EAI Endorsed Transactions
on Scalable Information Systems

Research Article EAI.EU

A Realizable Data Encryption Strategy

Pranjali[1], Srividya Ramisetty[2*], Vani B Telagade[3] and S Disha Adiga[4]

[1,2,3,4]CMR Institute of Technology, Bengaluru, India

Abstract

As technology continues to advance, data has become an increasingly important element in the sphere of Information Technology. However, enormous data generated by devices presents a major challenge in handling it in real time. Data encryption is a crucial component in ensuring data security and privacy during its transmission in network. Unfortunately, many applications disregard data encryption in order to achieve higher performance. The work proposes a solution to this problem by introducing a data encryption process that is, the Realizable Data Encryption Strategy (RDES) and Deoxyribonucleic Acid (DNA) computing, a revolutionary cryptographic method that improves information security by preventing authorized access to sensitive data, being used. Information security is improved by DNA symmetric cryptography being suggested. The outcomes show that plain-text encryption is a very secure procedure. The RDES approach is designed to improve privacy protection within the constraints of real-time processing. By implementing the RDES approach, data privacy and security can be significantly enhanced without compromising performance.

Keywords: Cryptography, DNA, RDES, Encryption, Decryption.

Received on 05 December 2023, accepted on 22 February 2024, published on 28 February 2024

doi: 10.4108/eetsis.5230

*Corresponding author. Email: srividya.ramisetty@gmail.com

1. Introduction

The Internet has connected the world like never before, and the future of technology is mobile computing. It allows access to the Internet and data on-the-go. Mobile cloud computing methods have enabled divergent applications, but concerns remain about data security and privacy [1]. The key privacy concern is due large volumes of data that needs to be transmitted which is not enciphered or encrypted. It may result in issues of privacy leakage. To address this problem, the paper proposes use of a Realizable Data Encryption Strategy (RDES) model. It is designed to protect the privacy of owners of data using networking facilities and devices. The RDES model includes two important terms for implementing data encryption: Pairs Matching and Paired Data Collision. It also includes two algorithms, S Table Generation (STG) and Weight Modularization (WM), to support the implementation of the RDES algorithm. When deciding on data encryption these algorithms identify means of recognizing privacy values.

Key contributions of the work carried are:

Firstly, the approach proposed selectively encrypts data maximizing protection level. Two modes of operation, encryption and non-encryption modes, are considered during the transmission phase.
Secondly, the algorithm proposed is expected to offer an optimal solution with utmost value for the total privacy weights. The constraints that may be involved are processing time and levels of privacy.
Thirdly, outcomes of the work provide solutions that focus on protecting data privacy and with an adaptive approach for data transmission.

Many scientists in numerous domains have used DNA computing to offer protection for the enormous quantity of data depicted by performing DNA computing. The term "DNA-BASED cryptographic approaches" refers to a variety of cryptographic techniques. DNA computing also acquired certain desirable qualities in its work, like high parallelism, large storage capacity and maximum energy efficiency, which is highly advantageous for the idea of

data concealing. The programme may offer a remedy for clustering, forecasting, and optimization issues in addition to signal processing. 108 TB of data might fit into one gram of DNA. The equivalent DNA base number is 1021.

The research work carried on aims at providing the following:

A comparative analysis of several DNA cryptography methods.
Provide a strong ciphertext using improved cryptographic technique.
Provide cutting-edge data concealing techniques with the improved strategy.

1.1 Deoxyribonucleic Acid

The natural mega particle known as Deoxyribonucleic Acid (DNA), is the source of living creatures. It is created when nucleotides combine. Deoxyribonucleotides are supposedly thought of as a monomer unit of DNA [2]. DNA is formed of four kinds of nucleotides bases A - Adenine, G- Guanine,T- Thymine and- Cytosine. Purines and Pyrimidines are two nitrogen bases of DNA. Single ring molecules, having bases C and T, are referred to as pyrimidines, whereas double ring molecules, A and G, are assumed as purines.

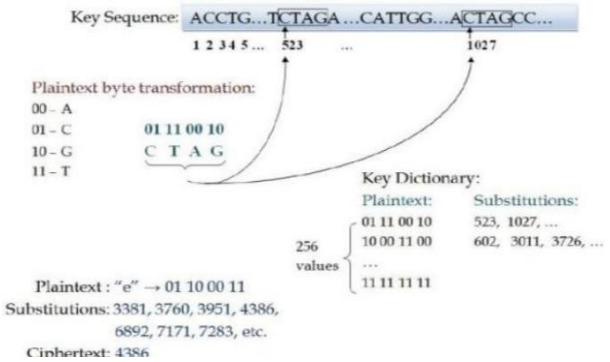

Figure 1. Encryption in DNA Indexing Algorithm

Vast industries that might be used for data and picture encryption, data concealment, steganography, etc. may be identified using DNA's capabilities. While performing DNA operations in various laboratories, consideration should be given to a number of factors, including temperature, pressure, oxygen ratio and so on.
Since DNA is a sensitive module, it might have a variety of outputs depending on the surroundings. Consequently, it may be claimed that the environmental factors depicted in Figure 1, completely determine how DNA computations are performed. The term 'binary coding scheme' refers to, binary conversion of DNA sequence. In this study, binary values 00, 10, 11 and 01 were represented by the DNA bases A, G, T and C. For instance, binary sequence for the DNA sequence "AATCGGAT" is 0000110110100011. The major bases of RNA (ribonucleic acid) are the same as those in DNA, with the exception that Uracil (U) is used in place of thymine (T).
Summarizing, the research work done proposes a naive and an effective method for safeguarding privacy for mobile cloud computing. By implementing DNA cloud computing, data privacy and security can be significantly enhanced without compromising performance.

2. Literature Review

Many applications disregard data encryption to achieve compliance. This work expresses privacy concerns about data and proposes Realizable Data Encryption Strategy (RDES), a revolutionary data encryption method. The goal of this addition is to broaden the breadth of privacy protection while minimizing time limitations.
Numerous cutting-edge applications in Cyber-Physical Systems are being driven by the anticipated enhanced network study and rising need for sharing data in mobile networks. However, the tensions between security and communication effectiveness limit the use of ITS today. The Security Aware Efficient Data Sharing and Transferring method, is proposed in this article as a solution to secure data transmission problem. The experimental assessment has demonstrated that the suggested paradigm performs well at safeguarding communications for ITS [3].
In the data from fusion reactors, a brand-new technique and its application for real-time blob filament detection and tracking are provided. Numerous additional uses in a diagnostic picture, depend on similar temporal properties. The method for obtaining these characteristics is presented in the study [4]. DDoS attack source traceback is an unsolved and difficult issue. A straightforward and efficient traceback approach is deterministic packet marking (DPM), however existing DPM-based traceback schemes are impractical because of their scaling limitations. They proposed, instead of designating every Internet node, as the current techniques do, simply label specific involved nodes for trace back purposes. The paper suggests an on-demand trace back approach rooted on the DPM mechanism. The conventional DPM method is used to designate these implicated ingress routers to identify the involved attack source [5].
The idea of DNA computing led to the evolution of several algorithms. In the past 20 years, the DNA computing concept has been used to tackle challenging problems like SAT and DES. The DNA computing concept is used by Lipton [6] to solve the NP-complete SAT problem. He also developed a DNA-based encoding system, and the author used the fewest possible variables to answer the SAT issue. Data Encryption Standard (DES) cracking was proposed by Boneh et al [7] using DNA computing. With their suggested approach, every cryptosystem might be broken with a key size of only 64 bits or less. Additionally, its encryption circuit is compact and can be broken in 916 steps. Each phase is equivalent to 32 extractions, and processing takes a minimum of one complete day for each

of the 10 extractor stages. To break DES using their suggested technique would take at least four months. The DNA computing approaches developed by Chen et al. [8] might be used to address challenging problems. Using DNA technology, they also cracked DES. The predicted algorithm by Chen et al. has three different functions: 1) The key space should be initialized with every possible using the initial function. 2) The encryption procedure; and 3) Finding the appropriate related key. The molecular sticker algorithm also includes tubes and short memory strands.

MingXin et al [9] proposed DNA cryptosystem, which incorporates DNA biotechnology with cryptography techniques. This technique was developed by utilizing DNA probes. Ciphertext was then included into the DNA chip. The development of DNA chips results in the method's privacy. A suggested approach for the effective, dependable, and secure transfer of data was recently made by Kar et al [10]. The method makes use of DNA keys, which comprises three keys: two for encryption and one for communicating the encryption key to the recipient. In this approach, the text itself will be transformed to binary sequence through a process called string to binary transformation.

The massive DNA pattern is converted to binary data making use of a binary coding algorithm for the encryption key. XOR is performed and adding procedures to the 64-bit raw and 64-bit key would provide the ciphertext. Shiu et al has made several cryptographic technique suggestions [11]. The following are the cryptographic methods that employ the DNA method: a) Insertion technique Methods two and three are replacement and complementing pair. They also demonstrated that the replacement strategy is more effective than the other two, by encrypting ciphertext with the DNA sequence, Liu et al [12] devised a data concealing technique. Their approach uses DNA coding and turns the plaintext into DNA sequence. Chebyshev map constructs one time key making use of the desired DNA sequence key and two pseudorandom DNA sequences, XXOR and YPrimer. The DNA messaging sequence might be turned into ciphertext by using shift procedures on the result that had been generated earlier. Following encryption, the word file might get transformed to PDF and created PDF delivered to the recipient.

Mandge et al [13] had suggested a powerful ciphertext method that combined key generation technology with matrix insertion modification in the encryption process. The several shifting and XOR operations will first turn the plaintext into a mini cipher. It is claimed that because this procedure incorporates safe key creation, if it were applied to the plaintext each time, we might acquire different mini ciphers for the same plaintext. After using a number of biotechnologies, including DNA primers, to transform this mini cipher into a final ciphertext. Table lookup substitution, devised by Taur et al [14] has improved the substitution approach using a single bit complementary rule. They expanded to include the replacement method's two bits complementing rule. TLSM, which makes use of lookup tables [15][16], converts two-bit plain text to the matching text.

Only one bit a time could be translated by a rule developed by Shiu et al. Additionally, the TLSM technique uses highly good ciphertext compression [17]. Zhang et al [18] presented a technique called SCLPV to fend off nefarious auditors. This method simultaneously provided resilience to malevolent audits and certificateless open validation to check outsourcing data accuracy in CPSS.

Wang et al [19] concentrated on creating a method enabling a secure system for the cloud that facilitated public audits while protecting privacy. The approach of defining adversaries was investigated in their study. User-machine interface problems were also covered in a separate investigation, albeit from an entirely distinct perspective, developed an investigative strategy that concentrated on the weaknesses brought on by the abuse of Graphical User Interface (GUI) elements. In the setting of Ui-based applications, this strategy took abuse of GUI element properties into account. There are numerous Encryption methods supported by various other biometric traits, other than DNA [20]. Srividya et al proposed encrypting data using fingerprint biometric as key.

Proposed Method

The absence of ability to monitor in internet browsers might cause privacy problems because the latest platforms do not allow for surveillance of enemies. The pace at which threats are amplified can also be decreased with an effective safe networking infrastructure. The proposed method works in two phases. The phases involve algorithms for data encryption and decryption.

Phase 1: Secret Data Encryption Algorithm

In step first binary data is converted to sequences representing DNA, as in equation (1), equation (2), equation (3) and equation (4)

A=00 . (1)
T=01 . (2)
C=10 . (3)
G=11 . (4)

In step second, Complementary pair rule is implemented. It is an unique pair assigned to every pair of nucleotides base.

Customer chooses to send original information M to cloud computing settings across the internet. To offer and upload the ultimate version of M(M''') to internet, there are 3 channel-phases. The data M is translated into binary form after it is read as an integer. The rule of base pairing must be applied in translating binary data to amino acids, represented as the sequence of DNA. In the actual world of biology, nucleotide synthesis follows set guidelines.

Algorithm 1: Encryption Process

Input: File
Output: Reference Sequence in Hexadecimal format
Start :
Let M be the Data
Convert Binary to DNA Nucleotides
Apply Complementary Rules on M'
Find Index for each Nucleotides Couple in DNA References Sequence M"
Reference Sequence in Hexadecimal format is generated, and this is called Secret Data M'"
Stop

Phase 2: Original data Extraction

Client 2 uses certain integers to represent the secret data. Phase two and its associated subphases will retrieve actual or native data from DNA reference sequences.

Algorithm 2: Decryption Process

Input: Reference Sequence in Hexadecimal format generated in the Encryption Process M'"
Output: Original Data
Start:
Find Index of each Couple of Nucleotides in Reference Sequence of DNA
Apply Complementary Rules on M'
Convert Binary to DNA Nucleotides
Stop

While a user uploads a file from his local computer to a website, the file must first be forwarded to DNA Encryption Service. This runs on a cloud server-1 and encrypts the file. Computer 1 transmits the ciphered file to cloud server-2, through the usage of internet service notion, where it is stored.
In decryption algorithm numeric data is Converted to sequences of DNA. Couple nucleotides are extracted from reference sequence of DNA. Rule of Complementary pair is again used, and unique value is assigned to base pair of every nucleotide. Finally sequences of DNA are converted to binary data. When the user downloads a file from web server, respective file in cloud server-2 is fetched and sent to DNA decryption, executing on cloud server-1. The file is downloaded to user's device once it is decrypted.

4. Results and Inference

Present section throws light on outcomes of research work carried out. The results are explained through figures. Website developed enable users to give the input in form of a file.

Figure 2. User Login

User Login web page is as shown in Figure 2. Once the user login, the file can be uploaded for encryption as depicted in Figure 3. The contents of the file is read and if any sensitive information is detected the file will be encrypted.

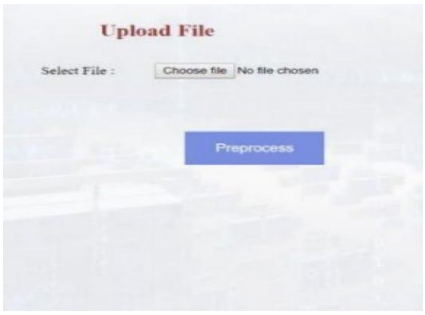

Figure 3. File upload page

The uploaded file is split into blocks irrespective of whether it is sensitive or not and is uploaded to the FTP cloud server.

```
Value -->99
Binary -->1100011
Binary Full -->01100011
Byte: 01100011
Output Phase 1 :TCAG
Output Phase 2:AGCT
First 2:AG
Second 2:CT
test a>>>>>>>>>>> :AG
test n >>>>>>>>>>>:3
Output 1 :3
test a>>>>>>>>>>> :CT
test n >>>>>>>>>>>:8
Output 2 :8
test a>>>>>>>>>>> :AG
test n >>>>>>>>>>>:3
test a>>>>>>>>>>> :CT
test n >>>>>>>>>>>:8
After Encryption -->0308
Value -->114
Binary -->1110010
```

Figure 4. DNA Encryption

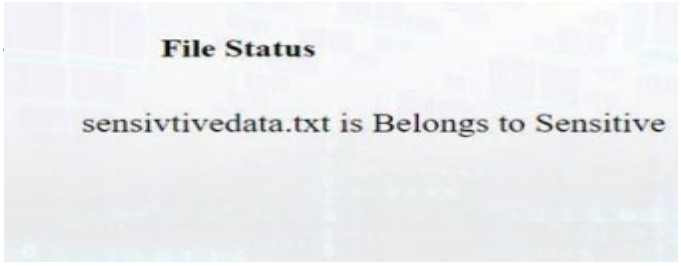

Figure 5. Display of File Type

Figure 4 depicts the results of DNA Encryption process, phase 1 and phase 2 output. If the uploaded file has sensitive data, the file is split, and only sensitive data is encrypted and later stored in cloud. Figure 5 shows the message being displayed as to the file is sensitive or not.

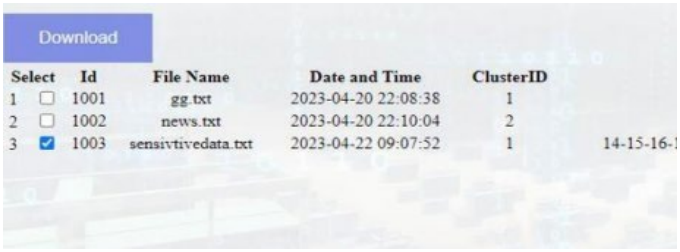

Figure 6. Download options

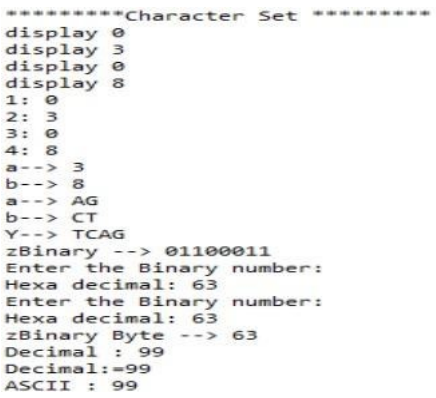

Figure 7. Files stored in cloud

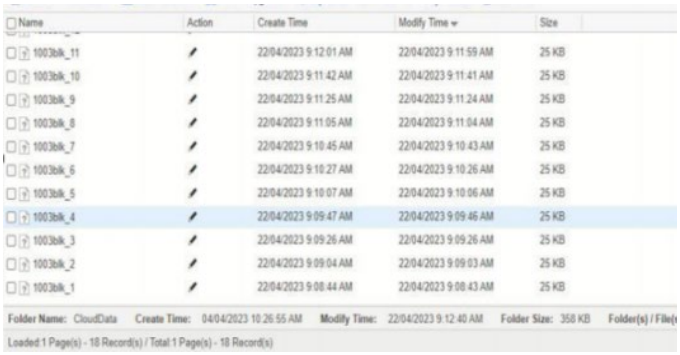

Figure 8. DNA Decryption

Figure 6 depicts the user's download options to download the required document and decrypt it. Figure 7, depicts the file stored as blocks in DriveHQ cloud and Figure 8, depicts the DNA Decryption process. DNA cryptography performed is better in many testing conditions. The results of our study satisfied our planning criteria and agreed with the predicted outcomes in theory.The dimensions used in practical assessments will be the subject of a subsequent study.

5. Conclusions

Big data concerns regarding privacy were the sole subject of this work. It also took into account actual cloud computing deployments. RDES, the suggested solution, was created to boost the effectiveness of privacy measures. The Secret Data Encryption algorithm, created to provide alternate data packages for encryption under varied scheduling limitations, was the main algorithm underlying the RDES paradigm. The results of the practical assessments demonstrated the suggested strategy's better and flexible efficacy.

References

[1] Sumit V T, Ritukar, Murthy B, Jagadish K S. Privacy-Preserving Data Encryption Strategy for Big Data in Mobile Cloud Computing Environment. International Journal of Innovative Research in Science, Engineering and Technology,2018. Vol. 7.

[2] Vikram, A., Kalaivani, S., & Gopinath. G. A Novel Encryption Algorithm based on DNA Cryptography. Proceedings of International Conference on Communication and Electronics Systems (ICCES). 17-19 July 2019, Coimbatore, India.IEEE, 20 February 2020.

[3] Shantha R, Mahender K, Jenifer A: Security analysis of hybrid one time password generation algorithm for IoT data. AIP Conference Proceedings. 2022; 2418:1-10.

[4] Balasubramaniam S, Joe V, Sivakumar TA: Optimization Enabled Deep Learning-Based DDoS Attack Detection in Cloud Computing. International Journal of Intelligent Systems. 2023; 2023:1-14.

[5] Guo. S, Guo. M, Zhou. W, Yu. A workable architecture for IP traceback using dynamically stochastic session tagging. IEEE Transactions on Computers, 2016. pp. 1418–1427.

[6] Lipton R J. DNA solution of hard computational problems. Science, New Series, American Association for Advancement of Science. 268(5210), 1995, pp. 542-545.

[7] Boneh. D, Dunworth. C, Lipton R J. Breaking DES Using a Molecular Computer. Proceedings of DIMACS workshop on DNA computing, Vol 27, pp. 37-51, published by the AMS, 1995.

[8] Chen Z, Geng X, Xu. J. Efficient DNA Sticker Algorithms for DES. Proceedings of IEEE 3rd International Conference on Bio-Inspired Computing, 28 September 2008 - 01 October 2008, Adelaide, SA, Australia, IEEE, 24 October 2008.

[9] MingXin L, XueJia L, GuoZhen X, Lei Q. Symmetric-key cryptosystem with DNA technology. Science in China Series F: Information Sciences, 2007. Vol. 50(3), pp.324-333.

[10] Kar N, Majumder A, Saha A, Jamatia A. An Improved Data Security using DNA Sequencing. MobileHealth '13: Proceedings of the 3rd ACM MobiHoc workshop on Pervasive wireless healthcare, 29 July 2013, Bengaluru, India. Association for Computing Machinery New York, NY, United States, pp.13-18.
[11] Shiu H J, Ng K L, Fang J F, Lee R C T. Data hiding methods based upon DNA sequences. Information Sciences, 2010, Vol.180.
[12] Liu H, Lin D, Kadir A. A novel data hiding method based on deoxyribonucleic acid coding. Computers and Electrical Engineering, 2013, Vol.39.
[13] Mandge T, Choudhary V. A DNA Encryption Technique Based on Matrix Manipulation and Secure key Generation Scheme. Proceedings of IEEE International conference on Information Communication and Embedded Systems,21-22 February 2013, Chennai, India IEEE, 25 April 2013 pp. 47-52.
[14] Taur J S, Lin H Y, Lee H L, Tao C W. Data hiding in DNA sequences based on Table Lookup Substitution. International Journal of Innovative Computing, Information and Control. 2012, Vol.8(10).
[15] Huang Y H, Chang C C, Wu C Y. A DNA-based data hiding technique with low modification rates. Multimedia Tools Applications, Springer, Volume 70, pages 1439–1451, (2014)
[16] Abbasy M R,Nikfard P, Ordi A, Torkaman M R N. DNA Base Data Hiding Algorithm. International Journal on New Computer Architectures and Their Applications (IJNCAA), January 2012, The Society of Digital Information and Wireless Communications, pp-183-193, Vol.2(1).
[17] Tornea O, Antonini M, Borda M. Multimedia Data Compression and Encryption using DNA Cipher. Communications Department, HAL Open Science. Technical University of ClujNapoca Roumanie 2013.
[18] Feng Z, Li X, Wei G. SCLPV: Secure Certificateless Public Verification for Cloud-Based Cyber-Physical-Social Systems Against Malicious Auditors. IEEE Transactions on Computational Social Systems. 2018. Vol 5(3), pp.1-4.
[19] Cong W, Sherman S, Qian W, KuiRen, Wenjing L. Privacy-Preserving Public Auditing for Secure Cloud Storage. IEEE TRANSACTIONS ON COMPUTERS, 2013. Vol.62.
[20] Srividya R and Ramesh B. Implementation of AES using biometric. International Journal of Electrical and Computer Engineering (IJECE), 2019, Vol.9.