A DNA-based Algorithm Methodology for Securing Cloud Data Storage

Animesh Kairi[1][0000-0003-2874-5998] ,Tapas Bhadra[2],Arindam Roy[3] and Debasmita Guha Biswas[4]

[1]Insititute of Engineering and Management, Kolkata, India
[2] Aliah University, Kolkata, India
[3]UEM Kolkata, India
[4]UEM Kolkata, India
ani.kairi@gmail.com
tapas.bhadra@aliah.ac.in
arindammtech@gmail.com
debasmitaguha3@gmail.com

Abstract Cloud computing, with its amazing performance, affordability, accessibility, and plethora of other benefits, has drastically changed the world and opened up new opportunities. It deals with the on-demand provision of computing resources, particularly computational power and data storage (cloud storage), through data centers that are reachable by many clients via the Internet. This technology makes it possible to employ computer resources quickly and affordably. This paper ensures cloud data security, comprehensive safety policies, a strong organizational safety culture, and cutting-edge solutions for cloud security are essential using DNA-based genetic algorithms. Encryption is the main technique among many that protect information exchange in the cloud, guaranteeing the integrity and confidentiality of data or messages transmitted. Genetics-based cryptography has been studied recently, looking at how genetic concepts can improve cryptographic techniques for data processing, storing, and sending.

This work presents a novel approach to protecting data stored in the cloud by incorporating DNA-based cryptography. This methodology technique, called Cloud-Based DNA Cryptography, provides a cutting-edge way to strengthen cloud data security by utilizing DNA's built-in capacity for data storage and strong encryption capabilities. The goal of the study is to investigate the foundations, methods, and possible benefits of using genetic cryptography in cloud systems. Using this investigation, the research aims to augment comprehension and provide significant discernments about the practicability and efficacy of genetic encryption in strengthening the security of cloud-based data storage.

Keywords: DNA, Cryptography, Cloud Computing, Genetics Algorithm, Cryptosystem.

1 Introduction

DNA, the essential building element of life, has amazing features that make it an ideal option for cryptographic applications. Its remarkable data storage capacities, along with dense information encoding and intrinsic cryptographic characteristics, make it an appealing option for safe data storage and encryption. Researchers have set an adventure to investigate DNA's potential as a storage medium and a cryptographic tool, using its unique properties. The combination of DNA's intrinsic characteristics and cryptographic approaches depicts prospects for revolutionizing data security in cloud-based storage systems [1].

Beyond its ability to solve existing cryptographic drawbacks, DNA cryptography opens up new avenues for encryption and safe data storage. Its connection with cloud-based storage systems promises new paradigms for data protection in addition to increased security. Using DNA's special qualities—such as its enormous store capacity and built-in encryption—this creative method provides a window into the data security of the future.

This research explores the intrinsic benefits of DNA cryptography and imagines a time when this multidisciplinary combination of scientific principles and data security is essential to protecting the secrecy and integrity of data kept on cloud servers. By accomplishing this, we add to the expanding body of knowledge on DNA cryptography and open the door to its useful application in real cloud settings.

1.1 Organization of the Paper

This paper is structured as follows: The first section includes the introduction, main objective, and motivation. The second section outlines the literature review. An overview of cloud computing is provided in the third section, followed by the proposed methodology of Cloud-Based DNA Cryptography in the fourth section. The fifth section compiles various security analyses used in cloud computing. The sixth section offers the discussion, and the final section concludes the paper.

2 Literature Review

The literature review encompasses various aspects of cloud computing, data security issues, and challenges associated with cloud-based DNA computing, DNA, and DNA cryptography. This section examines numerous DNA cryptography methods outlined in Table 1.

Table 1. Various Papers and Description for Literature Review

Paper Title	Author(s)	Description	Publishing Year
Towards DNA-based data security in the cloud computing environment [2]	Suyel et al.	Foundational exploration of DNA cryptography principles, encoding techniques, and potential applications. Demonstrates the potential of DNA as a cryptographic tool.	2022
Cloud Computing Security: Challenges and Solutions [3]	Abdukodir et al.	Surveys security challenges in cloud computing and advocates for inventive approaches to upgrade information security in cloud capacity. Lays the establishment for DNA cryptography integration.	2021
Biometrics security system: A review of multimodal biometrics-based techniques for generating crypto-key [4]	Indu et al.	Provides an overview of recent advancements in DNA data storage technologies, emphasizing the potential for compactly storing vast amounts of data using DNA.	2015
A comprehensive review of cloud computing security [5]	Shanmugasundaram et al.	A comprehensive audit tending to different cyber security viewpoints in cloud computing. Examines restrictions of conventional cryptographic strategies and underscores the requirement for novel approaches like DNA cryptography.	2017
Secure Data Storage in The Cloud by Using DNA and Chaos Cryptography [6]	Elminaam et al.	Investigates a DNA cryptography-based approach for secure cloud storage, highlighting its potential to significantly enhance data security and confidentiality in cloud environments.	2022
Privacy-preserving pattern matching over encrypted genetic data	Bing et al.	Investigates how DNA encryption procedures can be utilized to attain privacy-preserving information	2017

in cloud computing [7]		capacity. Examines strategies to guarantee information secrecy and protection inside cloud-based capacity frameworks.	
Data Security and Privacy Using DNA Cryptography and AES Method in Cloud Computing [8]	A. Kumar	Analyzes DNA encryption calculations planned to upgrade cloud information security. Examines cryptographic methods utilizing DNA and their potential applications in secure information capacity inside the cloud.	2021
Advances of DNA Computing in Cryptography [9]	Namasudra et al.	Examines the scalability aspects of DNA cryptography for cloud-based applications. Investigates how DNA's unique properties can be utilized to ensure flexible and secure data storage within the cloud environment.	2018
Secure Data Storage in The Cloud by Using DNA and ChaosCryptography [10]	Elminaam et al.	Examines cryptographic strategies custom-made to guarantee secure DNA information capacity. Look at different encryption strategies and their reasonableness for securing information put away utilizing DNA-based storage solutions.	2022

These research works collectively contribute to advancing the understanding and implementation of DNA cryptography in cloud computing, addressing challenges related to data security, privacy, integrity, and scalability.

2.1 DNA Computing

Biocomputing represents a current trend, leveraging bimolecular concepts' computational complexity, combined with traditional cryptographic methods, to generate secure algorithms for cryptography [12]. The application of bimolecular encryption techniques has spawned a novel cryptography branch known as DNA cryptography, with its principles primarily centered on DNA [11] and the central dogma of molecular biology. DNA, or Deoxyribonucleic acid, constitutes a double-stranded helix composed of nucleotides, where each nucleotide consists of one of four bases: A (adenine), G (guanine), C (cytosine), and T (thymine) [11, 12].

2.2 DNA and Cloud Computing

DNA data storage is an emerging technology that involves encoding digital information into the structure of DNA molecules. DNA offers a highly dense and stable storage medium, making it a potential solution for the growing need for efficient and long-term data storage. Cloud computing, on the other hand, involves the use of a network of remote servers hosted on the Internet to store, manage, and process data.

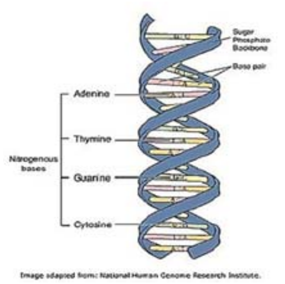

Fig. 1. DNA structure

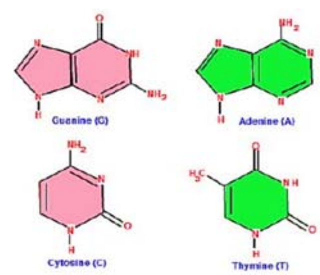

Fig. 2. DNA pairs

Base pair bonds, nitrogenous bases, and the sugar-phosphate backbone are all depicted in Figure 1's representation of the DNA helical helix. The link between adenine and thymine, cytosine and guanine, and the structure of DNA base pairs are depicted in Figure 2. Between the exons, the coding sections of eukaryotic genes, lie non-coding portions called introns.

3 Cloud Computing

Its foundation is the idea that an application framework may be viewed as a cloud, giving users and organizations global access. This method's basic tenet is to provide computing, storage, and software as a service. In essence, a cloud is a distributed, parallel system made up of networked, virtualized computers. Based on their unique needs, customers are charged for cloud storage, resources, and other services like Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). To meet the various demands of customers, this model provides processing power, storage, resources, and applications as a utility.

4 Proposed Methodology Cloud-Based DNA Cryptography

This section explains the methodology proposed for implementing Cloud-Based DNA Cryptography (CBDC) to strengthen data security within cloud storage environments. CBDC integrates DNA-based encryption technology with DNA steganography, encryption, and indexing techniques, offering a practical and innovative approach to data protection. The four nucleic acid bases A (adenine), C (cytosine), G (guanine), and T (thymine) constitute a DNA sequence; A pairs with T and G pairs with C, forming complementary bases. These DNA bases are shown in Table 3.

Table 3. Binary Table with DNA Bases

DNA Base	Binary Value
A	0000
G	0001
C	0010
T	0011

The architectural arrangement of the suggested DNA-based encryption-decryption is depicted in Figure 4.

Fig. 4. Cloud-Based DNA Encryption-Decryption Block Diagram

The DNA-Based Cloud Cryptography Encryption and DNA-Based Cloud Cryptography Decryption algorithms are shown in Algorithm 1 and Algorithm 2, respectively.

Algorithm 1: <u>DNA-Based Cloud Cryptography Encryption</u>

- **Input**: Digital data to be encrypted (plaintext) and Encryption key.

Step 1. Digital Encryption: Encrypt the digital data using a standard cryptographic algorithm with theencryption key to obtain encrypted digital data.

Step 2. DNA Encoding: Map the encrypted digital data into a DNA sequence based on a defined encoding scheme and then define a mapping table that associates digital symbols (e.g., bits or bytes) with specific DNA bases (A, T, C, G).

Step 3. Fragmentation and Concatenation: Fragment DNA the -encoded data into smaller segments to facilitate efficient storage and processing and concatenate metadata to each segment, indicating its position and any other relevant information.

Step 4. Cloud Storage and Replication: Store the fragments of DNA-encoded data in the cloud storage infrastructure

Algorithm 2: <u>DNA-Based Cloud Cryptography Decryption</u>

Input: DNA-encoded data fragments obtained from the cloud. Encrypted encryption key (retrieved from secure storage).

Output:

Step 1. **Retrieve Data:** Retrieve the DNA-encoded data fragments from the cloud storage.

Step 2. Concatenation and Reconstruction: Concatenate the DNA-encoded segments to reconstruct the complete DNA-encoded data.

Step 3. DNA-to-Digital Decoding: Decode the DNA sequence back into encrypted digital data using the reverse of the definedencoding scheme.

Step 4. Digital Decryption: Decrypt the encrypted digital data using the decryption key.

Table 4. DNA Encoding Library

AATC	ACGC	TTTG	ACAG	ACCG
ATCA	GCCT	CTCA	GTAA	ATTG
AGGT	TTAA	GTAT	ACCT	ACGG
CCAG	CCAC	GCCG	GCGC	TCTA

5 Result Analysis

The encryption process output is cipher text. Here the plain text is converted into DNA using DNA encode library, DNA padding table, and bond table and converted into decimals of 0,1,2 and 3 as described earlier in this paper. The cipher text snap is shown below after using the proposed cryptography technique in Figure 5.

```
23012012301232301123001230123230112303012123012302301012323011230123030
0121230123012303012301212301
23012012312302301301223013012012312302301123012301230012301232301301222
301012312301  21230230123011230301230120123301201233012301212301
21230230112301230230101230123012301233012301212303012230130120123012322
30112301230230130123012012312302301012312302301012323011230123001230122
32301123023010123123012303012123012300123230112301230012301230123230111
21230230101231230301223011230230123010123012301231
20123123023010123230101230123012312303012230123011230301212301230012333
```

Figure 5.Cipher Text of Proposed Technique

We take various size of plain text as different DNA cryptography algorithm does and we get the below table 7. From Table 7 and Figure 6, we can find that although the size of plaintext increases but the crypto time is quite less in comparison to other DNA cryptography techniques [14].

Figure 6 Encryption and Decryption Time of Proposed Technique

6 Conclusions

In conclusion, Cloud-Based DNA Cryptography is at the cutting edge of innovation, combining cloud computing, bioinformatics, cryptography, and DNA information capacity to tackle the evolving challenges of secure and efficient data storage and processing. This integration provides a potential solution for securely storing and managing large volumes of data using DNA molecules, leveraging the computational power and scalability of cloud computingIn summary, Cloud-Based DNA Cryptography promises to revolutionize data storage and security by integrating DNA data storage with cloud computing infrastructure. However, further research, experimentation, and technological advancements are needed to fully realize the potential of this innovative approach.

References

1. A. Karimov, I. Olimov, K. Berdiyev, U. Tojiakbarova and O. Tursunov, "Cloud Computing Security Challenges and Solutions," 2021 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2021, pp. 1-6, doi: 10.1109/ICISCT52966.2021.9670220.

2. S. Namasudra, D. Devi, S. Kadry, R. Sundarasekar, A. Shanthini, "Towards DNA based data security in the cloud computing environment,Computer Communications," 2020, pp. 539-547, doi: /10.1016/j.comcom.2019.12.041.

3. I. Verma and S. K. Jain, "Biometrics security system: A review of multimodal biometrics-based techniques for generating crypto-key," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2015, pp. 1189-1192.
4. G. Shanmugasundaram, V. Aswini and G. Suganya, "A comprehensive review on cloud computing security," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-5, doi: 10.1109/ICIIECS.2017.8275972.

5. D. S. A. Elminaam, M. A. W. Mousa and M. A. El Fattah, "Secure Data Storage in The Cloud by Using DNA and Chaos Cryptography," 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2022, pp. 175-182, doi: 10.1109/MIUCC55081.2022.9781704.
6. B. Wang, W. Song, W. Lou and Y. T. Hou, "Privacy-preserving pattern matching over encrypted genetic data in cloud computing," IEEEINFOCOM 2017 - IEEE Conference on Computer Communications, Atlanta, GA, USA, 2017, pp. 1-9, doi: 10.1109/INFOCOM.2017.8057178.

7. G. Hamed, M. Marey, S. A. El-Sayed and M. F. Tolba, "Comparative study for various DNA based steganography techniques with the essential conclusions about the future research," 2016 11th International Conference on Computer Engineering & Systems (ICCES), Cairo, Egypt, 2016, pp. 220-225, doi: 10.1109/ICCES.2016.7822003.

8. A. Kumar, "Data Security and Privacy using DNA Cryptography and AES Method in Cloud Computing," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1529-1535, doi: 10.1109/I-SMAC52330.2021.9640708.

9. Namasudra, S., & Deka, G.C. (Eds.). (2018). Advances of DNA Computing in Cryptography (1st ed.). Chapman and Hall/CRC.https://doi.org/10.1201/9781351011419

10. D. S. A. Elminaam, M. A. W. Mousa and M. A. El Fattah, "Secure Data Storage in The Cloud by Using DNA and Chaos Cryptography," 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2022, pp. 175-182, doi: 10.1109/MIUCC55081.2022.9781704.

11. A. Kairi, T. Bhadra, “Decoding The Future Using A Novel DNA-Based Cryptosystem”, in Journal of European Chemical Bulletin, Volume - 12, Special Issue-10: Page: 3597 –3609 2023

12. A. Kairi, T. Bhadra, A. Roy, “SECURED HEALTHCARE ECOSYSTEM USING DNA CRYPTOGRAPHY IN 5G NETWORK” Volume 23, Issue 2, September 2023 in China Petroleum Processing and Petrochemical Technology

13. A. Kairi, T. Bhadra,T Saha,S Saha, “Enhancing Healthcare Image Security with DNA Cryptography in the IOT Environment” Volume 12, Issue 10, Pages 1975–1986.

14. Kairi, A., Bhadra, T., Roy, A., Saha, T. (2024). An Innovative Method for DNA Cryptography-Based Digital Signature Verification. In: Dutta, S., Bhattacharya, A., Shahnaz, C., Chakrabarti, S. (eds) Cyber Intelligence and Information Retrieval. CIIR 2023. Lecture Notes in Networks and Systems, vol 1025. Springer, Singapore. https://doi.org/10.1007/978-981-97-3594-5_1.