –s
ATAA–W	TACC –E	GAAG–K	TATC –Q
AGAG –{	TTCC –2	GTAT –8	TTTA - ?
ACTG- b	TGCG - +	GGAT - *	TGTA - @
AATT –z	TCGT –h	GCTT –n	TCCC –t
ATTT –X	TAGA –F	GATA –L	TACG –R
AGTA –[	TTGG - 3	GTTG –9	TTCG - /
ACCC- c	TGGC - =	GCCG –o	TGCC - !
AACC –A	CCAG –i	GACG–M	CCTT –u
ATCG –Y	CAAT –G	GTCG –<	CATC –S
AGCG - }	CTAT –4	GGCC - ^	CTTC - :
ACGA –d	CGAA - _	GCGC –p	CGTA - ~
AAGG–B	CCTA –j	GAGG–N	CCCC –v
ATGC –Z	CATG –H	GTGT - >	CACC –T
AGGG -]	CTTG –5	GGGA - %	CTCG - ;
TCAT–e	CGTT - -	ACTC –q	CGCG - `
TAAT–C	CCCG –k	AATA –O	GCTA –w
TTAA–0	CACG –I	ATTA - ,	GATT –U
TGAA - |	CTCC –6	AGTT - $	GTTC – "
TCTG–f	CGCC -)	ACCG –r	GGTC - €
TATG–D	CCGG –l	AACG –P	GCCC –x
TTTT–1	CAGT –J	ATCC - .	GACC –V
TGTT- \	CTGA –7	AGCC - #	GTCC – '
ACAT- a	CGGG –(	GGTG - &	GGCG - £

The dynamic table is initiated with random character for each DNA sequence (A, T, G and C) as shown in Table 4.1.

Initially, both the sender and the receiver have the same DNA sequence table.

Table 4.2 represents the dynamic DNA sequence table during iterations. In initial table, DNA base positions can be changed according to users' wish while maintaining the format. Biological DNA replication to RNA i.e. mRNA, tRNA and conversion of RNA to amino acid table to get ciphertext are used to folding up the message into multiple times.

Table 4.3. Amino Acid Table

Symbol	Protein sequence
A	GCT, GCC, GCA, GCG
B	TAA,TAG
C	TGT, TGC
D	GAT, GAC
E	GAA, GAG
F	TTT, TTC
G	GGT, GGC, GGA, GGG
H	CAT, CAC
I	ATT, ATC, ATA
J	TGA
K	AAA, AAG
L	CTT, CTC, CTA, CTG
M	ATG
N	AAT, AAC
O	TTA,TTG
P	CCT, CCC, CCA, CCG
Q	CAA, CAG
R	CGT, CGC, CGA, CGG
S	TCT, TCC, TCA, TCG
T	ACT, ACC, ACA, ACG
U	AGA, AGG
V	GTT, GTC, GTA, GTG
W	TGG
X	AGT, AGC
Y	TAT
Z	TAC

Converting from RNA to amino acid uses protein sequence from Table 4.3 where the protein sequence is replaced by a capitalized letter i.e. 'A' for "GCT". According to the table, 26 letters are mapped to protein sequences.

There are three approaches included in this paper:

A. Iteration Process

During iterations, firstly initial position is checked and increased by a specific number. Then half of DNA bases of the table are filled at odd/ even positions of table sequentially according to initial position of each cycle and the rest of DNA bases are filled by opposite type position sequentially.

Fig 4.1. shows the iteration process of this technique.

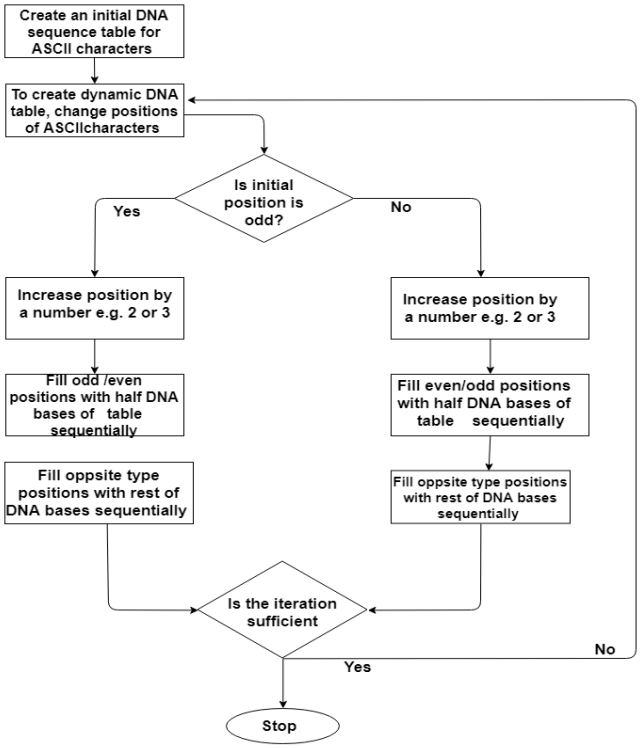

Fig 4.1. Iteration Process

B. Encryption Process

Step 1: At first, an initial DNA sequence table of ASCII characters is generated and known by both the sender and the receiver. Now it is converted to dynamic DNA sequence table. The table is destroyed after a period of time. The iteration number is sent to the receiver.

Step 2: Input plaintext is organized as the DNA sequence from the table. This is called the encoding format of input plaintext. Thus the plaintext is converted to DNA format.

Step 3: DNA sequences are converted to 2-bit binary value using the following mapping: A-00, T-01, C-10 and G-11.

Step 4: OTP key is generated which is fully random. The key size is same as the binary value of step 3.

Step 5: Binary XOR operation is applied between 2-bit binary value of input and OTP key.

Step 6: 00-A, 01-T, 10-C and 11-G mapping is applied to the XOR sequence. Thus the DNA representation of binary XOR value is obtained.

Step 7: Biological mRNA conversion process is applied on DNA sequence. The Thymine (T) is replaced with Uracil (U).

Step 8: mRNA is transferred to tRNA by replacing A→U, U→A, G→C and C→G. This is a real-life biological conversion process.

Step 9: tRNA is divided into two parts (1st and 2nd). 1st part and 2nd part interchange their position.

Step 10: 'A' or 'AC' is added with the tRNA sequence to replace amino acid table (if sequence is not divided by 3).

Step 11: Uracil (U) is replaced by Thiamine (T). This is called the reverse simulation process and the simulated tRNA sequence is obtained.

Step 12: Then the protein sequence is obtained from amino acid table. The table is taken as a sample of protein sequence. For each letter of alphabet, a number between 0 and 25 is assigned where A = 0, B = 1, …, Z = 25. By replacing the simulated tRNA sequence by letters using this table, we get the final ciphertext.

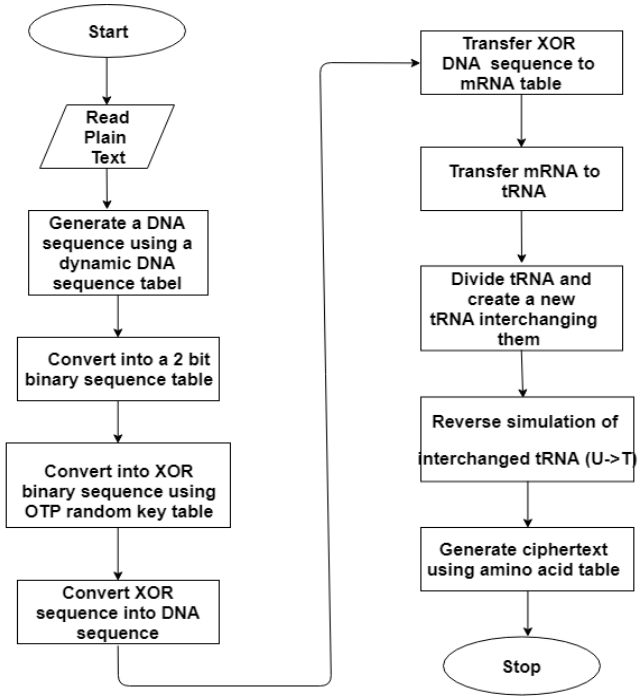

Fig 4.2. Encryption Process

C. Decryption Process

Step 1: Ciphertext and required iterations is received through a media. DNA encoding table and amino acid table are regenerated with the help of received information.

Step 2: Ciphertext and amino acid table are matched up. For each letter of the table the protein sequence is stored at the time of encryption. From the matching result of protein sequence and letter of ciphertext, simulated tRNA sequence of ciphertext is obtained.

Step 3: Interchanged tRNA sequence is formed by replacing Thiamine (T) with Uracil (U) from reverse simulated form.

Step 4: 'AC' or 'A' is removed from interchanged tRNA if added.

Step 5: By dividing the interchanged tRNA and changing their positions, the original tRNA is obtained.

Step 6: Then mRNA is acquired from the tRNA by replacing A→U, U→A, G→C and C→G.

Step 7: mRNA is transferred to XOR DNA sequence (by replacing 'U' with 'T').

Step 8: XOR DNA sequence is converted to XOR binary sequence.

Step 9: From XOR binary sequence, random sequence is generated.

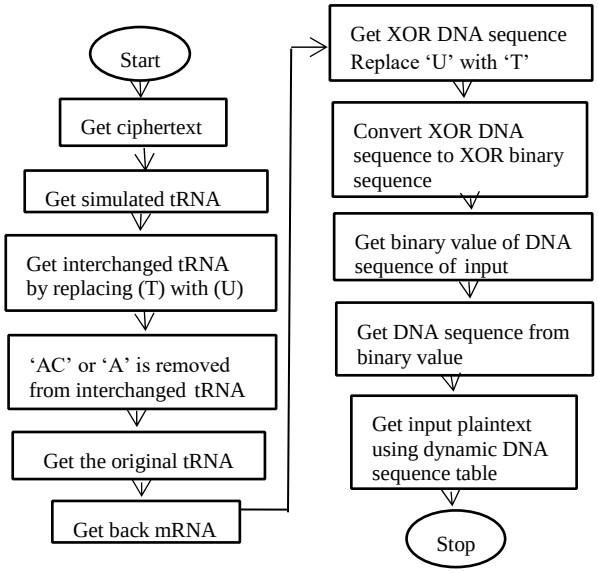

Fig 4.3. Decryption Process

V. CONCLUSION

The proposed DNA cryptographic technique using dynamic DNA sequence table in cloud computing increases the level of data security. In order to breach the security, the attacker needs to perform all kinds of checking which is not possible. The proposed system ensures the better data security compared to other techniques. And it is quite powerful against cryptographic attacks such as Brute-force attack, Birthday attack, Man-in-the-middle attack.

REFERENCES

[1] Atanu Majumder, Abhishek Majumdar, Tanusree Podder, Nirmalya Kar, Meenakshi Sharmas, "Secure Data Communication and Cryptography Based on DNA Based Message Encoding" 2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT).

[2] Prajapati Ashishkumar B., Prajapati Barkha , "Implementation of dna cryptography in cloud computing and using socket programming" 2016 International Conference on Computer Communication and Informatics (ICCCI -2016), Jan. 07 – 09, 2016, Coimbatore, INDIA.

[3] Mehdi Hojabri& Mona Heidari"Union of RSA algorithm, Digital Signature and KERBEROS in Cloud Computing" International Conference on Software Technology and Computer Engineering (STACE-2012).

[4] PrashantRewagad, YogitaPawar, "Use of Digital Signature with DiffieHellman Key Exchange and AES Encryption Algorithm to Enhance Data Security in Cloud Computing" 2013 International Conference on Communication System and Network Technologies (IEEE Computer Society).

[5] A. Atito, A. Khalifa and S. Z. Reda, "DNA-Based Data Encryption and Hiding Using Playfair and Insertion Techniques", Journal of Communications and Computer Engineering, Volume 2, Issue 3, pages 44: 49, 2012.

[6] Uma Somani, Kanika Lakhani, ManishaMundra, "Implementing Digital Signature with RSA Encryption Algorithm to Enhance the Data Security of Cloud in Cloud Computing"-2010 IEEE 1st International Conference on Parallel, Distributed and Grid Computing (PDGC2010).