International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072

Enhance Data Storage Security DNA Cryptography in Cloud

Reyona Mendez

Department of Computer Applications, Sree Narayana Guru Institute of Science and Technology, Ernakulam, India

---

Abstract – *Cloud computing has recently reached popularity and developed into a major trend in IT. We perform such a systematic review of cloud computing and explain the technical challenges facing in this paper. There have been several approaches applied by various researchers worldwide to strengthen security of the stored data on cloud computing. The Bi-directional DNA Encryption Algorithm (BDEA) is one such data security techniques, ASCII character set, ignoring the non-English user of the cloud computing. Existing system work focuses on enhancing the BDEA to use with the Unicode characters. This proposed system will include some data protection, various techniques evolved through years for Ciphers, Cryptography, Steganography and recently DNA based encryption for security is the trend. DNA cryptography was a breakthrough in the field of security which uses bio-molecular concepts and give us new hope of unbreakable algorithms but the concepts need to be exploited more especially in the cloud computing. This paper discuss cloud computing features, service models, and security issues and proposes a DNA based encryption algorithm for securing data in cloud environment which will be cost effective and secure by using bio-computational techniques . 5 level security uses in this proposed system DNA Digital Coding, Key Combination, RDNA (Random DNA), DNA Steganography techniques along with binary coding rules which make algorithm secure as it is an additional layer of biosecurity than conventional cryptographic techniques and also Morse code*

Key Words: Cloud Computing, Integrity, Confidentiality, Data Security, DNA Digital Coding, Key Combination, DNA Steganography techniques, RDNA (Random DNA), Morse code.

1. INTRODUCTION

Cloud computing is the latest technology in the field of distributed computing. It provides various online and on demand services for data storage, network services, platform services etc. Many organizations are unenthusiastic to use cloud services due to data security issues as the data resides on the cloud services provider's servers. To address this issue, there have been several approaches applied by various researchers worldwide to strengthen security of the stored data on cloud computing. The Bi-directional DNA Encryption Algorithm (BDEA) is such a data security techniques.

However, the existing technique focuses only on the ASCII character set, ignoring the non-English user of the cloud computing and to create a secure data storage in cloud using socket programming. The user has a provision to upload files and images into cloud server. Before do this, the server give access permission to access. Cloud computing has recently reached popularity and developed into a major trend in IT. We perform such a systematic review of cloud computing and explain the technical challenges facing in this paper. In Public cloud the "Pay per use" model is used. In Private cloud, In Hybrid cloud, the computing services is consumed both the private cloud service and public cloud service. Cloud computing has three types of services. Software as a Service (SaaS), in which customer prepared one service and run on a single cloud, then multiple consumer can access this service as per on demand. Platform as a Service (PaaS), in which, it provides the platform to create application and maintains the application. Infrastructure as a Service (IaaS), as per term suggest to provides the data storage, Network capacity, rent storage, Data centres etc. It is also known as Hardware as a Service (HaaS).

1.1 Principle

The main objectives of the project is to create more secure data storage in cloud using socket programming. The main modules are server, user and cloud. The objective of the project is 5 level security which is DNA Digital Coding, Key Combination, DNA Steganography techniques, RDNA (Random DNA), Morse code.

2. RESEARCH AREA

This study discuss about the security issues in cloud that facing in this developed world. This literature is related to Data security challenges in cloud, DNA computation, DNA Cryptography, DNA Steganography, and also Morse code. New fact are cloud computing problems arising is security, even though satisfactory solutions for many still will require significant developments. The combined contemporary and historical viewpoints allow us to identify a number of research topics that deserve more attention. On the other hand, we argue that two facets are to some degree new and fundamental to cloud computing: the complexities of multi-party trust considerations, and need for ensure mutual auditability. Cloud computing defined by NIST is that it is a model for ubiquitous enabling, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. The cloud model is composed of five essential characteristics which are on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service, Cloud service models are also known as SPI model. They are Software as a Service (SaaS), Platform as a Service (PaaS) and

Infrastructure as a Service (IaaS). Depending upon the customer requirements the Cloud services can be deployed as Private cloud, Community cloud, Public cloud and hybrid cloud. DNA based encryption algorithm for securing data in cloud environment which will be cost effective and secure by using bio-computational techniques. The suggested algorithm uses indexing and DNA steganography techniques along with binary coding rules which make algorithm secure as it is an additional layer of biosecurity than conventional cryptographic techniques. The two-level security level needs data storage into a medium-sized medium that has also been inserted into another media. Two covers can be the same or different depending on the features and applications. DNA cryptography is a neutral source that utilizes the evolution of the molecular and gives us the new hope of the algorithm that cannot be explored but the principles need to be even more explored in DNA analyzes. Focus on DNA security protection by using an image design using DNA. This template explains how to apply to DNA Cryptosystem for securing data. The Securing Portable Document Format file Using Extended Visual Cryptography (SPDFUEVC) technique proposes efficient storage to achieve data confidentiality and integrity verification with minimal computation, time complexity and storage space. Data security and privacy protection issues are relevant to both hardware and software in the cloud architecture. This study is to review different security techniques and challenges from both software and hardware aspects for protecting data in the cloud and aims at enhancing the data security and privacy protection for the trustworthy cloud environment. In this paper, we make a comparative research analysis of the existing research work regarding the data security and privacy protection techniques used in the cloud computing. DNA cryptography is a new era of enhanced security for the data transfer using the internet. DNA cryptography enhanced the cryptography in terms of time complexity as well as capacity. It uses the DNA strands to hide the information. The repeated sequence of the DNA makes highly difficult for unintended authority to get the message. This paper discusses DNA cryptography and the difference between the traditional and the DNA cryptography. This paper also brief the various work done in the field of the DNA cryptography.

3. RELATED WORK

In Cloud computing data security is prepared by the Authentication, Encryption and Decryption, Message authentication code, Hash function, and Digital signature and so on. Diffie-Hellman algorithm is used to generate keys for key exchange step. Then digital signature is used for authentication, thereafter AES encryption algorithm is used to encrypt or decrypt user's data file. Diffie- Hellman key exchange algorithm is vulnerable to main in the middle attack. The most serious limitation is the lack of the authentication. In final step the users send the request service to cloud service provider for using the cloud service and also cloud service, provide service to users. After doing this step user can used the cloud service provider. But for more security

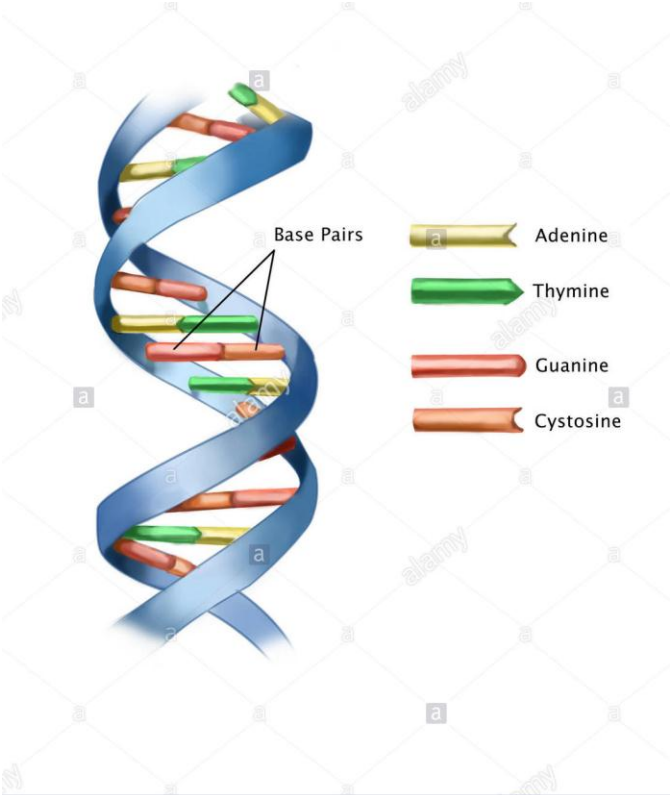

Fig-1: DNA Base strands

they performed RSA algorithm for encryption and decryption and then they use Digital Signature for Authentication. RSA algorithm and Digital signature are used for secure communication. The related work is to ensure two level security to the cloud data storage which is Key combination and the DNA Encryption. The disadvantages of this system is our confidential data may lost or can be theft by the third

4. PROPOSED WORK

4.1Problem Statement:

The existing system describes about the cloud computing, basics of cloud computing and security problems occurs in cloud. Here, the Bi-serial DNA encryption algorithm is performing, that providing the five level of security

(1)DNA Digital Coding

(2)Key Combination

(3)DNA Stenography techniques

(4) R-DNA [Random DNA]

(5)Morse code.

DNA, Deoxyribonucleic acid is a double-stranded helix of nucleotide with each nucleotide containing one of four bases A, G, C, T where A stands for adenine, G for guanine, C for cytosine and T for thymine respectively Methodology is

proposed is for ensure the security of cloud storage using DNA based encryption technique. DNA based coding, encryption technique, DNA Stenography indexing and Morse code these all method are proposed for securing data in the cloud. The Proposed DNA Cryptography technique is different from that of the DNA cryptography which uses real DNA Sequences or Oligos, as the computations performed are using the digital DNA. A DNA sequence consists of four nucleic acid bases A (adenine) C (cytosine) G (guanine) T (thymine)

Table -1: Binary code table

BASES	BINARY CODE
A	00
G	01
C	10
T	11

4.2Algorithms used:

Algorithm 1:

1: Begin

2: Input: Data to be stored D,Random DNA R-DNA.

3: Select the data D, which has to be stored securely in the cloud, Convert the data into binary.

4: Convert the binary data into DNA sequence based on the DNA coding rule, which generates a digital DNA to D'DNA.

5: Create a random DNA strand by selecting DNA sequences from the digital databases to R-DNA.

6: Select the R-DNA and index it. Select the coding and non-coding regions randomly or based on the index values.

7: Convert indexed R-DNA into short fragments based on the length of D'DNA base pair, and a key value based on D'DNA.

8: Remove the non-coding region and the generated DNA sequence is used as a cover for adding D'DNA.

9: Insert the D'DNA into non-coding regions of the generated R-DNA based on the index positions or random position depending on the indexing rule selected.

10: The resultant DNA sequence generated by DNA Stenography is then converted to Morse code

11: Upload the encrypted data in the binary form and store it in the cloud.

12: Output: Encrypted Data

13: End

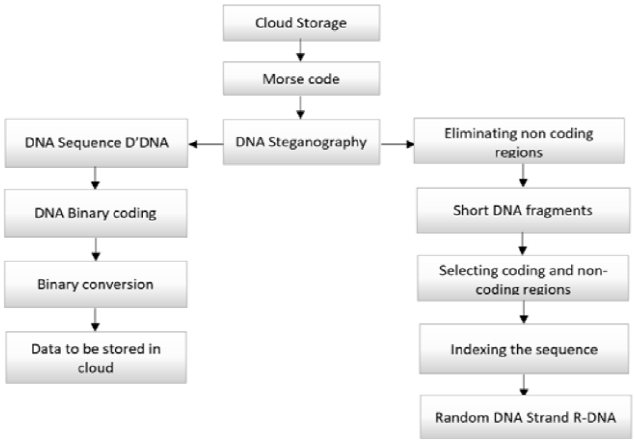

Fig -2: Encryption process

Algorithm 2:

1: Begin

2: Input: Encrypted Data.

3: Extract the encrypted data from the cloud which is in form of dits and dashes.

4: Covert Morse code back to the DNA sequence form as per rules that is DNA sequence which is a combination of R-DNA and D'DNA.

5: Based on the index value position, select the coding and non-coding regions of the DNA.

6: Retrieve DNA fragments from the non-coding region.

7: Extract and separate D'DNA and R-DNA from the DNA Sequence.

8: Append the fragments of D'DNA and apply DNA coding rule to get the binary data.

9: Convert the binary to ASCII. 10: Convert ASCII to text.

11: Generates the original data.

12: Output: Original Data and Random DNA

13: End

4.3Process step by step:

Step 1. Let us consider the data to be stored in the cloud by the user is, Confidential Data.

Step 2. Convert the data to be stored into the Binary form. Confidential Data -> 01001101 01111001 01000011 01101111 01101110 01100110 01101001 01100100 01100101 01101110 01110100 01101001 01100001 01000100 01100001 01110100 01100001

Step 3. Apply DNA binary Coding as per key combination Table to binary data of Table follows or on generates in DNA

form. A, G, C and T values can be altered and used as per user convenience depending on the selected

Step 4. Select Random DNA -> R-DNA Sample RDNA selected is: ACTGCTGAGAGTTGAGCTCACCCTC
AGTCCCTCACAGTTCCACACTGCCT
The random DNA is generated by downloading the sequence available from NCBI

Step 5. Indexing the R-DNA.
1C2T3G4C5T6G7A8G9A10G11T12
13G14A15G16C17T18C19A20C21C22C23
24C25A26G27T28C29C30C31T32

Step 6. Select the coding and non-coding regions randomly or based on the index positions. This sample is demonstrated using a random selection of coding and non-coding regions. Depending on the security of the data, complexity of the algorithm can be increased by defining index rules.

Step 7. Insert the D'DNA into non-coding regions of RDNA.

Step 8. Insert D'DNA into respective non-coding index positions. Random DNA selected act as cover medium to insert D'DNA and performs DNA steganography.
GATG GTCG GAAT GCTT GCTC GCGC GCCG GCGA
GCGG GCTC GTGA GCCG GCAG GCTA GAGA GCAG
GTGA GCAG -> D'DNA
Replacing Non-Coding region bases with 4bases (key value) of D'DNA per each base of non-coding region:
T6G7A8G9A10G11T12T13G14A15
G16C17T18C19A20C21C22C23 -> Cover DNA(noncoding region)
(GATG)6(GTCG)7(GAAT)8(GCTT)9(GCTC)10(GCGC)11(GCCG)12(GCGA)13(GGTT)14(ATGG)15(GAAT)16(GGAA)17(GGTT)18(GCCT)19(GAGA)20(GCAG)21(GTGA)22(GCAG)23 ->DNA Steganography

Step 9.Generate the DNA Sequence.
A1C2T3G4C5(GATG)6(GTCG)7(GAAT)8(GCTT)9(GCTC)10(GCGC)11(GCCG)12(GCGA)13(GCGG)14(GCTC)15(GTGA)16(GCCG)17(GCAG)18(GCTA)19(GAGA)20(GCAG)21(GTGA)22(GCAG)23T24C25A26 G27T28C29C30C31T32 -> In indexed form
ACTGCGATGGTCGGAATGCTTGCTCGCGCGCCGGCGAGCGGGC
TCGTGAGCCGGCAGGCTAGAGAGCAGGTGAGCAGTCAGTCCCT
CACAGTTCCACACTGCCT -> Stego DNA

Step 10. Converting DNA encrypted DNA to Morse code
- .... / .. ... / .- / ... .. – .–. .-.. . –..– / -.– . - / -.-. .... .- .-.. .-.. . -. –. .. -. –. / -.-. .- -.-. .... . .-.-

Step 11. Store the data into cloud in the binary form or in integer form by converting the binary data to integers depending on the convenience of user.

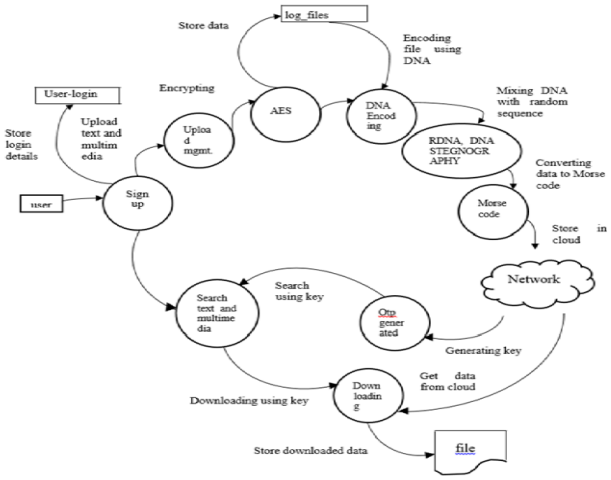

Fig -3: Data flow diagram

4.4Encryption and decryption by AES:

Rounds in AES Length of the input output block and the State is 128 for AES algorithm. This is represented by Nb= 4, which reflects the number of 32-bit words (number of columns) in the State. For the AES algorithm 128, 192, or 256 bits is the length of the Cipher Key, K. The key length of the block is denoted by Nk and its value is 4, 6, or 8. This value reflects the number of 32-bit words (number of columns) in the Cipher Key. For the AES algorithm, during the execution of algorithm the numbers of rounds to be performed are dependent on the key size. Nr is used to represent the number of rounds. AES algorithm uses a round function for both its Cipher and Inverse Cipher. This function is composed of four different byte-oriented transformations: A) Using a substitution table (S-box) byte substitution, B) By different offsets shifting rows of the State array, C) mixing the data within each column of the State array, and D) adding a Round Key to the State. No.of round keys generated by key expansion algorithm is always one more than the actual no. of rounds present in the algorithm.

The Algorithm:

The input (block size Nb, also known as plaintext) of the AES algorithm is converted into a 4 x 4 array, called a state. Four transformations, Add Round Key, Sub Bytes, Shift Rows and Mix Columns, perform various operations on the state to calculate the output state (the final cipher text). Except for AddRoundKey each of these operations. Transformation in AES to perform all these transformations above, some mathematical operations are needed to understand which are given as below.

AddRoundKey: The AddRoundKey routine is simple XOR addition of round key and a portion of expanded key into plaintext.

Sub byte: Sub byte is the SBOX for AES. It operates on each byte in the state and performs a nonlinear substitution in the GF (28) field, which is what makes AES a non-linear cryptographic system. In order to be invertible each value of b' must be generated from a unique value of b. A look up table

can also be implemented for Sub Bytes. Sub Byte operation performs an affine transformation on the inverse of byte b, and adds it to 0xC6.

Shift Rows: operates on individual rows of the state. It provides diffusion throughout the AES algorithm. The first row is not changed. The second row is shifted one byte to the left, with the left most byte wrapping around. The third row shifts two bytes to the left, and the fourth row shifts three bytes to the left with appropriate wrapping to the right. This description is for AES-128, the number of shifts for each row changes based on the key size.

Mix Columns: Mix Columns operates on individual columns of the state .It provides diffusion throughout the AES algorithm. The columns are considered polynomials over GF (28) and multiplied modulo x4+1 with a(x) where a(x) = 03x3 + 01x2 + 01x + 02 NOTE: x4+1 is relatively prime to a(x). This can be represented as a matrix equation.

4.5DNA Digital Coding:

In information science, 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 below. There are possibly 4! =24 pattern by encoding format like (0123/ATGC).

1. Adenine (A);
2. Thymine (T);
3. Cytosine (C);
4. Guanine (G).

The easiest way to encode is to represent these four units as four figures: 1. A (0) –00; 2. T (1) –01; 3. C (2)–10; 4. G (3)–11. Obviously, by these encoding rules, there are 4!= 24 possible encoding methods. For DNA encoding, it is necessary to reflect the biological characteristics and pairing principles of the four nucleotides. Based on this principle, we know that: A (0) – 00 and G (3) – 11 make pairs, T (1) –01 and C(2) – 10 make pairs.

4.6 Key Combination:

Here in this work, we are using ATGC as a key. Every bit have 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. From the table 2, we can generate 64 bit key values and adding ATGC, we can generate 72-bit key (64 bits of key combination and 8 bits of ATGC). ATGC key is sending to the receiver side by using Diffie-Hellman key sharing algorithm. In this work, every time the key value will be randomly change.

Table -2: Key combination values

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	15
CA	1110	14
CT	1011	11
CG	0000	0
CC	1101	13

4.7 DNA Steganography:

The Random DNA sequence is selected and those sequence is indexed with the original DNA sequence. So the original DNA sequence is hidden. These new DNA sequence is converted in to binary for further process.

4.8 Morse Coding:

The RDNA sequence is then converted into International Morse code, is a method of transmission of text information between sender and receiver. It is invented by Samuel F.B. Morse in the telegraphy fi el. This code encodes the original text to non-English natural language called “dots and “dashes.

Table -1: Morse code

DNA SEQUENCE	MORSE PATTERN
A	.-
G	..
C	-.
T	–

DECRYPTION PROCESS:

Now at receiver side, the receiver receives the amplified message and ATGC key for decryption and performing encryption operation. This is all contains in the user module. The server module is only for accept connections from client and view cloud storage, machine name and IP address. The cloud module is for control all requests send to the server and also cloud can leave from the network whenever it needs

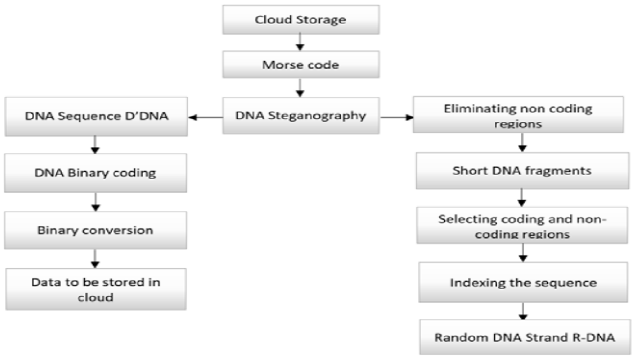

Fig -4: Decryption process

5. CONCLUSIONS

Cloud data security is the main challenge for cloud usability. There are various algorithms to ensure security for cloud data like RSA, Diffie-Hellman, DNA encryption etc. Two level security is currently provided for cloud data security. But in this study more security is given for cloud security. Key combination, DNA Encryption, Random-DNA (RDNA), DNA Steganography, Morse code technique. The purpose of future enhancement is to update the developed system as the need arise and when new technologies comes.

The future enhancement is subjected to the user needs and technological growth.

REFERENCES

[1] L. M. Adleman, "Molecular computation of solutions to combinatorial

[2] E. S. Babu, C. N. Raju, and M. H. K. Prasad, "Inspired pseudo biotic

[3] R. Bhadauria and S. Sanyal, "Survey on security issues in cloud computing and associated mitigation techniques," arXiv preprint arXiv:1204.0764, 2012

[4] K. Brindha and N. Jeyanthi, "Securing portable document format file using extended visual cryptography to protect cloud data storage," International International Journal of Network Security, Vol.20, No.x,PP.xxx-xxx, xxx. 2018 (DOI: 10.6633/IJNS.2018xx.20(x).xx) 8

[5] Z. Cao, C. Mao, L. Liu, "Analysis of one secure anticollusion data sharing scheme for dynamic groups in the cloud," International Journal of Electronics and Information Engineering, vol. 5, no. 2, pp. 68–72, 2016.

[6] C. T. Clelland, V. Risca, and C. Bancroft, "Hiding messages in dna microdots," Nature, vol. 399, no. 6736, pp. 533–534, 1999.

[7] Cold Spring Harbor Laboratory, Exons and Introns, Oct. 23, 2016. (http https://www.dnalc.org/view/15549-transcritiontranslation-exonsand introns.html)

[8] M. Hogan, F. Liu, A. Sokol, and J. Tong, "NIST cloud computing standards roadmap," NIST Special Publication, vol. 35, 2011.

[9] X. Li, C. Zhou, N. Xu, "A secure and efficient image encryption algorithm based on DNA coding and spatiotemporal chaos," International Journal of Network Security, vol. 20, no. 1, pp. 110-120, 2018.

[10] L. Liu, W. Kong, Z. Cao, J. Wang, "Analysis of one certificate less encryption for secure data sharing in public clouds," International Journal of Electronics and Information Engineering, vol. 6, no. 2, pp. 110115, 2017.

[11] P. Mell, T. Grance et al., "The NIST definition of cloud computing,"2011.

[12] M. Misbahuddin and C. Sreeja, "A secure imagebased authentication scheme employing dna crypto and steganography," in Proceedings of the Third International Symposium on Women in Computing and Informatics, pp. 595–601, 2015.

[13] A. Mosa, H. M. El-Bakry, S. M. Abd El-Razek, S. Q. Hasan, "A proposed E-government framework based on cloud service architecture," International Journal of Electronics and Information Engineering, vol. 5, no. 2, pp. 93–104, 2016.291-303, 2016.

[14] D. Patel, "Accountability in cloud computing by means of chain of trust dipen contractor," International Journal of Network Security, vol. 19, no. 2, pp. 251-259, 2017.

[15] B. T. Rao and N. vurukonda, "A study on data storage security issues in cloud computing," Procedia Computer Science, vol. 92, pp. 128–135, 2016.

[16] C. Sreeja, M. Misbahuddin, and N. Mohammed Hashim, "Dna for information security: A survey on dna computing and a pseudo dna method based on central dogma of molecular biology," in International Conference on Computer and Communications Technologies (ICCCT'14), pp. 1–6, 2014.

[17] Y. Sun, J. Zhang, Y. Xiong, and G. Zhu, "Data security and privacy in cloud computing," International Journal of Distributed Sensor Networks, 2014.

[18] W. L. Tai, Y. F. Chang, "Comments on a secure authentication scheme for IoT and cloud servers," International Journal of Network Security, vol. 19, no. 4, pp. 648-651, 2017.

[19] Tutorvista, Deoxyribonucleic Acid, Nov. 23, 2016.(http://chemistry.tutorvista.com/biochemistry/deoxyribonucleicacid.html)

[20] N. Vaanchig, H. Xiong, W. Chen, Z. Qin, "Achieving collaborative cloud data storage by key-escrowfree multi-authority CP-ABE scheme with dualrevocation," International Journal of Network Security, vol. 20, no. 1, pp. 95-109, 2018.

[21] Virtual Genetics Education Centre, DNA, Genes and Chromosomes,Nov. 22, 2016. (http://www2.le.ac.uk/departments/genetics/vgec/highereducation/topics/dnageneschromosomes)

[22] Securing Cloud Data using Dna and Morse Code: A Triple EncryptionScheme(https://www.researchgate.net/publication/330740693-Securing-Cloud-Data-using-DNA-and-Morse-Code-A-Triple-Encryption-Scheme)