Knowledge and Information Systems
https://doi.org/10.1007/s10115-023-01937-z

REGULAR PAPER

Merkle tree-blockchain-assisted privacy preservation of electronic medical records on offering medical data protection through hybrid heuristic algorithm

M. Lakshmanan[1] · G. S. Anandha Mala[2]

Received: 5 April 2023 / Revised: 16 June 2023 / Accepted: 10 July 2023

Abstract

A growing number of medical records are changed by the electronic folders that can be shared and transmitted in real-time in recent years as a result of the speedy improvement of data mechanisms and network methodologies. Yet, security breaches and privacy issues could affect medical data sent over open communication channels. Due to its distinctive qualities including immutability, blockchain technology, anonymity, decentralization, and, verifiability has drawn more notice in various fields. With a variety of multimedia techniques, healthcare organizations all over the globe are evolving into more user-centered, harmonized, and effective models. The management of enormous amounts of information, like records and photographs of all the individual improve the person's labor requirements and the security hazards. Recent advances in the healthcare industry have led to the creation of enormous numbers of electronic health records (EHRs). The owner of the data can manage the EHR data and share thanks with certain individuals in the EHR system. It is challenging for information to guarantee protection and detection procedures due to the enormous amount of information in the medical care models. Hence, in order to secure the EHR, the development of artificial intelligence techniques like blockchain and a new privacy preservation model is recommended for this work. Particularly, the Merkle tree is a key component of blockchain technology. In this proposed model, the Merkle tree is applied that helps for effective and safe authentication of huge data frameworks with the purpose of verifiability of the posted patient data. This model consists of two main phases, such as data sanitization and data restoration. Here, the sanitization operation is based on the creation of the optimal key using a new iteration-based firefly reptile search algorithm (IFRSA). The creation of the optimal key is established by functioning a multi-objective function which includes the factors like Euclidean distance, hiding ratio, and, data preservation ratio among encrypted data using the

✉ M. Lakshmanan
lakshmanan1909@gmail.com

G. S. Anandha Mala
gs.anandhamala@gmail.com

[1] Computer Science and Engineering, Sri Venkateswara College of Engineering, Sriperumbudur, India

[2] Computer Science and Engineering, Easwari Engineering College, Chennai, India

Published online: 04 September 2023

original key and the key with a variation. Finally, the electronic health care model becomes more private, dependable, and helpful through our suggested system. The comparison shows the supremacy of the developed approach.

Keywords Merkle tree-blockchain · Electronic medical records · Privacy preservation · Optimal key generation · Iteration-based firefly reptile search algorithm

1 Introduction

Personal health records (PHR), electronic medical records (EMR), and HER are the three primary stages of health-based electronic records that are currently utilized extensively for developing healthcare detection and present-time medicine. EMR and EHR are the health care documents that are saved in the medical organizations, while private health documents are gathered by private medical detecting devices. Both EHR and EMR are health care documents saved in desktops, and have the goal of authorizing doctors to develop health quality and handle health expenses [1, 2]. The environment of medical care can be made more tailored and individualized to patient's requirements by integrating EMR, EHR, and PHR [3]. EHR is described as inter-organizational transmission, whereas EMR is a file that is due within the organization. PHR is a patient-run online system that aims to make physiological data about sick persons clearly. So the user can comprehend and be involved in the therapeutic operation [4]. Additionally, it is safe to share, store, and protect medical electronic records, allowing telemedicine services, analysis of healthcare-related examinations, and increasing the transparency of the medical process [5].

At present a lot of methods are implemented to prevent the information from being obstructed, listened on, and altered in the operation of sending data because electronic records are sent over a network and the data sent over the internet have the issue of data security while data transmission [6]. Many researchers are attempting to offer Blockchain methodology to multiple standard modules in order to protect person privacy protection by utilizing the distinctive functionalities of Blockchain methodology, which consists of anonymity, traceability, information not being varied, node information synchronization, and finally decentralization [7]. This is done to efficiently end the issues connected to data protection [8]. All of the tools are modified to advantage the sick person because conventional medical databases will contain heterogeneous, time-varying amounts, of large, and complex data and because the health care organizations are forwarding toward sick person-centric systems [9]. Blockchain technology can deliver high-quality services for a reasonable price when used for digitized medical data [10]. This study goals to create a method, which enables authorized group members to access sick person EMR in order to achieve remote sick person monitoring and diagnosis [11].

The blockchain is a decentralized electronic dataset that helps to keep and upgrade any type of data, including documents, functions, and data transactions [12]. The blockchain may grow simultaneously as block prices rise, and the idea of a link is created by utilizing the hash function to identify the last section [13]. Contrasted to conventional record-keeping methods, blockchain technology is more dependable and trustworthy. By utilizing a smart contract to automate the conventional processes, operating costs and productivity are enhanced. Contrarily, one of the biggest drawbacks of blockchain technology is how much energy it requires to maintain a real-time distributed database [14]. Nonetheless, the area of data science is greatly impacted by blockchain technology in terms of being able to govern and change the raw data [15]. Consensus and cryptography approaches are utilized with the aid of blockchain technology to validate data and enable data analytics. Yet, from the standpoint of patients, location-connecting programs depend on blockchains and are susceptible to illegal security breaches [16]. It is already known that security maintenance, as a significant risk in the health care surroundings and various other implementations, has been intensively investigated. Common methods for protecting location privacy include differential privacy and K-anonymity [17]. However, these methods will result in information loss, making it impossible for us to obtain the intact region [18]. A privacy-preserving solution must be used in order to achieve various-stage security protection for the documents with the region in the blockchains because a blockchain can only guarantee user anonymity, not the security preservation of the region data stored on the blocks.

The major offerings of the implemented system are explained here.

- To design the framework of Merkle tree-blockchain-based privacy preservation of the EMR on providing medical data protection that helps to make the data as efficient, secure, and reliable preservation.
- To create the IFRSA algorithm, which infers the idea of the old firefly algorithm (FA) and reptile search algorithm (RSA) for optimizing the keys during the generation process.
- To accomplish the data sanitization operation for the better data preservation by employing the optimal key generation where the parameters are optimized by the suggested IFRSA system for enhancing the model reliability.
- To estimate the efficiency of the model employing multiple variables and comparing it with various conventional algorithms this shows the system functionality.

The structure of the paperwork is described here. Section 2 depicts the related works of the improved system. Section 3 describes about the novel architecture of the Merkle tree-blockchain-based privacy preservation of the EMR. Then Sect. 4 shows the demonstration of the developed system of the Merkle tree-blockchain-based privacy preservation model. Section 5 elaborates on the developed IFRSA algorithm. Section 6 explains the outcomes and discussions of the developed model. The last Sect. 7 summarizes the suggested scheme.

2 Existing works

2.1 Related works

In 2021, Ammeta and Parthiban [19] developed the HBESDM-DLD model, which combined secured healthcare information handling with deep learning-assisted diagnostics. The concept that was shown included many stages of activities such as encryption, the improvement of the optimal keys, secure data handling utilizing the Hyperledger blockchain, and detection. The displayed system gave the user the ability to manage data access, grant hospital administrators access to both reading and writing the information, and alarm important contacts. The "SIMON" block-cipher methodology was employed for encryption. A "Group Teaching Optimization Algorithm (GTOA)" was used for the SIMON technique's best key creation at the same time to improve its effectiveness. Also, the sharing of medical data has been done utilizing the multi-channel hyper ledger blockchain, which made use of a blockchain to store patient visit data and to record linkages for the "EHRs" that were saved in outside databases by the medical institutions. Finally, a "Variational AutoEncoder (VAE)"-assisted detection method was employed to identify the visibility of the disorder later the information had encrypted at the side of the receiver. The benchmark medical dataset was utilized for the functionality validation of the HBESDM-DLD model, and the outcomes were examined using a variety of performance metrics. The outcomes of the research explained that the HBESDM-DLD methodology was superior to cutting-edge techniques.

In 2021, Zhu et al. [20] have developed an enhanced “Merkle tree-based blockchain EMR storage” solution. The recommended scheme’s distinguishing feature was the replacement of the raw binary tree framework in the suggested convolutional Merkle tree with a convolutional layer framework, which could significantly enhance the efficiency. Studies have revealed that, for the similar count of input information, the amount of stored nodes had fallen noticeably, and the number of layers in the modified convolutional Merkle tree and the hash calculation count have both been drastically reduced. The effectiveness and security research further demonstrated that the suggested plan could offer a dependable option for the future advancement of data storage security.

In 2020, Huang et al. [21] have presented a blockchain-based security-preserving model that achieved safely sending of medical data between several entities involved in semi-trusted cloud servers, research institutions, and patients. In the meantime, it was possible to gain consistency and data availability between sick persons and research institutions. And proxy re-encryption methodology was utilized to verify that experiment organizations could decrypt the middle ciphertext. The zero-knowledge proof was applied to confirm that the sick person’s medical data complies with the requirements put forth by research institutions without disclosing the identity of the sick person. Additionally, this plan could carried out split consensus depended on the PBFT method for the transactions between research institutions and sick persons under the predetermined conditions. Theoretical analysis has shown that the suggested scheme could meet privacy and security requirements like performance evaluation, availability, integrity, and confidentiality and has shown that it was effective and practical in comparison to other common schemes.

In 2021, Iqbal et al. [22] have recommended a framework blockchain-assisted “Reliable and Intelligent Veterinary Information Management System (RIVIMS)” utilizing smart contracts and machine learning methods. Blockchain-assisted safe veterinary data handling and the assumed and data analytics phases made up the two major components of the suggested RIVIMS. First, + Hyperledger Fabric was utilized to create a trustworthy and secure data administration model for veterinary clinics. Second, the permission blockchain architecture was employed to create smart contracts, predictive analytics, and modules for data. The goal of the predictive and data modules was to evaluate sick person appointment information from veterinary clinics to identify the modern trends and create a reliable assumption method utilizing machine learning methods. Predictive and Data analytics have aided veterinary management in making better business choices for the future so that veterinary sick persons could receive better healthcare. The developed blockchain-assisted system’s efficiency was assessed using Hyperledger Caliper as a benchmarking tool in terms of latency, throughput, transaction success rate, and transactions per second. Additionally, machine learning performance metrics like “MAE, RMSE, and R2” scores have been utilized to assess the general effectiveness of the assumption method. The research findings showed that the suggested RIVIMS was more effective and reliable.

In 2020, Li et al. [23] explained a Blockchain-based data aggregation system for healthcare surroundings. Besides, to develop local health care detection, they have developed a set of authentication methodologies for various authenticated persons to freely access a sick person’s personal health data. A group session key would be decided upon by the authorized group members and utilized to secure sensitive patient data. The session key needed to be upgraded whenever a new person adds the health care group or a departing person joined the group. Finally, our suggested plan was to enhance the utility, dependability, and safety of the computerized healthcare system.

In 2019, Nguyen et al. [24] have developed a special EHRs sharing architecture that combined the blockchain and the decentralized “Interplanetary File System (IPFS)” on a moving cloud module. Importantly, they have developed a dependable enter control system utilizing smart assignments to achieve safe EHR transmission between various sick persons and healthcare professionals. With a moving app employed to Amazon cloud computing, they demonstrated a framework creation utilizing Ethereum blockchain in the present information exchanging scenario. The empirical findings demonstrated that the expert’s idea offered a practical approach for trustworthy data replacing on moving clouds during safeguarding private health data from potential risks. The security analysis and system assessment also proved the functionality gained in easy access handling design, low network latency with maximum safety, and information privacy stages, contrasted to the previous information sharing methods.

In 2020, Ji et al. [25] have experimented with the region-sharing assisted on blockchains for telecare medical information systems. First, experts have outlined the fundamental conditions for verifiable, retrievable, multi-level, confidential, unforgeable, and decentralized region sharing on blockchains. Then, researchers suggested a blockchain-assisted multi-level region-sharing model named as BMPLS, which utilized order-preserving encryption and Merkle trees. The analysis’s findings demonstrated that the expert’s plan met the aforementioned criteria. Lastly, the effectiveness of the expert’s plan was assessed, and the results of the experiment demonstrated that it was effective and workable for both sick patients and medical professionals. In a nutshell, the researcher’s plan might be utilized to implement region sharing for telecare medical information methods while protecting privacy.

In 2020 Rathee et al. [26] developed a safety architecture of the medical care multimedia information by establishing the hash of entire information so that any variation in the information or breach of pharmaceuticals might be affected by every user of the blockchain module. The outcomes have been contrasted to the old way and proven with enhanced results that, the Blockchain technique, offered an 86% success rate over scenarios involving probabilistic authentication, wormhole attack, falsification attack, and product drop ratio.

In 2023 Zhang et al. [32] designed a privacy-preserving EMR sharing system model according to consortium blockchain that was called as EMRShareChain. Particularly, an Improved PBFT consensus algorithm (I-PBFT) was proposed for improving the blockchain block-out stability and efficiency that is helped for making it more suitable for medical scenarios. At last, a prototype development of the aforementioned model was executed based on the Hyperledger Fabric approach.

In 2023 Sharma et al. [33] developed a secure blockchain-based proposed application (PA) to maintain, validate, and generate healthcare certificates. The PA acts as a communication medium among the backend blockchain application and network entities such as doctors, IoT devices, patients, and hospitals for generating and verifying medical certificates. At last, the experimental evaluation of the designed work has shown that it effective solution than the other baseline schemes.

In 2023 Cerchione et al. [34] introduced the information processing theory (IPT) which allowed us to validate and design a blockchain-based EHR system for enhancing the storage of medical records. The improvements of this system helped to attain better organizational outcomes, managerial outcomes, and distributed networks regarding clinical outcomes. At last, the experimental outcomes of the designed method were utilized to explore the better solutions of the current system.

2.2 Research gaps and challenges

The existing EHR analysis systems are not helpful for meeting the complicated problems as they do not include a permanent architecture for reliability and data safety policies. Thus, a novel solution is needed for guaranteeing accountability of the utilization of medical record systems, for regulating the government security policies, and for maximizing the data accessibility. In addition, it is also necessary for analyzing the historical data of the patients through several data mining approaches, machine learning algorithms, and blockchain-based approaches, which are intended for predicting the future scheduling requests in future. Thus, there is a need of protecting the medical data through blockchain strategies. Numerous privacy preservation schemes are discussed in Table 1. GTOA [19] has alerted emergency contrasts by encrypting the health records and has been used for assisting the efficient disease diagnosis process. Conversely, this model is not suitable for applying heuristic algorithms for performing the learning rate schedules and hyperparameter optimizer. Merkle tree based-blockchain [20] obtains a superior performance in the electronic medical records and query application and guarantees the convenience and security of data storage. However, it takes higher computational time. The "Practical Byzantine Fault Tolerance (PBFT)" algorithm [21] has satisfied the privacy and security requirements like availability, integrity, and confidentiality and has used a proxy re-encryption module for transferring the encrypted clinical data to the intermediary cipher text. It suffers from conspiracy attacks. RIVIMS [22] has ensured performance regarding transaction latency, transaction throughput, transaction success rate, and transaction per second. It has been suitable for managing the veterinary clinic resources clinically and reliably. This model does not evaluate the interoperability. Blockchain transformation with group authentication [23] has protected the private medical records of patients from improper access and has reduced the medical errors. Though, it does not meet some security functional requirements and addresses common security attacks. Asymmetric encryption algorithm [24] has achieved higher flexibility for mobile users, provided data privacy, system integrity, and higher security levels, and shared medical data over mobile cloud frameworks more quickly and reliably. Conversely, it is not accessible for designing new real-time applications. Merkle tree [25] is more suitable for telecare medical information systems and specifies their efficiency in real-time applications. It faces a higher computational overhead. Blockchain Technology [26] has offered higher security chains of the recorders with the generation of hash function and shown outperformance by evaluating over illegal activities. It does not evaluate the multi-media data over transaction cost or time. Thus, these challenges motivate the research workers to develop a new safety preservation system.

3 Novel framework of Merkle tree-blockchain-assisted privacy preservation of electronic medical records

3.1 Dataset details

The developed Merkle tree-blockchain-based privacy preservation model utilizes two standard datasets which contain medical data for individual patients. The first dataset includes 764 medical records, whereas the second dataset contains 304 medical records. By employing these two datasets the better results are obtained for the improved system.

Table 1 Features and challenges of conventional privacy preservation of HER data

Author [citation]	Methodology	Features	Challenges
Sammeta and Parthiban [19]	GTOA	It has alerted emergency contrasts by encrypting the health records It has been used for assisting the efficient disease diagnosis process	This model is not suitable for applying heuristic algorithms for performing the learning rate schedules and hyperparameter optimizer
Zhu et al. [20]	Merkle tree based-blockchain	It obtains a superior performance in the electronic medical records and query applications It guarantees the convenience and security of data storage	However, it takes higher computational time
Huang et al. [21]	Practical byzantine fault tolerance (PBFT) algorithm	It has satisfied the privacy and security requirements like availability, integrity, and confidentiality It has used a proxy re-encryption scheme for sending the encrypted clinical information to the intermediary cipher text	It suffers from conspiracy attacks
Iqbal et al. [22]	RIVIMS	This model has ensured the performance regarding transaction latency, transaction throughput, transaction success rate, and transaction per second It has been suitable for managing the veterinary clinic resources clinically and reliably	This model does not evaluate the interoperability
Li et al. [23]	Blockchain transformation with group authentication	This model has protected the private medical records of patients from improper access It has reduced the medical errors	Though, it does not meet some security functional requirements and addresses common security attacks
Nguyen et al. [24]	Asymmetric encryption algorithm	This system has achieved higher flexibility for mobile users and provided data privacy, system integrity, and higher security levels This model has shared the medical data over mobile cloud frameworks more quickly and reliably	It is not accessible for designing a new real-time application

Table 1 (continued)

Author [citation]	Methodology	Features	Challenges
Ji et al. [25]	Merkle tree	It is more suitable for telecare medical information systems It specifies their efficiency in real-time applications	It faces a higher computational overhead
Rathee et al. [26]	Blockchain technology	The proposed framework has offered higher security chains of the recorders with the generation of the hash function It shows outperformance by evaluating over illegal activities	It does not evaluate the multi-media data over transaction cost or time

3.2 Proposed privacy preservation model of electronic medical records

Nowadays, most of the EMRs are employed to replace the conventional recording techniques. Besides, EMR includes a multiple amounts of user's personal privacy data like the individual's general information, reports from the examination, advice of the doctors, a few of the ultrasound pictures, and, so on. This type of medical information is created every day. Hence, handling and securing this kind of information effectively and safely has become very important. The blockchain is the latest application technique that contains data storage in a distributed manner, consensus technique, point-to-point data transfer, cryptographic methodology, and various computer applications. The blockchain method is developed from the topic of bitcoin, which is split transmitted ledger datset architecture. The blockchain could be categorized into 3 namely, public chain, personal chain, and association chain [35]. The details of the information in the blockchain are saved in blocks with the help of cryptographic technologies here the sections are connected like a link in a orderly manner. The Merkle tree is an important section in the blockchain [36] methodology. It does not require downloading the entire transaction information to check if a transaction has been permitted through the entire network. The conventional Merkle tree is a binary tree framework which saves multiple hash values. Merkle tree framework can reduce the amount of hash evaluation in creating the root node. The upcoming Fig. 1 shows the diagrammatic illustration of the developed system.

The main objective of this paper is to enhance a Merkle tree-blockchain-assisted privacy preservation of EMR by providing medical data protection via the developed IFRSA algorithm. In this implemented system, the Merkle tree is performed that helps to secure and efficient confirmation of the multiple data structures. This data structure offers verifiability of the shared sick person information. This model contains two modules, such as data restoration and information sanitization. The sanitization of the data works according to the optimal key creation employing the suggested IFRSA algorithm. The generation of the best key is established by applying the various-objective function that contains the variables like information preservation rate, hiding rate, and the Euclidean distance among the encrypted information utilizing the varied key and the original key. At last, the electronic medical system becomes more reliable, useful, and secure by the developed method. The comparison explains the supremacy of the suggested scheme.

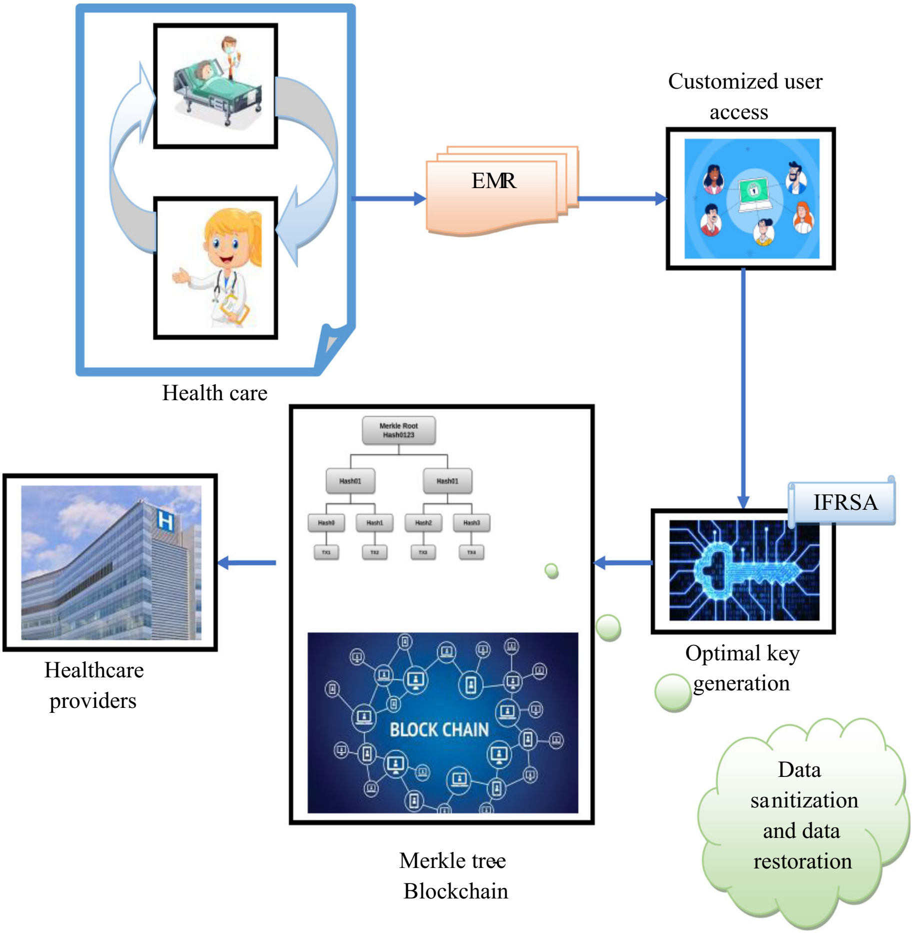


Fig. 1 The architectural diagram of the developed model

3.3 Problems of privacy preserving the data

The issues that are faced in the data safety preservation method are explained here.

- There are several efficient approaches to data mining that preserves privacy. However, the majority of these techniques could have some unintended consequences, such as diminished data value and decreased data mining effectiveness.
- The amount of noise produced is sufficient to prevent the recovery of the record's individual values.
- Frequently, the data records are made publicly accessible by simply omitting crucial identifiers from individual information, including the name and social security number.
- The k-anonymity techniques largely concentrate on a general strategy that preserves all individuals equally, without taking into account their specific requirements. The result could be that certain groups of people receive insufficient protection, while others receive overly intrusive privacy controls.

- A malicious opponent will do anything to deduce sensitive information. They have the power to terminate the protocol at any time, send fake or spoof communications, collaborate with other (evil) parties, and other actions.
- Although they properly execute the protocol, semi-honest opponents may attempt to deduce the other parties' secret information from the data they observe.
- Digital signature techniques require knowledge of the underlying data for verification, which makes it easy to discover the true identity of a data source.

4 System demonstration of Merkle tree-blockchain-assisted privacy preservation scheme

4.1 System model of blockchain

The blockchain-assisted EMR secure storage method includes four modules, such as candidates, the scheme server, the cloud server, and the blockchain server. The details of these four modules are expressed below.

Users: This module refers to the sick persons who wish to update and secure EMR or the other modules that need to query the information. At first, the entire user must register the scheme server and obtain the special ID after the registration. This ID must be kept secure and set as a symmetric key Ky_t to encrypt the information for privacy.

Scheme server: The responsibility of this server is to secure the entire EMR communication scheme that creates the mission variables and sends conditions between the candidates, the blockchain server, and the cloud server.

Cloud server: This server saves the encrypted cipher text D_L of the user and results in the time message N and the storage region.

Blockchain server: This server builds an enhanced convolution-assisted Merkle tree with the help of encrypted cipher text D_L which belongs to the users. Next, the encrypted message, the pre-block's hash merit, the root merit of the Merkle tree, and the timestamp in the section are included in the block of the suggested blockchain also.

The data storage contains the request, the storage, encryption, and the record. The description of this stage is given below.

Request: To secure the cipher text D_L, the candidate V transmits a request to the scheme server. Then the request R_T stage is splitted into two methods.

- The storage request R_T is transmitted to the scheme server. Here u is the time request.
- The scheme server permits the cipher text D_L that the candidate needs to secure after checking the candidate's V request R_T and the identity ID. Next, the scheme server sends the cipher text D_L to the blockchain server.

Storage: The cloud server secures the cipher text D_L. Then it will send the storage information N and the storage location to the scheme server. Besides, the scheme server transmits N to the candidate V.

Encryption: After gaining the data V, the user N encrypts V and the symmetric key Ky_t joins to create a tag S by $S \leftarrow F\{N||\mathrm{Ky}_t\}$ then transmits the tag S to the scheme server.

Record: The candidate N wants to store the information V and cipher text D_L on the blockchain.

The scheme request, cloud query, blockchain query, verification, and decryption process should perform if the candidate wants to query the submitted EMR. These stages are expressed here.

Request: To extract the cipher text D_L, the candidate N sends the request to the scheme server.

Blockchain query: This process is done in the below steps.

- The scheme server questions the root section Q_0 and the tag S from the blockchain server based on the index I_G.
- Next, the blockchain sends the particular tag S and Merkle root value Q_0 to the scheme server later gaining the index I_G.
- At last, the scheme server sends the tag S to the candidate N.

Cloud query: The candidate N queries the cipher text D_L from the cloud server with the help of a tag S.

Verification: The scheme server creates the root merit $Q_{0'}$ of the cipher text D_L with the enhanced Merkle tree architecture and checks it with the root section Q_0. The scheme server transmits the cipher text D_L to the candidate N, if $Q_{0'} = Q_0$. Or else, the scheme server provides the candidate N review that the file G has been invaded.

Decryption: Later the candidate N gains the cipher text D_L, the candidate N decrypts, and lastly obtains the file G. Figure 2 displays the architecture of the blockchain for the

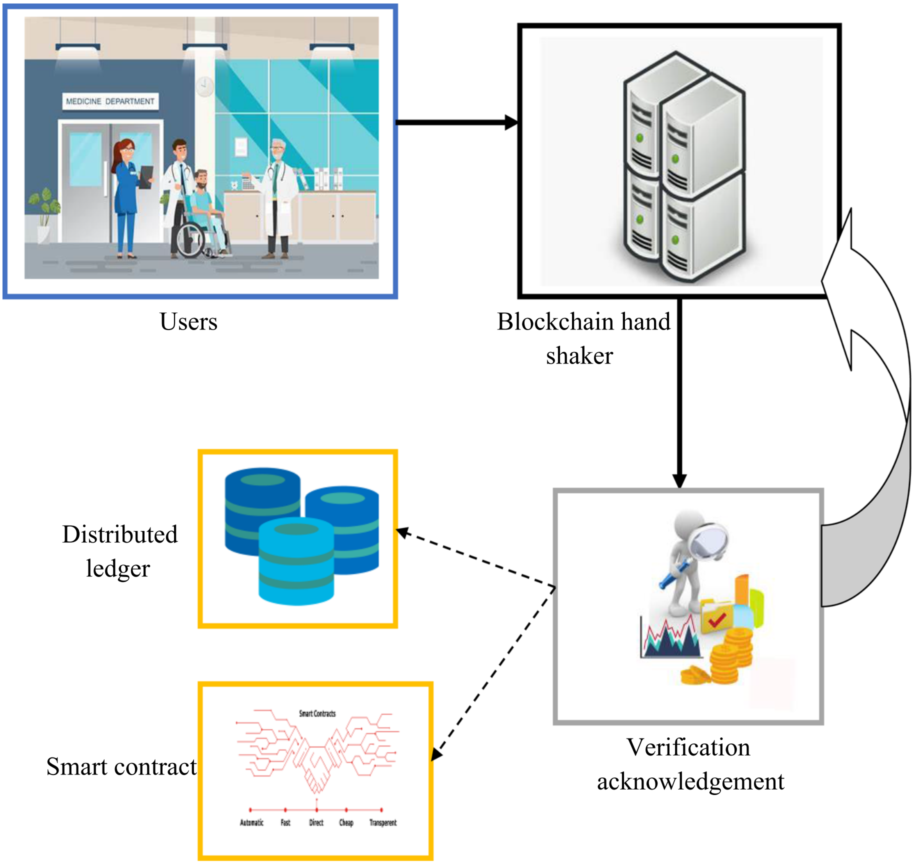


Fig. 2 The framework of the blockchain model for the suggested scheme

developed model.

4.2 Merkle tree structure

The conventional Merkle tree [27] is a binary tree framework which saves multiple hash merits. The five-layer Merkle tree contains the root node $I_{4,1}$ at the peak of the Merkle tree, a set of middle nodes $I_{3,n},\ n = 1, 2$, in the center of the Merkle tree $I_{2,k},\ k = 1, 2, 3, 4$, and a collection of child nodes $I_l,\ l = 1, 2, ..., 8$ at the base of the Merkel tree. The evaluation method of the parent node $I_{4,1}$ is described as below.

At first, assume E_i be the ith input data. Here $i = 1,\ 2,\ \ldots,\ 8$. Then the term $I(.)$ is a hash function. Next, the input data E_i is encrypted with $I(.)$ to obtain the leaf nodes I_i that are formulated in Eq. (1).

$$I_i = I(E_i),\ i = 1,\ \ldots,\ 8. \tag{1}$$

Moreover, the forwarding layer is offered by the child nodes I_i and the middle nodes could be evaluated with the help of Eq. (2).

$$I_{2,i} = I(I_{2i-1}, I_{2i}),\ i = 1, ..., 4 \tag{2}$$

Next, the previous step will be repeated and create the third layer with the utilization of Eq. (3).

$$I_{3,n} = I(I_{2,2n-1}, I_{2,2n}),\ n = 1, 2 \tag{3}$$

At last, the root node resultant can be acquired from Eq. (4).

$$I_{4,1} = I(I_{3,1}, I_{3,2}) \tag{4}$$

Commonly, the child nodes save the hash merits of a particular collection of information and the middle nodes save the hash values estimated by the points of the two low-level linked leaf nods. The parent node saves the final hash measure of the points of the last two middle nods. The Merkle tree has a strong bond so the particular middle node will vary if any of the child nodes get varied. Simultaneously, the root node also gets changed. Figure 3 shows the architecture of the Merkle tree.

4.3 Key generation using IFRSA

In the suggested system, the key retrieve operation gives an important part in the two information restoration and the data sanitization operation, here optimization process is carried out by the developed IFRSA. The answer transformation is the initial stage for the creation of the key operation. Here the key K is transformed into the latest form employing the Kronecker process. The key K is transformed into K_1 by utilizing Eq. (5), here the size is denoted as $\sqrt{no''} \times Q_{\max}$. For the sample, for the key $K = \{5, 6, 2\}$, the key matrix is provided in Eq. (5).

$$K_1 = \begin{bmatrix} 5 & 5 & 5 \\ 6 & 6 & 6 \\ 2 & 2 & 2 \end{bmatrix}_{\left[\sqrt{no''} \times Q_{\max}\right]} \tag{5}$$

In the above Eq. (5), the count of the transaction is represented as no the nearest maximum optimal power of no is given as no'', and the highest transaction length is indicated as $Q_{\max}$.

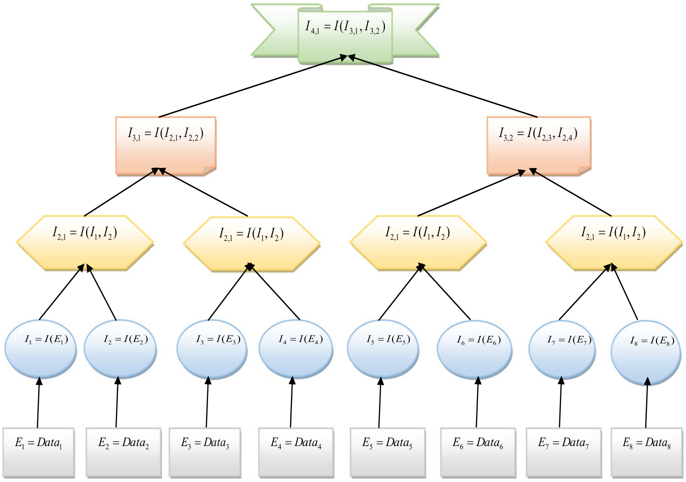


Fig. 3 The fundamental architecture of the Merkle tree for the developed system

According to Eq. (5), the regenerated key matrix K_1 is generated by running row-wise replication. Moreover, the key matrix K_2 is created by the Kronecker operation and it is shown in Eq. (6).

$$K_2 = K_1 \otimes K_1 \tag{6}$$

The symbol $\otimes$ denotes the product of Kronecker. The length of K_2 is also referred to as $\sqrt{no''} \times Q_{\max}$.

4.4 Privacy preservation with optimal key generation

Information of sanitization is the operation of obscuring the sensitive information or data in a Merkle tree that helps to ignore the information from the publishing on to an unknown point. On another side, the information restoration process is the opposite operation of the information sanitization operation that is performed by the evaluation of sanitization effectiveness.

In the sanitization operation, the binary conversion is performed for the both key matrix generation and the Merkle tree data. Here, the recommended IFRSA is utilized for creating the best key. The gathered binary information went to the XOR function for gaining the sanitized information, the sanitized information is created as expressed in Eq. (7).

$$\mathrm{DS}' = \mathrm{DS} \oplus K_2 \tag{7}$$

In the above Eq. (7), the sanitized information is represented as DS′ the raw data is denoted as DS, and the correctly created key is pointed by K_2. In the suggested model the data E_i are employed for the sanitization and obtained the data as $E_{i'}$ later during the sanitization method. Hence, the quickly responsive data are obscured in the sanitization operation that is then sent to the Merkle tree. Therefore, the information is secured for the next usage, which could be enhancing the functionality of the security in the Merkle tree module without any malicious attacks. In the restoration process, the raw information is recovered by employing

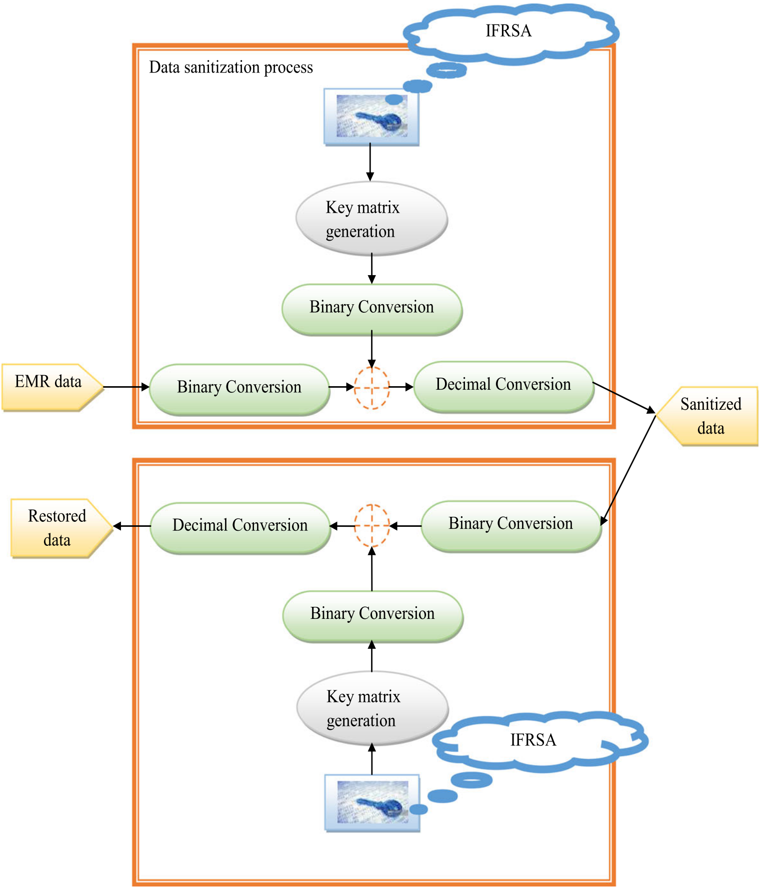


Fig. 4 Pictorial representation of the privacy preservation with the optimal key generation for the developed system

a similar key utilizing the recommended IFRSA algorithm. The step is given in Eq. (8).

$$\overset{\wedge}{\mathrm{DS}} = \mathrm{DS}' \oplus K_2 \tag{8}$$

Here, the parameter $\overset{\wedge}{\mathrm{DS}}$ denotes the restored data. Figure 4 shows the privacy preservation with optimal key generation.

5 Iteration-based firefly reptile search algorithm for privacy preserving the electronic medical records

5.1 Basic FFA

FA [28]: This algorithm is mainly focused on the character of the fireflies in the tropical summer sky. Fireflies contact with each other, search the prey and locate their mates utilizing the bioluminescence with different flashing patterns. Imitating the nature, multiple metaheuristic algorithms can be performed. The rules which are applied in the existing algorithms are described here.

- All the fireflies are unisex therefore one firefly will be engaged with other fireflies concerning the sex.
- The attracting character is proportional to the brightness of the firefly. The less bright one is moving forwards to the shiner one. If there are no brighter fireflies when compare to the significant firefly, it will forward arbitrarily in the space.
- The shining of the firefly is connected to an analytical section of the cost performance.

The evaluation model of the old algorithm is described below.

The starting places of agents are established in the search space given in Eq. (9).

$$y_{j,k}^{(0)} = y_{j,\min} + rnd \cdot (y_{j,\max} - y_{j,\min}),\ \ j = 1, 2, ..., O \tag{9}$$

Here, the variable $y_{j,k}^{(0)}$ represents the starting value of the jth parameter for the kth agent. $y_{j,\min}$ and $y_{j,\max}$ are the lowest and the highest acceptable measures for the jth variable.

The fireflies could split as a subgroup into multiple distinct groups so they can handle the multimodel issues by themselves. Then the attractiveness is based on the cost function. The traditional FA algorithm represents two main issues. One change is the formulation and the intensity of light of the attractiveness. The Gaussian form of the both absorption and the inverse square law is formulated in Eq. (10).

$$J(s) = J_0 e^{-\alpha s^2} \tag{10}$$

Here the parameter J indicates the intensity of light, J_0 is pointed to the original intensity of the light and the value α denotes the coefficient of the light absorption which is a constant value. The attractiveness δ of the firefly is estimated in Eq. (11).

$$\delta(s) = \delta_0 e^{-\alpha s^2} \tag{11}$$

Here, the variable δ_0 is a constant and indicates the attractiveness at the range at $s = 0$.

The space between the two fireflies j and k at the yj and yk accordingly could be explained as the Cartesian distance $sjk = |yj - yk|$.

The motion of the firefly j that is allured to the other brighter firefly k is calculated in Eq. (12).

$$\Delta y_j = \delta_0 e^{-\alpha s_{jk}^2}(y_k^u - y_j^u) + \chi \gamma_j,\quad y_j^{u+1} = y_j^u + \Delta y_j \tag{12}$$

Here the first parameter is a because of the attraction since the next variable is randomization with χ being the randomization variable. Then the parameter γ_j is a vector of the arbitrary numbers that are shown in the Gaussian distribution. The step size which is an arbitrary number estimated from Eq. (13).

$$M(t) = Bt^{-(1+\varphi)},\ \ B = \varphi \Gamma(\varphi) \sin\left(\frac{\varphi\pi}{2}\right)\frac{1}{\pi} \tag{13}$$

The parameter $\Gamma(\varphi)$ is a Gamma function and φ is indicated as the exponent for the distribution function. The function which is utilized for the geometrical annealing schedule is shown in Eq. (14). This is initializing from the start χ_0.

$$\chi = \chi_0 \kappa^u \tag{14}$$

Here $0 < \kappa < 1$ is the randomness reduction function.

Algorithm 1: FA

```
Consider the entire population and iteration count
Initialize the population and estimate the objective function.
Initialize the search space
For $j = 1 to u$
        For $k = 1$ to $u$
            Utilizing Eq. (9) evaluate the search space.
            Employing eq. (10) find out the Gaussian form.
            Estimate the movement of the firefly using Eq. (11)
            $k = k + 1$
            End
        Rank the fireflies and find out the optimal values
End
Returns the optimal candidate results acquired by the FA algorithm
```

5.2 Basic RSA

RSA [29]: This algorithm is mainly focused on the hunting features, encircling features, and the natural characteristics of the crocodiles. This mechanism is a gradient-free model and population-assisted because of that can be utilized to handle the complicated optimization issues. The experimental model is described here.

This optimization model initiates with a set of solution answers Y which is given in Eq. (15).

$$Y = \begin{bmatrix} y_{1,1} & \cdots & y_{1,k} & y_{1,m-1} & y_{1,m} \\ y_{2,1} & \cdots & y_{2,k} & \cdots & y_{2,m} \\ \cdots & \cdots & y_{j,k} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{M-1,1} & \cdots & y_{M-1,k} & \cdots & y_{M-1,m} \\ y_{N,1} & \cdots & y_{M,k} & y_{M,m-1} & y_{M,m} \end{bmatrix} \tag{15}$$

The term Y is a collection of the candidate answers which are created arbitrarily by employing Eq. (15), $y_{j,k}$ which explains the kth place of the jth answer, M denotes the count of the candidate answers, and m the is dimension size of the explained issue.

$$y_{j,k} = \text{rnd} \times (\text{ub} - \text{lb}) + \text{lb}, \quad k = 1, 2, ..., m \tag{16}$$

The parameter rnd is the arbitrary value ub and lb indicates the higher and the base bounds of the explained issue. The position up gradation of the suggested RSA algorithm is estimated in Eq. (17).

$$y_{(j,k)}(u+1) = \begin{cases} bst_k(u) \times -\alpha_{(j,k)}(u) \times \beta - S_{(j,k)}(u) \times \text{rnd}, & u \leq \frac{U}{4} \\ bst_k(u) \times y_{(s1,k)} \times es(u) \times \text{rnd}, & u \leq 2\frac{U}{4} \text{ and } u > \frac{U}{4} \end{cases} \tag{17}$$

The variable $bst_k(u)$ is the place in the good-acquired answer then the term rnd denotes the random number that falls in 0 and 1, u is the amount of the present loop, and U denotes the highest number of the iteration. Next, the variable $\alpha_{(j,k)}$ represents the hunting operator of the kth place in the jth answer that is estimated employing Eq. (18). β is called a sensitive variable which handles the exploration correctness for the encircling section over the loops that is similar to 0.1. The Reduction operation $S_{(j,k)}$ is a value that is estimated in Eq. (19) and utilized to decrease the search region. The random number which limits between [1 M] is represented as s_1 and the position of the random variable is indicated as $y_{j,k}$ of the jth answer. M is pointed out by the candidate's answers. The term $es(u)$ denotes Evolutionary Sense which is estimated in Eq. (19).

$$\alpha_{(j,k)} = bst_k(u) \times Q_{(j,k)} \tag{18}$$

$$S_{(j,k)} = \frac{bst_k(u) - y_{(s_2,k)}}{bst_k(u) + \sigma} \tag{19}$$

$$es(u) = 2 \times s_3 \times \left(1 - \frac{1}{U}\right) \tag{20}$$

The parameter σ is a little value and s_2 is the arbitrary value that is limited between [1 , M]. The value 2 is employed as a correlation measure to offer the measures between 2 and 0, which also term s_3 indicates the arbitrary numbers $-$ 1 and 1. The variable $Q_{(j,k)}$ is the probability variation between the kth place of the good-acquired answer and the kth place of the present answer that is formulated in Eq. (21).

$$Q_{(j,k)} = \alpha + \frac{y_{(j,k)} - N(y_j)}{bst_k(u) \times (\mathrm{ub}_{(j)} - \mathrm{lb}_{(j)}) + \sigma} \tag{21}$$

Here, the parameter $N(y_j)$ is the mean place of the jth answer that is evaluated utilizing Eq. (22). $\mathrm{lb}_{(j)}$ and the term $\mathrm{ub}_{(j)}$ is the base and the maximum of the kth places accordingly. α is denoted as a sensitive number that handles the exploration correctness for the hunting process over the set of iterations that is similar to 0.1.

$$N(y_j) = \frac{1}{m}\sum_{k=1}^{m} y_{(j,k)} \tag{22}$$

Algorithm 1: RSA

```
Note the entire population and looping count
Draw the candidate matrix $Y$ .
Initialize the population and analyze the objective function.
For $j = 1 \; to \; N$
        Evaluate the candidate solutions employing Eq. (15)
        Upgrade the position of the entire member utilizing Eq. (16)
        Estimate the evolutionary sense by applying Eq. (18)
        $j = j + 1$
        End
    Update the optimal candidate solution
End
Returns the optimal candidate results acquired by RSA
```

5.3 Recommended IFRSA

The functionalities of the two existing algorithms are described. The FA is based on the environmental behaviors of the fireflies. Fireflies can handle the multimodal issues naturally. The RSA algorithm depends on the functional behaviors of the crocodiles. The major motivation for this existing method is the hunting and encircling behaviors of the crocodiles. The merits of the FA algorithm are it can perform.the parallel implementation. It can be utilized in the constraint handling. The benefit of the RSA algorithm is it can perform easily. It has a fast response and has smooth converged curves. However, the FA algorithm has less efficiency and is slow. Also, the RSA algorithm is a miss self-learning technique; there is a change in the solution technique because of the fitness value updating. To tackle the aforementioned problems, a new algorithm has been established. In our proposed model executed based on the iteration counts If $u < \mathrm{Max}_{\mathrm{itr}}/4$ then the FA algorithm is executed. Or else, the RSA algorithm is executed. Figure 5 displays the flowchart for the developed algorithm.

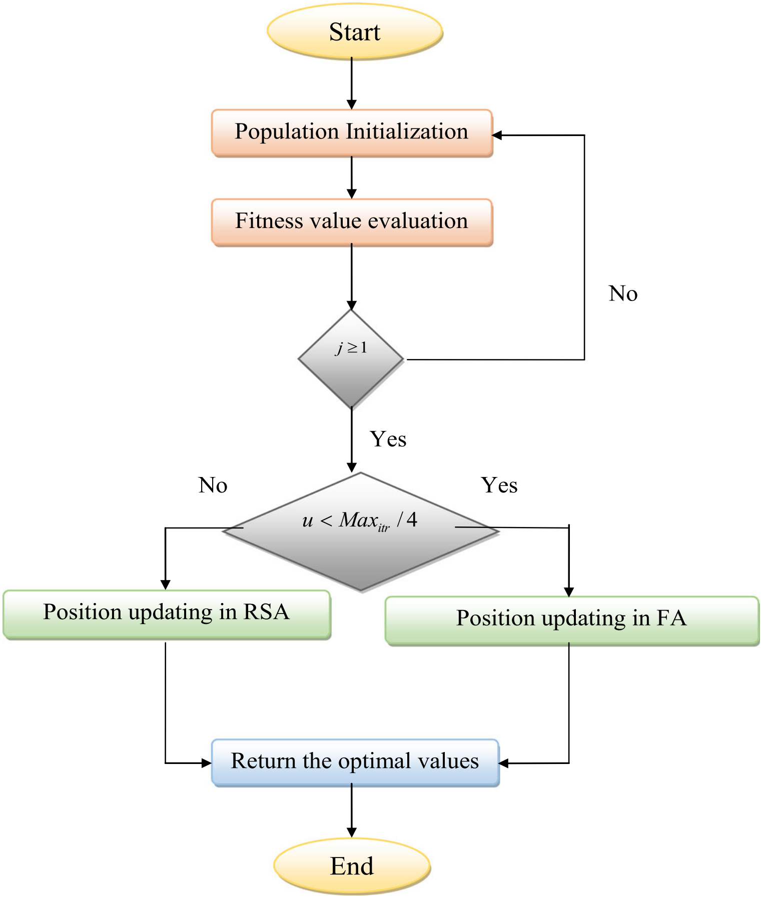
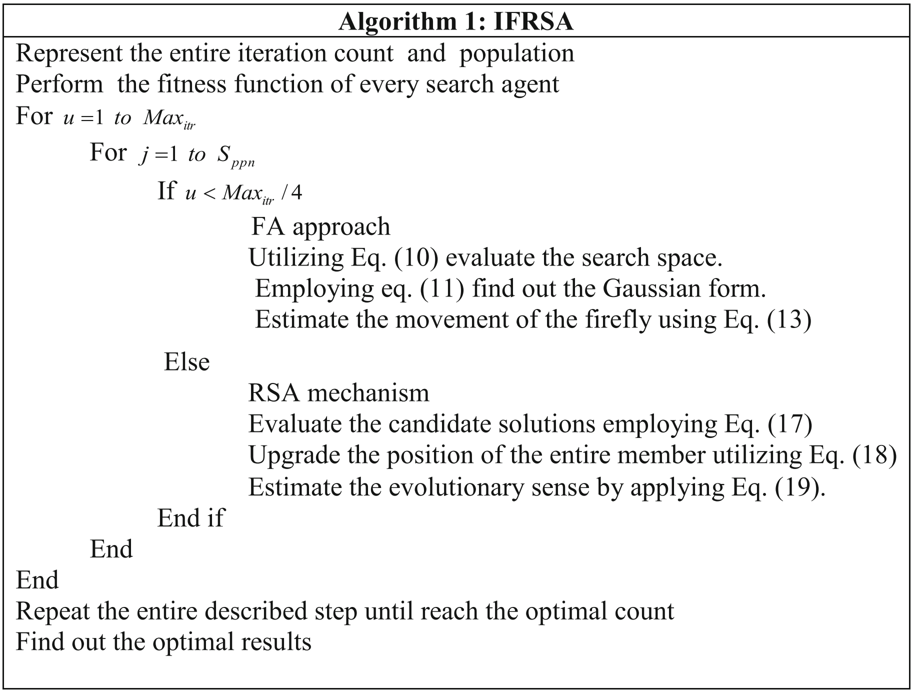


Fig. 5 Flowchart for the improved IFRSA algorithm



Algorithm 1: IFRSA

```
Represent the entire iteration count and population
Perform the fitness function of every search agent
For $u = 1$ to $Max_{itr}$
        For $j = 1$ to $S_{ppn}$
                If $u < Max_{itr} / 4$
                        FA approach
                        Utilizing Eq. (10) evaluate the search space.
                        Employing eq. (11) find out the Gaussian form.
                        Estimate the movement of the firefly using Eq. (13)
                Else
                        RSA mechanism
                        Evaluate the candidate solutions employing Eq. (17)
                        Upgrade the position of the entire member utilizing Eq. (18)
                        Estimate the evolutionary sense by applying Eq. (19).
                End if
        End
End
Repeat the entire described step until reach the optimal count
Find out the optimal results
```

5.4 Multi-objective function with its constraints

The objective function of the suggested method is formulated here. The key to the privacy-preserving data is optimized with the help of the proposed IFRSA algorithm. The objective function Obf is given in Eq. (23).

$$ogj = \underset{\{K^d\}}{\arg\min}\,[FF] \tag{23}$$

The terms w_1, w_2, and w_3 are the weights of the data which are having the values such as 0.001, 0.3, and 0.6 accordingly.

$$FF = (w_1 * D) + \left(w_2 * \frac{1}{\mathrm{hr}}\right) + \left(w_3 * \frac{1}{\mathrm{pr}}\right) \tag{24}$$

The variable K^d is referred for the optimal data key which limits from 0 to 1. The parameter HR denotes the hidden ratio and the term PR represented the preservation ratio. The term D denotes the Euclidean distance.

To find out the objective function the Euclidean distance is measured. The distance between the original information and the transferred information is employed to measure the Euclidean distance D which is expressed in Eq. (25)

$$D = \sqrt{\sum\nolimits_{j=1}^{m} ((E_t) - (E_i))^2} \tag{25}$$

Here, the parameter E_t is denoted as the transferred data and E_i is indicated as the original data. The objective function Obf is given in Eq. (25). The mathematical expression of two ratios is expressed here.

Hidden ratio HR: "It is explained as the rate of sensitive items that are correctly hidden in DS'." It is mathematically expressed in Eq. (26).

$$\mathrm{HR} = \frac{\mathrm{len}_1}{U_i} \tag{26}$$

The parameter len_1 has denoted the length of the nonzero indexes. Then the term U_i represented as the total count of indexes which has to be hidden.

Preservation ratio PR: "It is defined as the rate of non-sensitive rules not hiding in DS'. This is the reciprocal of the hidden ratio". The preservation ratio is shown in Eq. (27).

$$\mathrm{PR} = \frac{\mathrm{len}_2}{U_q} \tag{27}$$

The variable len_2 is denoted as the count of the zero indexes and the term U_q is indicated as the total number of the data indexes that have to be preserved.

Figure 6 depicts the solution encoding for the developed model.

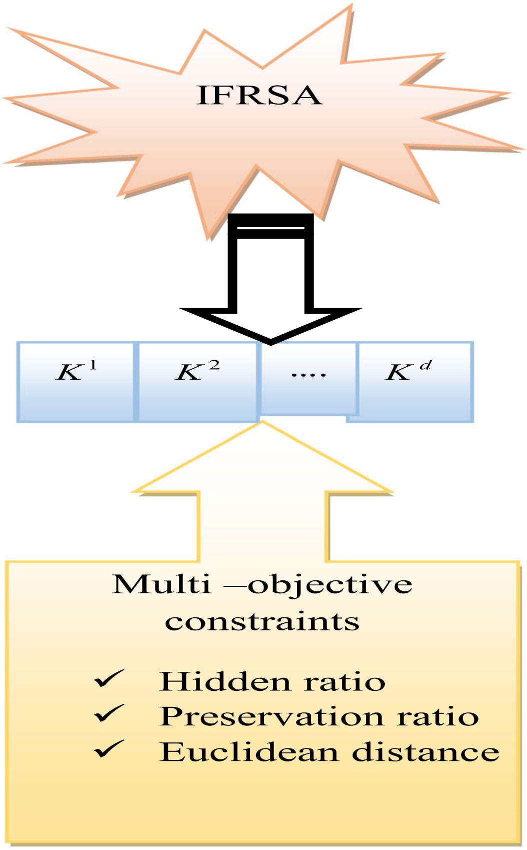


Fig. 6 Diagram of solution encoding for the suggested model

6 Results and discussions

6.1 Experimental setup

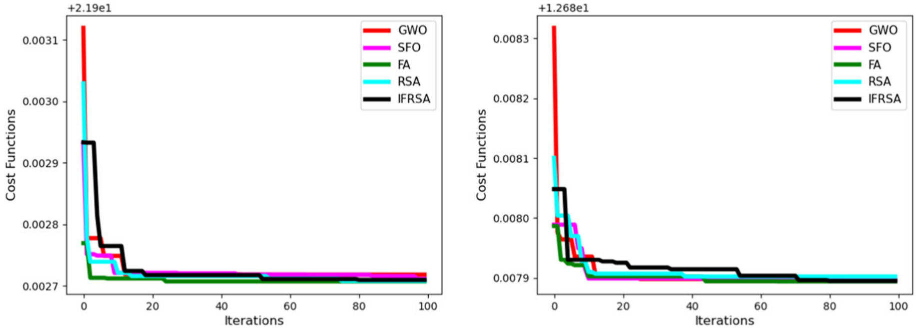
The suggested IFRSA system was implemented in the Python phase and the required results were gathered. The improved algorithm executed in 100 highest iteration counts and also the population count was 10. The chromosome length was measured based on the data size. Multiple calculations were done to evaluate the proposed system. Hence, the compared algorithms such as Grey Wolf Optimization (GWO) [30], Sun Flower Optimization (SFO) [31], FA [28], and RSA [29] were considered accordingly.

6.2 Performance measures

(a) *Euclidean distance:* It is measured in Eq. (25).
(b) *Hidden ratio:* It is estimated in Eq. (26).
(c) *Preservation ratio:* It is calculated in Eq. (27).
(d) *Known-plaintext attack (KPA):* "The known-plaintext attack is an attack model for the cryptanalysis where the attacker has access to both the plaintext and its encrypted version".
(e) *Chosen-plaintext attack (CPA):* "In Chosen Plaintext Attack is attacker selects the random plaintext to be encrypted and obtains the corresponding plaintext".
(f) *Restoration efficiency:* "It is defined as a function of the error amount measured by the distance between the aberrant shape and its original version".

6.3 Cost function estimation of the suggested model for the two datasets

Figure 7 shows the cost function estimation of the developed system contrasted with the distinct algorithms. Figure 7b depicts the cost function for the second dataset. When considering the value is 40 for the dataset 2, the cost function of the developed system was reduced to 98.02% of GWO, 98.05% of SFO, 98.07% of SA, and finally 97.97% of RSA. This result



(a)

(b)

Fig. 7 Estimation of the cost functions for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms in terms of **a** dataset 1 and **b** dataset 2

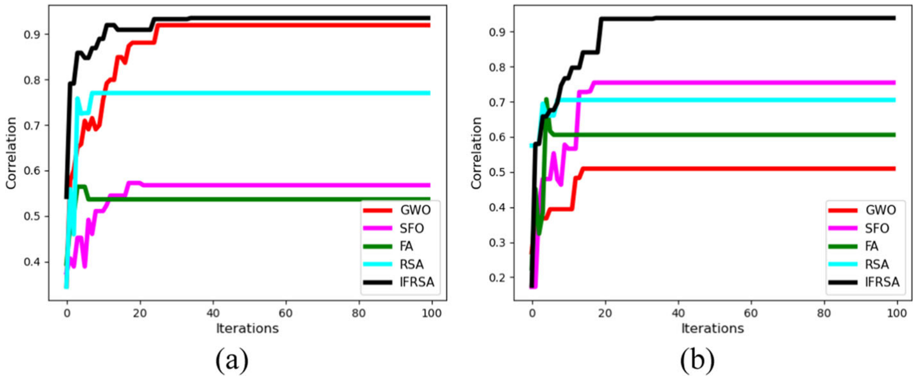


Fig. 8 Measurement of the correlation for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms in terms of **a** dataset 1 and **b** dataset 2

explains the less cost of the suggested model which increases the availability of the developed model.

6.4 Correlation measurement of the developed system for the two datasets

Figure 8 displays the correlation measurement of the developed system contrasted with the various algorithms. Figure 8a depicts the correlation for the first dataset. When considering the value is 20 for dataset 1, the correlation of the suggested system improved by more than 95.5% of GWO, 97.15% of SFO, 97.35% of SA, and finally 96.15% of RSA. This result explains the better system functionality of the developed model.

6.5 Analysis of the CPA attack for the recommended model in two datasets

Figure 9 depicts the CPA attack estimation of the developed system contrasted with the various models. When analyzing the CPA attack for the developed model, it preserves the

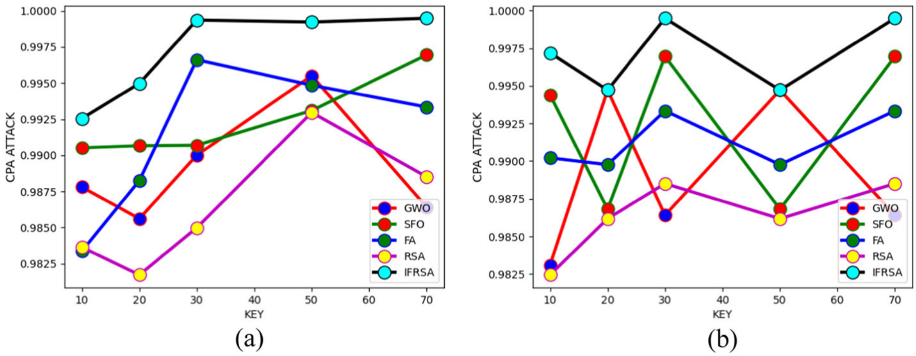


Fig. 9 Analysis of CPA attack for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms concerning **a** dataset 1 and **b** dataset 2

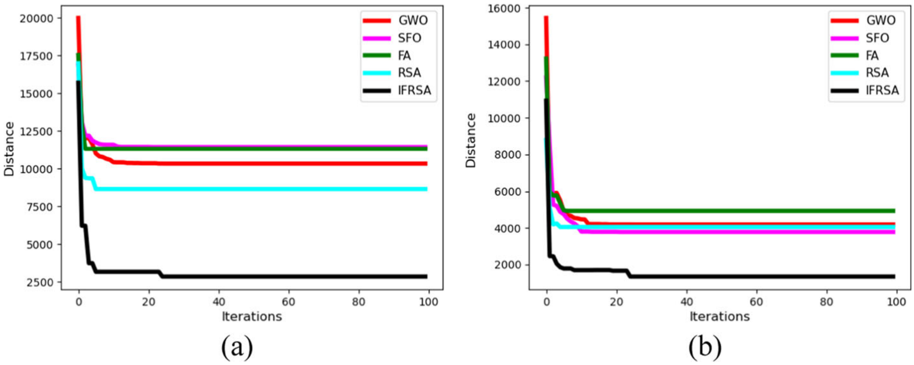


Fig. 10 Measurement of the distance for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms concerning **a** dataset 1 and **b** dataset 2

data with more privacy rather than other existing approaches. Thus, the proposed work acquires promising performance for managing EMR data over the blockchain.

6.6 Distance measurement of the developed model for the two datasets

Figure 10 shows the distance measurement of the developed system contrasted with the various algorithms. Figure 10b explains the measurement of the distance for the second dataset. When noticing the iteration value is 20 for the dataset 2, the distance measurement of the recommended model is enhanced by 20.4% of GWO, 19.9% of SFO, 28.9% of SA, and finally 20.15% of RSA. This result explains the better measurement of the distance for the developed model.

6.7 Euclidean distance measurement of the recommended system for the two datasets

Figure 11 displays the Euclidean distance measurement of the developed system contrasted with the various algorithms. Figure 11a explains the Euclidean measurement of the distance for the first dataset. When noticing the value is 60 for the dataset 1, the Euclidean distance measurement of the recommended model enhanced than 99.36% of GWO, 99.66% of SFO, 99.25% of SA, and finally 99.33% of RSA. This result explains the better distance measurement in the Euclidean system for the developed model.

6.8 Hiding ratio estimation of the suggested model for the two datasets

Figure 12 describes the evaluation of the hiding ratio of the developed system contrasted with the various algorithms. Figure 12a describes the evaluation of the hiding ratio for the first dataset. When noticing the value is 20 for the dataset 1, the hiding ratio of the recommended system is enhanced by 98.6% of GWO, 99.25% of SFO, 98.3% of SA, and finally 98.35% of RSA. This result explains the better ratio estimation for the developed model which improves the system reliability.

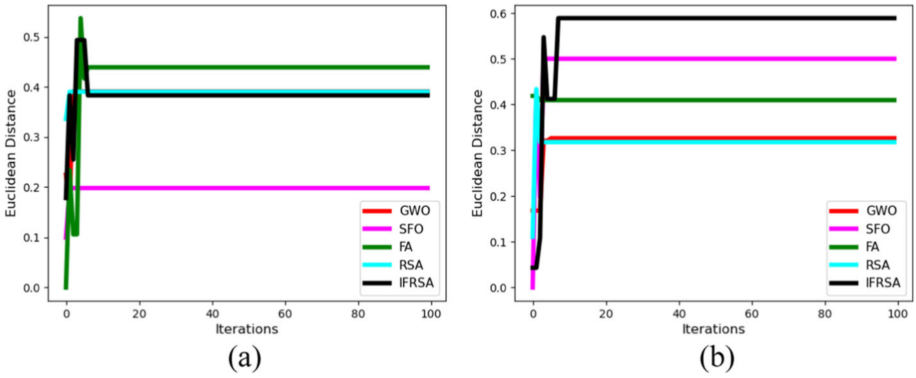


Fig. 11 Euclidean measurement of the distance for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms regarding **a** dataset 1 and **b** dataset 2

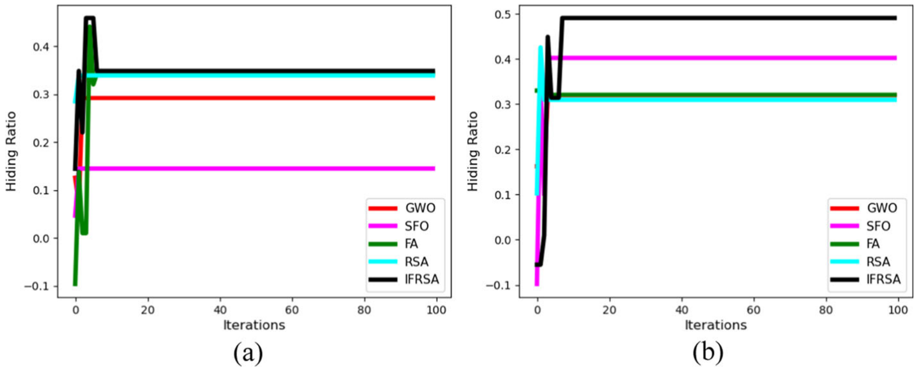


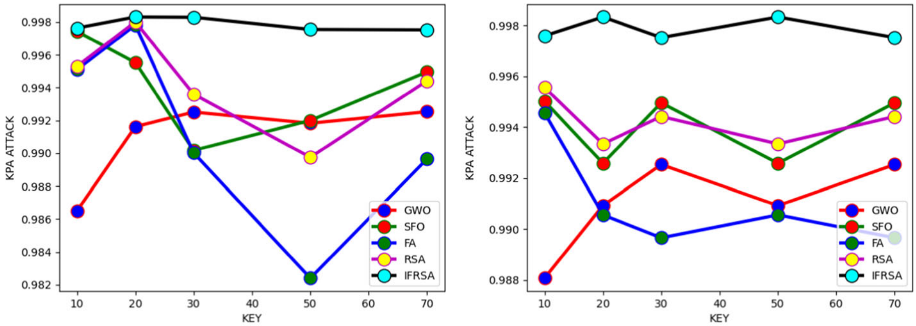
Fig. 12 Estimation of hiding ratio for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms regarding **a** dataset 1 and **b** dataset 2

6.9 Analysis of the KPA attack for the suggested system in two datasets

Figure 13 depicts the analysis of the KPA attack for the developed system contrasted with the various algorithms. When analyzing the KPA attack for the improved model, it preserves the data with more privacy rather than other existing approaches. Thus, the suggested work acquires promising performance for managing the EMR data over the blockchain.

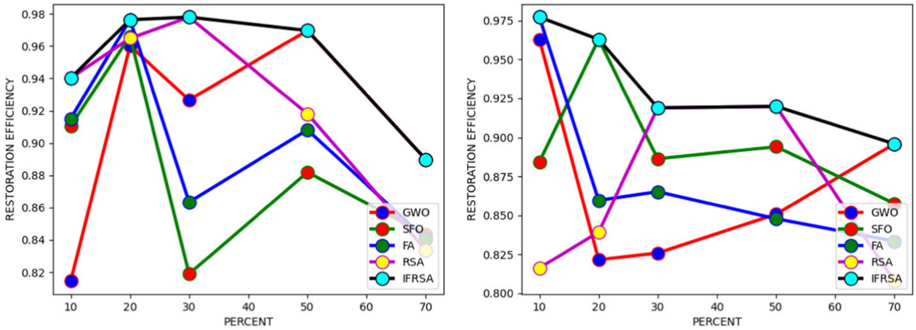
6.10 Measurement of the restoration efficiency of the improved system for the two datasets

Figure 14 expressed the measurement of the restoration efficiency for the developed system contrasted with the various algorithms. Figure 14b explains the analysis of the restoration efficiency for the second dataset. When noticing the value is 50 for the dataset 2, the restoration efficiency of the recommended model enhanced than 98.29% of GWO, 98.19% of SFO,



(a) (b)

Fig. 13 KPA attack analysis for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms regarding **a** dataset 1 and **b** dataset 2



(a) (b)

Fig. 14 Measurement of the restoration efficiency for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms regarding **a** dataset 1 and **b** dataset 2

98.29% of SA, and finally 98.15% of RSA. This result explains the better restoration for the developed model which enhances the efficiency.

6.11 Analysis of the preservation ratio for the suggested model for the two datasets

Figure 15 expressed the preservation ratio analysis for the developed system contrasted with the various algorithms. Figure 15a explains the analysis of the preservation ratio for the first dataset. When noticing the value is 20 for dataset 1, the preservation ratio of the recommended system enhanced than 97.05% of GWO, 96.1% of SFO, 96.95% of SA, and finally 97% of RSA. This result explains the better ratio analysis for the developed model which enhances the system performance.

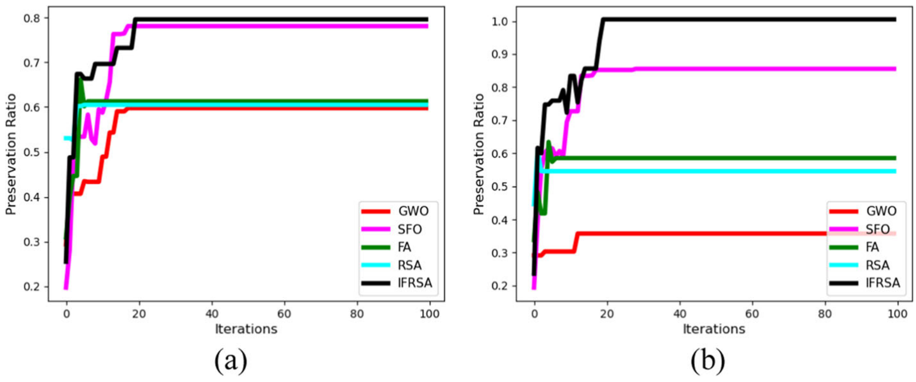


Fig. 15 Preservation ratio analysis for the recommended Merkle tree-blockchain model for privacy preservation of data contrasted with the existing various algorithms regarding **a** dataset 1 and **b** dataset 2

6.12 Overall statistical evaluation of the recommended model over distinct algorithms

The assimilative evaluation of the recommended system is compared with the traditional systems of distinct algorithms that are displayed in Table 2. The suggested IFRSA system gives the successful results when compared to the other conventional algorithms. According to dataset 1, the suggested IFRSA algorithm has 31, 31, 31, and 31% over GWO, SFO, FA, and RSA methods employing the best statistics. This explains the enhanced functionality of the developed system.

Table 2 Overall statistical analysis of the recommended medical data protection over diverse algorithms for two datasets

Statistics	GWO [30]	SFO [31]	FA [28]	RSA [29]	IFRSA
Dataset 1					
Best	21.90271	21.90271	21.90271	21.90271	21.90263
Worst	21.90272	21.90271	21.90272	21.90271	21.90271
Mean	21.90271	21.90271	21.90271	21.90271	21.90271
Median	21.90271	21.90271	21.90271	21.90271	21.90271
Standard deviation	3.92E–06	1.08E–06	3.93E–06	4.61E–07	1.66E–06
Dataset 2					
Best	12.68789	12.68789	12.6879	12.68789	12.68781
Worst	12.6879	12.6879	12.6879	12.6879	12.68791
Mean	12.6879	12.6879	12.6879	12.6879	12.6879
Median	12.6879	12.6879	12.6879	12.6879	12.6879
Standard deviation	2.89E–06	3.43E–06	2.21E–06	3.47E–06	5.12E–06

7 Conclusion

This paperwork has implemented a Merkle tree-blockchain employed privacy preservation of EMR on medical data. In this method, the Merkle tree was utilized which provided secure and efficient verification of multiple data structures. This method utilized two main modules such as data restoration and data sanitization. The data sanitization method was depended on the creation of the optimal key employed by a suggested IFRSA algorithm. The creation of the optimal key was established to derived various-objective functionality which contained the parameters such as hiding rate, Euclidean distance, and data preservation rate among the encrypted data employed the original key and the varied key. At last, the developed electrical medical system became more reliable, useful, and secure. Then the compared outcomes showed that the supremacy of the improved system. By employing dataset 1, the cost function of the developed model was reduced to 99.92% of GWO, 99.95% of SFO, 99.93% of SA, and finally 99.98% of RSA correspondingly. Therefore, the results showed that the recommended IFRSA system has performed more successfully than the existing systems. The system has executed slower because of the Merkle tree. This problem will be resolved in future research applying the modern technologies. The major fault of this work is the verification process. Here, we need more concentration on the verification process, which will help to improve the effectiveness of the system. In the future, the verification acknowledgment scheme will be added to the designed method, which will help to reveal the patient consent and control over their data. It helps to provide more security for the healthcare data, and also we will try to implement our designed model into any one of the real-time applications in the upcoming work.

Author contributions All authors have made substantial contributions to conception and design, revising the manuscript, and the final approval of the version to be published. Also, all authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Yongjoh S, So-In C, Kompunt P, Muneesawang P, Morien RI (2021) Development of an internet-of-healthcare system using blockchain. IEEE Access 9:113017–113031
2. Mamun AA, Azam S, Gritti C (2022) Blockchain-based electronic health records management: a comprehensive review and future research direction. IEEE Access 10:5768–5789
3. Yang X, Li T, Xi W, Chen A, Wang C (2020) A blockchain-assisted verifiable outsourced attribute-based signcryption scheme for EHRs sharing in the cloud. IEEE Access 8:170713–170731
4. Tan L, Yu K, Shi N, Yang C, Wei W, Lu H (2022) Towards secure and privacy-preserving data sharing for COVID-19 medical records: a blockchain-empowered approach. IEEE Trans Netw Sci Eng 9(1):271–281
5. Rajput AR, Li Q, Taleby Ahvanooey M, Masood I (2019) EACMS: emergency access control management system for personal health record based on blockchain. IEEE Access 7:84304–84317
6. Akkaoui R, Hei X, Cheng W (2020) EdgeMediChain: a hybrid edge blockchain-based framework for health data exchange. IEEE Access 8:113467–113486
7. Zarour M, Ansari MTJ, Alenezi M, Sarkar AK, Faizan M, Agrawal A (2020) Evaluating the impact of blockchain models for secure and trustworthy electronic healthcare records. IEEE Access 8:157959–157973
8. Zhang Heyan (2020) Secure routing protocol using salp-particle swarm optimization algorithm. J Netw Commun Syst 3(3):1–10

9. Ray PP, Chowhan B, Kumar N, Almogren A (2021) BIoTHR: electronic health record servicing scheme in IoT-blockchain ecosystem. IEEE Internet Things J 8(13):10857–10872
10. Tang F, Ma S, Xiang Y, Lin C (2019) An efficient authentication scheme for blockchain-based electronic health records. IEEE Access 7:41678–41689
11. Madine MM, Battah AA, Yaqoob I, Salah K, Jayaraman R, Al-Hammadi Y (2020) Blockchain for giving patients control over their medical records. IEEE Access 8:193102–193115
12. Wang Y, Zhang A, Zhang P, Wang H (2019) Cloud-assisted EHR sharing with security and privacy preservation via consortium blockchain. IEEE Access 7:136704–136719
13. Shahnaz A, Qamar U, Khalid A (2019) Using blockchain for electronic health records. IEEE Access 7:147782–147795
14. Stafford TF, Treiblmaier H (2020) Characteristics of a blockchain ecosystem for secure and sharable electronic medical records. IEEE Trans Eng Manag 67(4):1340–1362
15. Zhou X, Liu J, Wu Q, Zhang Z (2018) Privacy preservation for outsourced medical data with flexible access control. IEEE Access 6:14827–14841
16. Deebak BD, Al-Turjman F, Aloqaily M, Alfandi O (2019) An authentic-based privacy preservation protocol for smart e-healthcare systems in IoT. IEEE Access 7:135632–135649
17. Sahi MA, Abbas H, Saleem K, Yang X, Derhab A, Orgun MA et al (2018) Privacy preservation in e-healthcare environments: state of the art and future directions. IEEE Access 6:464–478
18. Liu X, Wang Z, Jin C, Li F, Li G (2019) A blockchain-based medical data sharing and protection scheme. IEEE Access 7:118943–118953
19. Sammeta N, Parthiban L (2021) Hyperledger blockchain enabled secure medical record management with deep learning-based diagnosis model. Complex Intell Syst 8:625–640
20. Zhu H, Guo Y, Zhang L (2021) An improved convolution Merkle tree-based blockchain electronic medical record secure storage scheme. J Inf Secur Appl 61:102952
21. Huang H, Zhu P, Xiao F, Sun X, Huan Q (2020) A blockchain-based scheme for privacy-preserving and secure sharing of medical data. Comput Secur 99:102010
22. Iqbal N, Jamil F, Ahmad S, Kim D (2021) A novel blockchain-based integrity and reliable veterinary clinic information management system using predictive analytics for provisioning of quality health services. IEEE Access 9:8069–8098
23. Li C-T, Shih D-H, Wang C-C, Chen C-L, Lee C-C (2020) A blockchain-based data aggregation and group authentication scheme for electronic medical system. IEEE Access 8:173904–173917
24. Nguyen DC, Pathirana PN, Ding M, Seneviratne A (2019) Blockchain for secure EHRs sharing of mobile cloud based E-health systems. IEEE Access 7:66792–66806
25. Ji Y, Zhang J, Ma J, Yang C, Yao X (2018) BMPLS: blockchain-based multi-level privacy-preserving location sharing scheme for telecare medical information systems. J Med Syst 42:147
26. Rathee G, Sharma A, Saini H, Kumar R, Iqbal R (2020) A hybrid framework for multimedia data processing in IoT-healthcare using blockchain technology. Multimed Tools Appl 79:9711–9733
27. Ahamad D, Hameed SA, Akhtar M (2020) A multi-objective privacy preservation model for cloud security using hybrid Jaya-based shark smell optimization. J King Saud Univ Comput Inf Sci 34(6):2343–2358
28. Gandomi AH, Yang X-S, Talatahari S, Alavi AH (2013) Firefly algorithm with chaos. Commun Nonlinear Sci Numer Simulat 18:89–98
29. Abualigah L, Abd Elaziz M, Sumari P, Geem ZW, Gandomi AH (2022) Reptile search algorithm (RSA): a nature-inspired meta-heuristic optimizer. Expert Syst Appl 191:116158
30. Naserbegi A, Aghaie M, Zolfaghari A (2020) Implementation of grey wolf optimization (GWO) algorithm to multi-objective loading pattern optimization of a PWR reactor. Ann Nucl Energy 148(107703):1
31. Navaneetha Krishnan S, Yuvaraj D, Banerjee K, Josephson PJ, Kumar T, Ayoobkhan MUA (2022) Medical image enhancement in health care applications using modified sun flower optimization. Optik 271:170051
32. Zhang X, Xi P, Liu W, Peng S (2022) EMRShareChain: a privacy-preserving EMR sharing system model based on the consortium blockchain. International symposium on bioinformatics research and applications. Springer, Cham, pp 343–355
33. Sharma P, Namasudra S, Crespo RG, Parra-Fuente J, Trivedi MC (2023) EHDHE: enhancing security of healthcare documents in IoT-enabled digital healthcare ecosystems using blockchain. Inf Sci 629:703–718
34. Cerchione R, Centobelli P, Riccio E, Abbate S, Oropallo E (2023) Blockchain's coming to hospital to digitalize healthcare services: Designing a distributed electronic health record ecosystem. Technovation 120:102480
35. Sameera KM, Vinod P, Rehiman KR, Jifhna PN, Sebastian S (2022) Blockchain federated learning framework for privacy-preservation. International conference on advancements in smart computing and information security. Springer, Cham, pp 250–261

36. Ndzimakhwe M, Telukdarie A, Munien I, Vermeulen A, Chude-Okonkwo UK, Philbin SP (2023) A framework for user-focused electronic health record system leveraging hyperledger fabric. Information 14(1):51

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

M. Lakshmanan received his B.E. in Computer Science and Engineering from Anna University, Chennai, in 2011. He completed his Master of Engineering in Software Engineering in 2013 from Anna University Regional Centre, Coimbatore. Currently, he is pursuing a PhD at Anna University, Chennai, Tamil Nadu, India. Presently, he serves as an Assistant Professor in the Department of Computer Science and Engineering at Sri Venkateswara College of Engineering, Tamil Nadu, India. He has authored more than 25 research articles covering various domains. His primary research interests focus on Blockchain Technologies, Cryptography, Cloud Computing, and Adhoc Networks.

G. S. Anandha Mala received B.E. degree in Computer Science and Engineering from Bharathidhasan University in 1992, M.E. degree from the University of Madras in 2001, and PhD degree from Anna University in 2007. Currently, she serves as a Professor at Easwari Engineering College in Chennai, India, and holds the position of the head of the Department of Computer Science and Engineering. With an extensive teaching experience of over 25 years at both graduate and post-graduate levels, she has authored 70 technical papers published in various international journals and conferences. Her area of interest includes Natural Language Processing, Software Engineering, Data Mining, Image Processing, and Grid Computing.