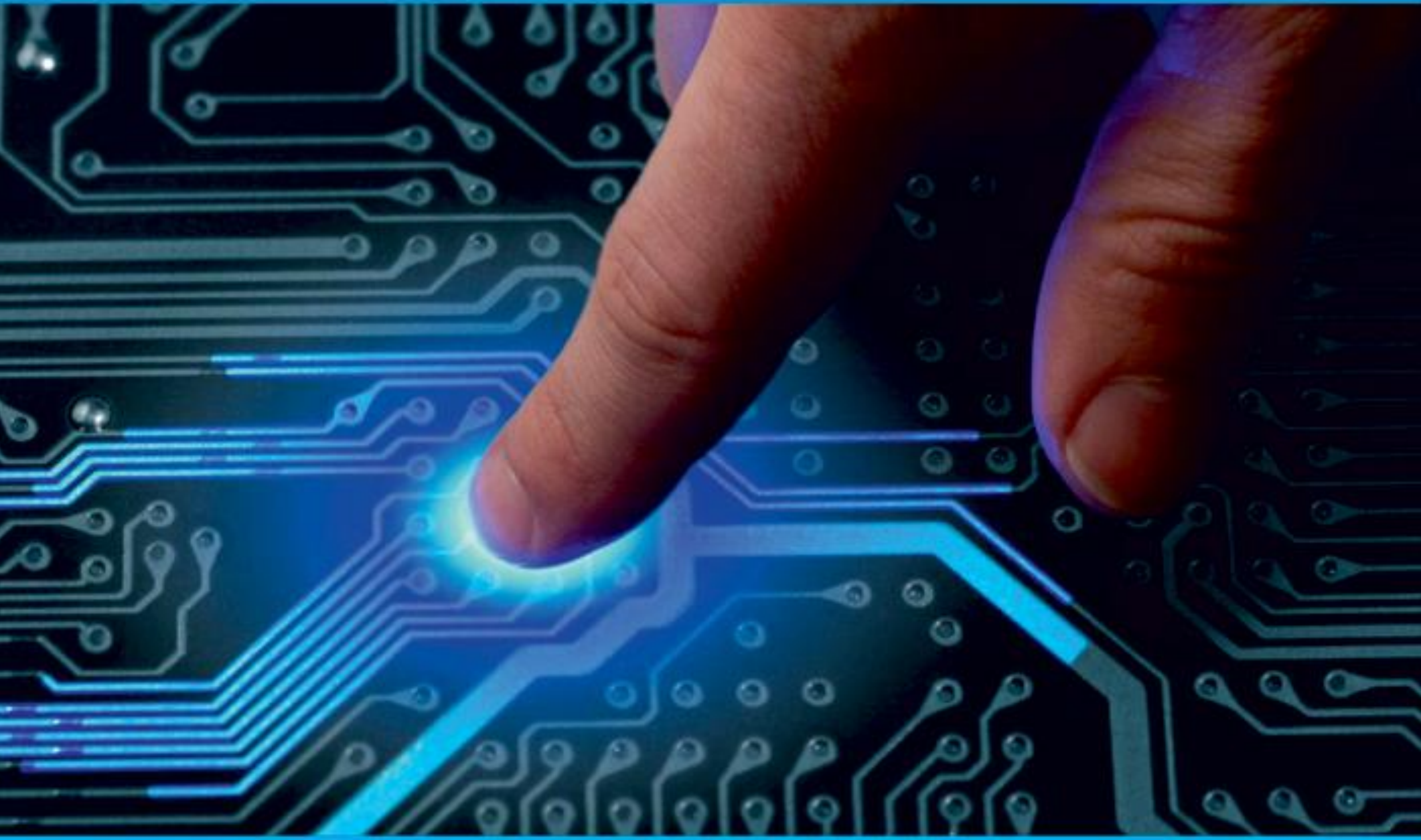
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205351 |

A Novel Hybrid Framework for Securing Health Insurance Data

Rohit Hardikar , Anish More , Kartik Bade , Nishant Joshi , Mrs. J.M. Kanase

UG Student, Dept. of Computer Engineering, P.E.S Modern College of Engineering, Pune, India

UG Student, Dept. of Computer Engineering, P.E.S Modern College of Engineering, Pune, India

UG Student, Dept. of Computer Engineering, P.E.S Modern College of Engineering, Pune, India

UG Student, Dept. of Computer Engineering, P.E.S Modern College of Engineering, Pune, India

Assistant Professor (M.E) , Dept. of Computer Engineering, P.E.S Modern College of Engineering, Pune, India

ABSTRACT: The healthcare industry has witnessed a significant digital transformation with the adoption of electronic health records and the management of health insurance data. As health information becomes increasingly digitized, the need for robust security measures to protect sensitive health insurance data has never been more critical. This project presents a novel Hybrid Framework designed to enhance the security of health insurance data, ensuring confidentiality, integrity, and availability while mitigating potential risks and vulnerabilities. Key components of the Hybrid Framework include a secure data storage system, a real-time monitoring and alerting system, and a user authentication and access control mechanism. In this project, we are going to take the insurance data from the user and hide it in a DNA sequence using dynamic DNA encoding. Then the data will be stored in a cloud database. This project aims to address the growing concerns regarding the security and privacy of health insurance data, particularly in an era of increasing cyber threats and data breaches. By adopting this innovative Hybrid Framework, healthcare organisations can safeguard sensitive information, maintain compliance with data protection regulations, and build trust among patients and stakeholders. Ultimately, the framework contributes to the overall improvement of data security in the healthcare industry, ensuring the confidentiality and integrity of health insurance data in an increasingly interconnected digital world.

KEYWORDS: Dynamic DNA Encoding, Compression, Encryption, Steganography, AES, Cyber-Security.

I. INTRODUCTION

In recent times, the healthcare domain has undergone a significant digital transformation, propelled by the adoption of electronic health records (EHRs) and the management of health insurance data. While this shift promises enhanced operational efficiency and accessibility to healthcare services, it brings forth unprecedented challenges, particularly concerning the security and confidentiality of data. The realm of health insurance data, encapsulating intricate details of individuals' medical histories, treatments, and coverage, has emerged as a prime target for cyber threats. Breaches in health insurance data pose multifaceted risks, spanning from financial exploitation to breaches of patient privacy and potential identity theft. Consequently, there is an urgent imperative for robust security protocols to safeguard this critical information. In response to these challenges, this paper presents a novel hybrid framework designed explicitly for securing health insurance data. Our framework combines the strengths of multiple security technologies, including encryption, authentication, to provide a comprehensive and resilient defense against evolving cyber threats. Sensitive and extremely personal information on a person's medical history, current medical conditions, and financial situation is contained in health insurance data. The healthcare industry is now a top target for cyberattacks due to its worth on black markets. Patient privacy and data integrity are seriously at risk from these threats, which have increased in frequency and complexity. They include ransomware attacks, identity theft, and data breaches. There has been a discernible increase in reported data breaches over the last 14 years, with 2021 seeing a particularly high number. For healthcare organizations, building and maintaining trust with stakeholders and patients is still crucial. Data breaches can damage a person's reputation over time in addition to undermining trust. enabling sensitive data interchange between healthcare organizations in a secure manner.

In our proposed technique we are using two encoding methods namely huffman encoding and LZ77 encoding. One technique for lossless data compression is Huffman encoding. It is determined by how frequently each data item appears in a file. Shorter codes are assigned to more common symbols and longer codes to less frequent symbols in Huffman encoding. In this sense, it minimizes the average number of bits needed to represent each symbol, improving data representation. LZ77 (Lempel-Ziv 1977) is a dictionary-based compression algorithm. It works by finding repeated sequences of data in the input stream and replacing them with references to previous occurrences. This approach exploits redundancy in the data to achieve compression. The security of health insurance data is achievable through the technique of dynamic DNA encoding along with data compression. By dynamically modifying the encoding technique in accordance with the input data, dynamic DNA encoding seeks to maximize the utilization of DNA sequences. This reduces the amount of nucleotides needed to encode the information while enabling effective data storage and retrieval. The contribution of the proposed work is described as follows.

- To provide a secure means of communication between users, hospitals and insurance companies.
- To hide the data using dynamic DNA encoding.
- To lower insurance security data breaches.
- To securely retrieve the data and display it to the authorized person

II. RELATED WORK

Various methods are introduced in the existing medical insurance data security. Some of the recent works are discussed in this section. Swetha Gadde, J. Amutharaj, S. Usha introduced a security model to protect the isolation of medical data in the cloud using hybrid cryptography. To protect the integrity and confidentiality of medical data, hybrid cryptography is used, which combines symmetric and asymmetric cryptosystems. An improved iteration of the Advanced Encryption Standard (AES) that utilizes a robust S-box is suggested, incorporating the Runge-Kutta Optimization (RKO) method. RKO improves security by using calculations based on the Mackey-Glass equation to produce an enhanced RS-box. Initially, Improved Huffman Coding (IHC) is used to compress medical data. The Modified Elliptic Curve Cryptography (MECC) technique, which generates keys based on Deoxyribonucleic Acid (DNA), is presented. The Bald Eagle Search optimization algorithm (BES) is used to help find the optimal key. The medical data are then safely saved in the cloud and encrypted using the IRS-AES algorithm.

Hongmin Li, Tie Li, Wei Feng, Jing Zhang, Jun Zhang, Lixia Gan, and Chunlai Li introduced a novel image encryption scheme based on non-adjacent parallelable permutation and dynamic DNA-level two-way diffusion. This study presents IES-NPDDT, a novel approach to picture encryption that combines dynamic DNA-level two-way diffusion with non-adjacent parallel permutation. The main goal of the technique is to increase the efficiency of encryption by making each encryption process parallelisable. The plain image is first split into permutation blocks and neighbouring blocks. Multiple computing units then permute these blocks simultaneously. The permuted image is then encoded into four DNA planes, each of which has two bits that are dynamically encoded using chaotic sequences. The substantial nonlinearity of the diffusion process is ensured by this dynamic encoding in conjunction with dynamic DNA decoding. Finally, using the hash value, key components, and chaotic values as guides, multiple two-way DNA-level diffusions are carried out in parallel on the DNA planes. Gunjankumar Bhoi, Raj Bhavsar, Priteshkumar Prajapati and Dr. Parth Shah presented a Review of Recent Trends in DNA Based Cryptography. The concept of DNA cryptography originated because, according to studies, even a small amount of DNA can store enormous amounts of information, analyze that information, and communicate it. An overview of DNA-based security research is provided in this paper. Encryption algorithms based on coupled map lattices, random DNA, polymerase chain reaction, and other popular encryption algorithms are discussed. Algorithms can be used for specialized or broad purposes; some are just meant to encrypt certain types of images, such as text data or photos from the medical field, while others are meant to be used generally for both text and image data.

III. PROPOSED METHODOLOGY

A. *Dynamic DNA Encoding:*

DNA is made up of four nucleotides: adenine (A), cytosine(C), guanine (G), and thymine (T). Data will be converted to DNA sequence by converting its binary form to A, T, C and G. For example “dna” to “01100100 01101110 01100001”.Binary will be converted to DNA sequence using these values: A-00, T-01, C-10, G-11. To make this process dynamic, we use four different values to encode the data:

Table1 (A-00, T-01, C-10, G-11)
Table2 (A-01, T-10, C-11, G-00)
Table3 (A-10, T-11, C-00, G-01)
Table4 (A-11, T-00, C-01, G-10)

For example: the string ”dna” will be converted to say
”GCGCTTCGAAAT CAGTCGAGCATTCTC TAGTGTTCTCAGGATGGTCTTGGCTCAAAGCACAGCTTTA
CCATGGAGTCGATTTAGAGCGAGGTGAGGACCTTGTGGCCCGGTGTACTACGGCAGTCTCTTGCT”
and the key required to decipher it is ”33434342241341414213414144312243143141232423134132
3141223221221211124411322142111411243311213233213313433324324311312323412133214212”.The key is the sequence of the name of the table used to encode its respective character. In this way, the string ”dna” is converted to a sequence of DNA nucleotides.

B. *Compression:*

The suggested solution uses LZ77 and Huffman Coding as its compression techniques. The two most often used letters in English are typically "e" and "t." What is the most typical pair, then? You might have guessed "ee," "et," or "te," but it's actually "th." These kinds of popular words and syllables, which occur significantly more frequently than you may assume based just on letter frequencies, are easy to find and compress with LZ77. While letter-oriented Huffman may effectively identify and compress files based only on letter frequencies, it is unable to identify correlations between consecutive letters, which are found in common words and syllables.

An original file is compressed by LZ77 into an intermediate string of literal letters and "copy items." After that, Huffman compresses the intermediate sequence much more. Those "copy items" are frequently already substantially shorter than the original substring would have been if we had only compressed the original file using Huffman rather than the LZ77 phase. Furthermore, Huffman compresses the actual letters in the intermediate sequence exactly as effectively as it would have in the original file.

C. *Encryption:*

Sensitive data is encrypted by the encryption module using the Advanced Encryption Standard (AES) 256 method to safeguard it against illegal access or interception. It guarantees the integrity and confidentiality of data while it is being sent and stored. AES 256 encryption: Makes use of the widely recognized AES 256-bit encryption method norm for safe encryption. Before storing, sensitive data is encrypted. or transmission to stop data breaches or illegal access. securely maintains encryption keys to guarantee the integrity and confidentiality of data that is encrypted. gives approved individuals the ability to decode to obtain encrypted data as required.

In our proposed methodology, one of the project modules is Authentication. The Authentication module is responsible for verifying the identity of users accessing the application. It includes functionalities such as user registration, login, password management, and session management. Allows users to create new accounts by providing necessary information and verifying their identity. Authenticates users based on their credentials to grant access to the application. Enables users to reset passwords, change passwords, and implement password policies for security Another module used is API Integration. The API Integration module facilitates integration with external systems, services, or APIs to extend the functionality of the application and enable seamless data exchange. Defines endpoints to communicate with external APIs or services for data retrieval, updates, or synchronization. Implements mechanisms to authenticate and authorize access to external APIs securely. Maps data between different formats and structures to ensure compatibility and consistency between systems.

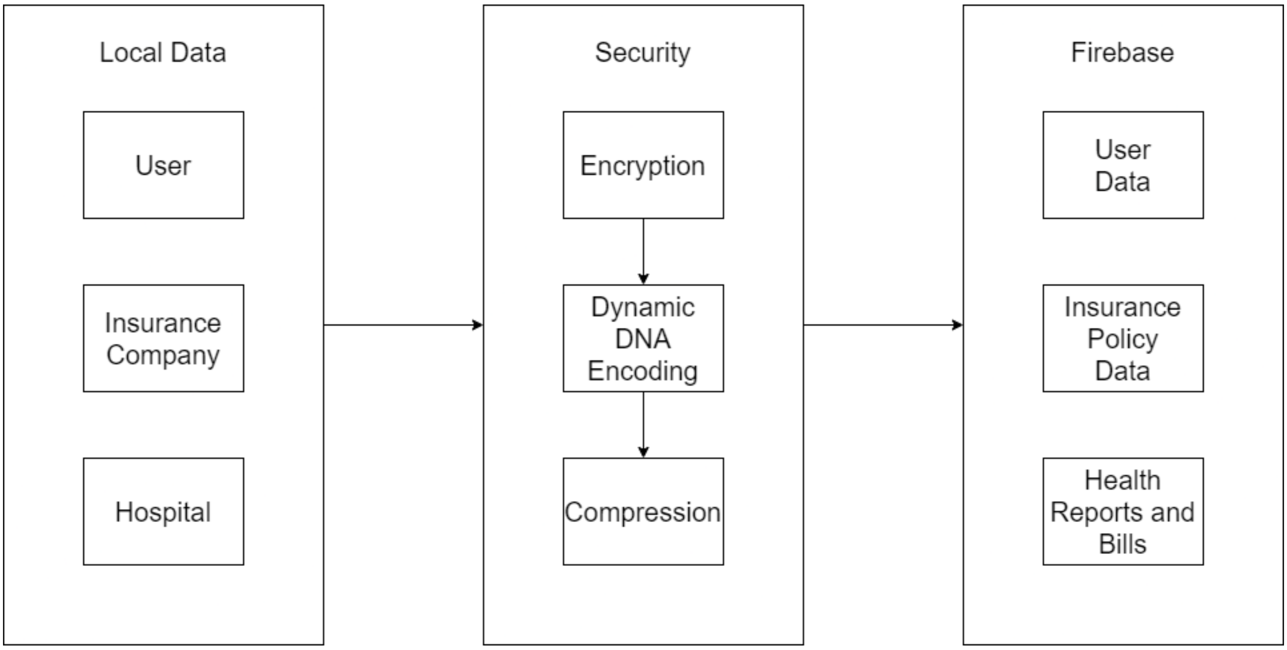


Fig 1: Architecture Diagram - 1

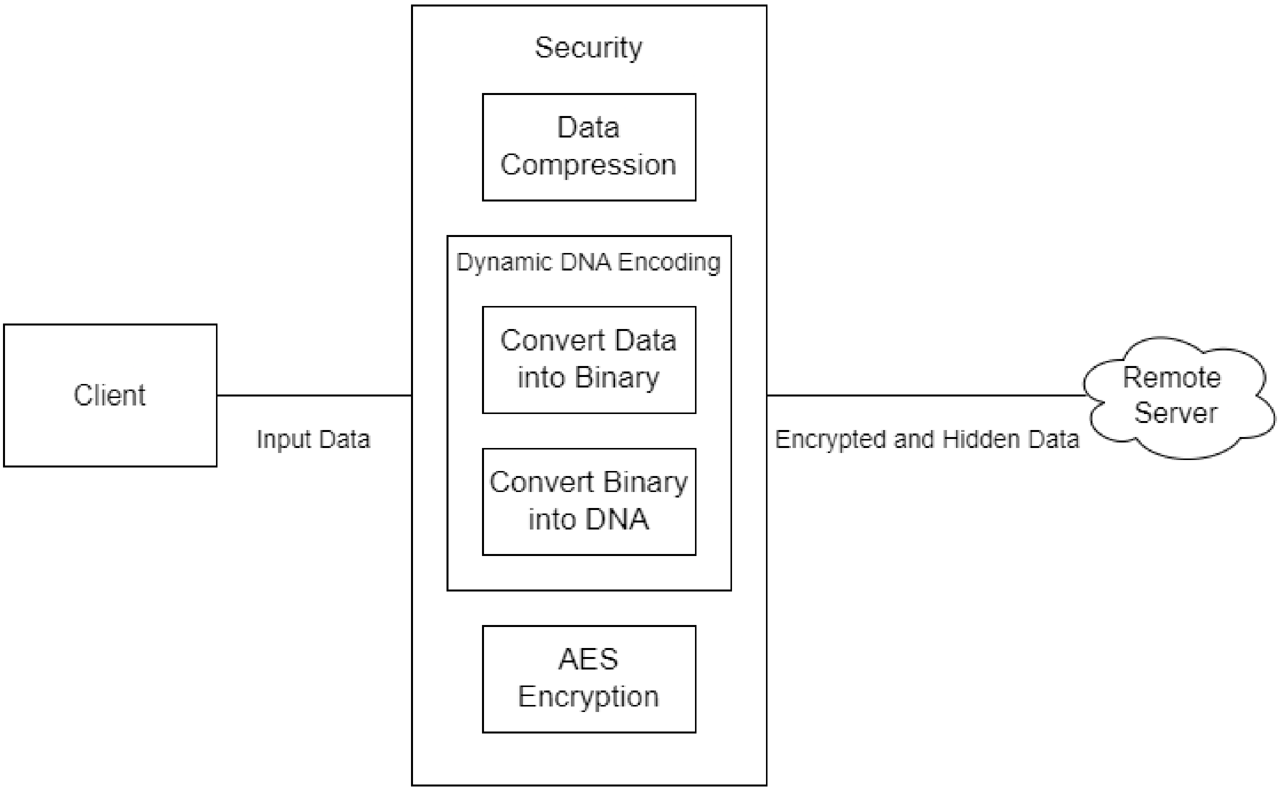


Fig 2: Architecture Diagram

IV. Result and Discussion

An overview of the text datasets used in the research is given in this section. These text datasets are used to test the success of the suggested system using a variety of performance criteria. The datasets are described after a discussion of these performance metrics.

4.1. Datasets utilized

Basic patient data, including phone numbers, addresses, and links to uploaded files, are included in our dataset. In healthcare settings, this data is critical for managing patient records and enabling communication and data access.

4.2. Performance Metrics

Different performance criteria are assessed for the text dataset in the proposed work. Key generation time and encryption time is used to measure the performance.

Key generation time: It is calculated by measuring how long the encryption algorithm takes to produce the authentication keys.

Encryption time: The term refers to the length of time the system takes to encrypt data. The amount of time required for the suggested method to encrypt the data is measured based on the input data.The other metrics that can be used are file upload time and decryption time.

File upload time: For text data, the term ”file upload time” describes how long it takes to move a text file from one place to another, usually over a file transfer protocol (FTP) or a network. It gauges the effectiveness and speed of data transfer processes and is affected by file size, network latency, and capacity on the network.

Decryption time: The time required by the cryptosystem to decode the provided data set is referred to as the decryption time. The anticipated amount of time needed for the suggested method to decode the data.

4.3. Text Dataset Performance Analysis

The next section discusses the performance metrics that were used to compare the suggested method's efficacy with other approaches that were already in use on the text dataset.

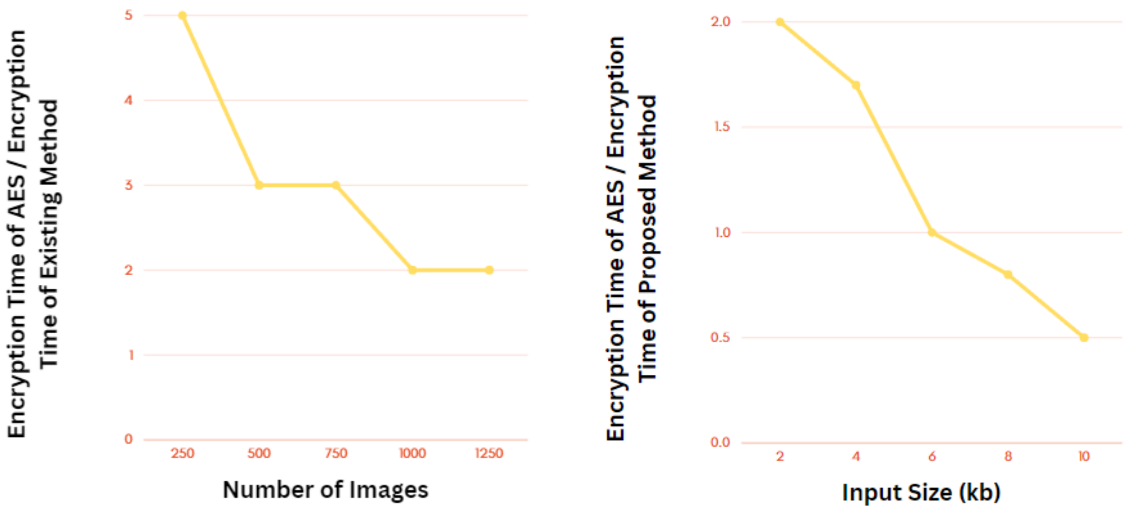


Fig 3: Comparison of Existing and Proposed Method's Encryption Time with reference to AES Encryption.

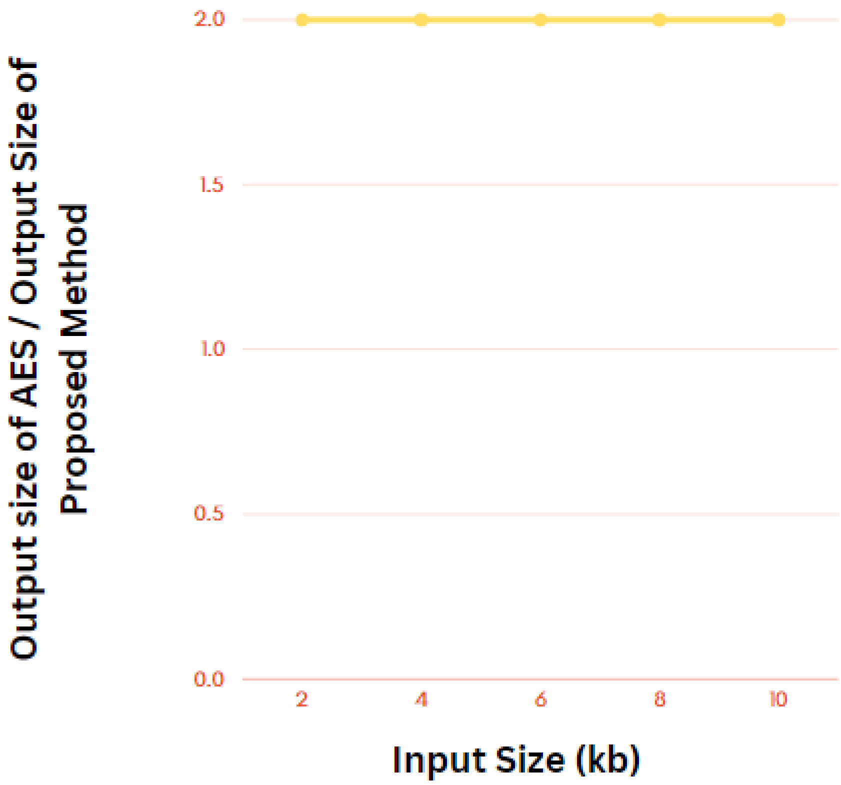


Fig 4: Output Size Graph

V. SUMMARY AND CONCLUSIONS

Digital technologies are being used more and more by the healthcare sector to manage health information, particularly private health insurance data. In order to improve the security of health insurance data, this article suggests a Hybrid Framework that combines user authentication procedures, real-time monitoring, and safe data storage. Using dynamic DNA encoding, the framework incorporates a novel method of encoding insurance data into DNA sequences that are subsequently stored in a secure cloud database. The project is to solve privacy and security issues related to health insurance data, especially in light of the increasing number of data breaches and cyberattacks. Healthcare businesses may safeguard confidential data, follow data protection laws, and build patient and stakeholder trust by implementing this framework.

In conclusion, this study offers a Hybrid Framework that tackles the pressing requirement for improved health insurance data security in the digital era. Health insurance data confidentiality, integrity, and availability are guaranteed by the framework through the use of secure data storage, real-time monitoring, and user authentication procedures. Healthcare data security problems can be uniquely solved by the creative method of encoding data into DNA sequences. By means of this initiative, healthcare establishments can alleviate risks and vulnerabilities linked to cyber attacks and data breaches, consequently upholding adherence to data protection laws and fostering confidence among patients and relevant parties. By protecting the confidentiality and integrity of health insurance data in a linked digital world, the hybrid framework helps to improve data security in the healthcare sector as a whole.

REFERENCES

1. Swetha Gadde, J. Amutharaj b, S. Usha c, 2021, Science Direct, ”A security model to protect the isolation of medical data in the cloud using hybrid cryptography.”https://doi.org/10.1016/j.jisa.2022.103412
2. Hongmin Li a,b, Tie Li a,b, Wei Feng c, Jing Zhang c Jun Zhang c ,Lixia Gan c ,Chunlai Li a,b,2021, Science Direct, ”A novel image encryption scheme based on non-adjacent parallelable permutation and dynamic DNA-level two-way diffusion.” https://doi.org/10.1016/j.jisa.2021.102844

3. Rajkumar Ettiyan , Geetha V, 17 February 2023, Science Direct, "A hybrid logistic DNA-based encryption system for securing the Internet of Things patient monitoring systems." https://doi.org/10.1016/j.health.2023.100149
4. Gunjankumar Bhoi; Raj Bhavsar; Priteshkumar Prajapati; Parth Shah "A Review of Recent Trends on DNA Based Cryptography" DOI: 10.1109/ICISS49785.2020.9316013
5. Batista BG, Ferreira CH, Segura DC. A QoS-driven approach for cloud computing addressing attributes of performance and security. Future Gener Comput Syst 2017;68:260–74.
6. Bharadwaja AV, Ganesan V. A novel hybrid image hiding technique using elliptic curve cryptography and DNA computing technique. Int J Electron Secur Digit Forensics 2021;13(4):460–73.
7. Chen C-M, Deng X, Kumar S, Kumari S, Islam SK. Blockchain-based medical data sharing schedule guaranteeing security of individual entities. J Ambient Intell Humaniz Comput 2021:1–10. Springer.
8. Chen F, Tang Y, Wang C, Huang J, Huang C, Xie D, Wang T, Zhao C. Medical cyberphysical systems: a solution to smart health and the state of the art. IEEE Trans Comput SocSyst 2021.
9. Chinnasamy P, Deepalakshmi P. HCAC-EHR: hybrid cryptographic access control for secure EHR retrieval in healthcare cloud. J Ambient Intell Humaniz Comput 2022;13(2):1001–19
10. Denis R, Madhubala P. Hybrid data encryption model integrating multi-objective adaptive genetic algorithm for secure medical data communication over cloudbased healthcare systems. Multimed Tools Appl 2021;80(14):21165–202.
11. Hureib ES, Gutub AA. Enhancing medical data security via combining elliptic curve cryptography and image steganography. Int J Comput Sci Netw Secur (IJCSNS) 2020;20(8):1–8.
12. Jintcharadze E, Iavich M. Hybrid implementation of Twofish, AES, ElGamal and RSA cryptosystems. In: 2020 IEEE East-West Design Test Symposium (EWDTS); 2020. p. 1–5.
13. Kumar P, Bhatt AK. Enhancing multi-tenancy security in the cloud computing using hybrid ECC-based data encryption approach. IET Commun 2020;14(18):3212–22.
14. Kumari A, Kumar V, Abbasi MY, Kumari S, Chaudhary P, Chen C-M. Csef: cloudbased secure and efficient framework for smart medical system using ecc. IEEE Access 2020;8:107838–52.
15. Maitra T, Obaidat MS, Giri D. Elgamal cryptosystembased secure authentication system for cloud-based IoT applications. IET Netw 2019;8(5):289–98.
16. Manaa ME, Hadi ZG. Scalable and robust cryptography approach using cloud computing. J Discrete Math Sci Cryptogr 2020;23(7):1439–45.
17. Khan JS, Ahmad J, Abbasi SF, Arshad, Kayhan SK. DNA sequence based medical image encryption scheme. In: 2018 10th computer science and electronic engineering. 2018, p. 24–9. http://dx.doi.org/10.1109/CEEC.2018.8674221.
18. Wang X, Liu L, Zhang Y. A novel chaotic block image encryption algorithm based on dynamic random growth technique. Opt Lasers Eng 2015;66:10–8. http://dx.doi.org/10.1016/j.optlaseng.2014.08.005.
19. Diaconu A-V. Circular inter-intra pixels bit-level permutation and chaos-based image encryption. Inform Sci 2016;355–356:314–27. http://dx.doi.org/10.1016/ j.ins.2015.10.027.
20. Wu X, Wang D, Kurths J, Kan H. A novel lossless color image encryption scheme using 2D DWT and 6D hyperchaotic system. Inform Sci 2016;349–350:137–53. http://dx.doi.org/10.1016/j.ins.2016.02.041.
21. Yin Q, Wang C. A new chaotic image encryption scheme using breadth-first search and dynamic diffusion. Int J Bifurcation Chaos 2018;28(04):1850047. http://dx.doi.org/10.1142/S0218127418500475.
22. Zahmoul R, Ejbali R, Zaied M. Image encryption based on new Beta chaotic maps. Opt Lasers Eng 2017;96:39–49. http://dx.doi.org/10.1016/j.optlaseng.2017.04. 009.
23. Ping P, Fan J, Mao Y, Xu F, Gao J. A chaos based image encryption scheme using digit-level permutation and block diffusion. IEEE Access 2018;6:67581–93. http://dx.doi.org/10.1109/ACCESS.2018.2879565.
24. Zefreh EZ. An image encryption scheme based on a hybrid model of DNA computing, chaotic systems and hash functions. Multimedia Tools Appl2020;79(5187):24993–5022. http://dx.doi.org/10.1007/s11042-020-09111-1.
25. Zhang Y, He Q, Xiang Y, Zhang LY, Liu B, Chen J, Xie Y. Low-cost and confidentiality-preserving data acquisition for internet of multimedia things.IEEE Internet Things J 2018;5(5):3442–51. http://dx.doi.org/10.1109/JIOT. 2017.2781737
26. Zhang Y, He Q, Chen G, Zhang X, Xiang Y. A low-overhead, confidentialityassured, and authenticated data acquisition framework for IoT. IEEE Trans Ind Inf 2020;16(12):7566–78. http://dx.doi.org/10.1109/TII.2019.2957404.
27. Zhang LY, Wong K, Zhang Y, Zhou J. Bi-level protected compressive sampling. IEEE Trans Multimed 2016;18(9):1720–32. http://dx.doi.org/10.1109/ TMM.2016.2581593

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE
SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details