Intelligent Automation & Soft Computing
DOI: 10.32604/iasc.2023.025719
Article

Tech Science Press

Novel Block Chain Technique for Data Privacy and Access Anonymity in Smart Healthcare

J. Priya* and C. Palanisamy

Department of Information Technology, Bannari Amman Institute of Technology, Erode, 638401, Tamilnadu, India
*Corresponding Author: J. Priya. Email: priyaajothimani@gmail.com
Received: 02 December 2021; Accepted: 27 January 2022

Abstract: The Internet of Things (IoT) and Cloud computing are gaining popularity due to their numerous advantages, including the efficient utilization of internet and computing resources. In recent years, many more IoT applications have been extensively used. For instance, Healthcare applications execute computations utilizing the user's private data stored on cloud servers. However, the main obstacles faced by the extensive acceptance and usage of these emerging technologies are security and privacy. Moreover, many healthcare data management system applications have emerged, offering solutions for distinct circumstances. But still, the existing system has issues with specific security issues, privacy-preserving rate, information loss, etc. Hence, the overall system performance is reduced significantly. A unique blockchain-based technique is proposed to improve anonymity in terms of data access and data privacy to overcome the above-mentioned issues. Initially, the registration phase is done for the device and the user. After that, the Geo-Location and IP Address values collected during registration are converted into Hash values using Adler 32 hashing algorithm, and the private and public keys are generated using the key generation centre. Then the authentication is performed through login. The user then submits a request to the blockchain server, which redirects the request to the associated IoT device in order to obtain the sensed IoT data. The detected data is anonymized in the device and stored in the cloud server using the Linear Scaling based Rider Optimization algorithm with integrated KL Anonymity (LSR-KLA) approach. After that, the Timestamp-based Public and Private Key Schnorr Signature (TSPP-SS) mechanism is used to permit the authorized user to access the data, and the blockchain server tracks the entire transaction. The experimental findings showed that the proposed LSR-KLA and TSPP-SS technique provides better performance in terms of higher privacy-preserving rate, lower information loss, execution time, and Central Processing Unit (CPU) usage than the existing techniques. Thus, the proposed method allows for better data privacy in the smart healthcare network.

Keywords: Adler 32 hashing algorithm; linear scaling based rider optimization algorithm with integrated KL anonymity (LSR-KLA); timestamp-based public and private key schnorr signature (TSPP-SS); blockchain; internet of things (IoT); healthcare

1 Introduction

The application and development of healthcare information system management systems have typically been hampered by privacy and security concerns [1]. The healthcare sector is a data-intensive enterprise that generates, distributes, sorts, and accesses huge amounts of data every second [2]. IoT-based smart medical applications have now developed to the extent that medical services are automatically provided to all users [3]. Because such IoT devices gather and analyze a lot of sensitive data, information security and privacy is among the most important requirement [4]. Patients' personal information is normally kept on a server and processed remotely to protect their privacy. As a result of this, patients' are concerned about the privacy and confidentiality of their information [5]. This is because medical data is constantly subject to various security concerns such as information leakage, data breaches, unauthorized tampering, and other issues [6]. Personal data, end-to-end security, and privacy are the top concerns when sharing healthcare data between users. This is significant, mainly because the technology demands additional infrastructures such as the cloud, Internet of Things, mobile devices, etc [7].

For scalability, flexibility, and cost considerations, some cloud-based healthcare data exchange systems have been suggested using data encryption and operation anonymization. However, consumers are often cautious about sending their private and sensitive data to the cloud due to the potential consequences [8]. Healthcare data is currently stored in a centralized cloud-based database, and medical records are mostly un-portable in the traditional approach. Centralization raises security concerns and mandates trust in a single authority [9,10]. It can handle the vast storage of medical data and the sharing of patients' medical data via the cloud platform, thereby improving the efficiency and quality of medical services [11,12]. Blockchain technology is well-suited to the maintenance of secure and decentralized networks. It can alter the way data is stored and exchanged [13].

Nowadays, blockchain is used in practically every field of research and development. IoT and blockchain technologies are being heavily explored and deployed in numerous fields, particularly e-healthcare [14]. In the blockchain, transactions are represented as blocks linked to building a chain of blocks. If one block or transaction is compelled to modify, the blockchain's complete chain header information should be modified as well [15]. The parties do not need to build a trust relationship when the transaction is performed and resolved on the ledger; instead, they must trust the blockchain itself to accomplish this task [16]. Many researches have been carried in the health industry to assess the possibilities of blockchain technology in that sector [17]. The majority of them are still dealing with blockchain challenges [18]. To address the challenges, a secure blockchain-based architecture is proposed that ensures authenticity by strengthening data privacy and integrity through LSR-KL Anonymity and the TSPP-SS authorization approach.

The main aim of this research work is the ensure data privacy and access anonymity using novel blockchain-based LSR-KL and TSPP-SS algorithms. Numerous research and methodologies are introduced, but the performance is not ensured significantly. The existing approaches have drawbacks with data privacy and integrity. The LSR-KL and TSPP-SS method is proposed to overcome the above-mentioned issues to improve security and data privacy, improving overall performance. The main contribution of this research is ensuring data privacy and access anonymity. The proposed method is used to provide better results using effective approaches. The remaining sections of the article are organized as follows: Section 2 reviews some relevant works, Section 3 elaborates the proposed architecture, Section 4 describes the experimental results, and Section 5 concludes the paper with future directions.

2 Literature Survey

Omar et al. [19] presented a patient-centric healthcare data management system that stores data are employing blockchain technology, allowing for privacy. To encrypt the patient data and assure authentication, cryptographic functions are deployed. To provide sufficient cryptographic capability to the

users, Elliptic Curve Cryptography (ECC) was employed as the cryptographic instrument. According to an experimental performance evaluation, the platform performed well in a blockchain context. However, the strategy failed to account for data loss.

Xu et al. [20] offered health-chain, a blockchain-based large-scale health data privacy-preserving solution in which health data has been encrypted for fine-grained access control. Users were effectively removed or added as authorized doctors by exploiting user transactions for key management. The information was encrypted and stored in the InterPlanetary File System (IPFS), which minimized communication and computation overhead while maintaining confidentiality. After security research and trial findings, the Health-chain was suitable for a smart healthcare system. The method, however, had concerns with scalability and privacy.

Arava et al. [21] presented a fine-grained k-anonymity algorithm, which uses a systematic procedure of seed selection. This method exhibits a minimum information loss than existing clustering algorithms. Data-sensitive information is a crucial concern of every individual. Hospitals lag their trust in privacy to take up the newest technologies of the cloud-like Information-as-a-service, storage-as-a-service to deploy their patient's data for better health management. An intensive study is being undertaken to run over the shortcomings of data privacy for the published information and the publisher. One of the methods is privacy by statistics using data mining techniques such as k-anonymity. The fundamental approach of k-anonymity is to anonymize sensitive information published that could not be determined from at least (k-1) instances. The best way to attain k-anonymity is by grouping similar records into a cluster by choosing the best seed value to balance utility and privacy in the published data. However, it has issues with optimal attribute selection.

Nagasubramanian et al. [22] developed the keyless signature infrastructure (KSI) to ensure digital signatures' secrecy and authentication features. Furthermore, blockchain technology has been used to maintain data integrity. The KSI used a hash function to give security to the system. The results showed that the system's response time with blockchain technology was nearly half that of traditional solutions. In addition, the authors stated that the method using blockchain had a lower cost of storage than previous systems. However, the technique performed poorly when it came to protecting the privacy of small data.

Prabha et al. [23] presented a suppressed K-Anonymity Multi-Factor Authentication Based SchmidtSamoa Cryptography (SKMA-SC) multifactor authentication solution for privacy-preserving data access in cloud computing. Registration, authentication, and data access were the three important steps in the SKMA-SC approach. Clients registered their personal identity information during the registration phase, and it was securely saved in the cloud server (CS) using the suppression method. The SKMA-SC technique authenticated clients' identities during the authentication phase. It allowed customers to access requested data services during the data access phase by executing authorization using the Schmidt-Samoa data encryption or decryption procedure. According to the results, the SKMA-SC technique improved the privacy-preserving rate (PPR) and reduced authentication's computational complexity (CC). The centralized storage of patient-sensitive data resulted in security and privacy concerns.

Deebak et al. [24] introduced a seamless, secured anonymous authentication scheme (S-SAAS) to establish a secure session in cloud-based mobile edge computing to strengthen the authentication process and eliminate security risks and vulnerabilities. This approach eliminated the problem of clock synchronization to reduce computing costs. In addition, this protocol used a random integer to thwart potential attacks and meet essential security requirements. Real-time multimedia medical server systems were used to reduce signal transmission and routing overhead. Furthermore, according to the trial results, the S-SAAS approach used less signaling congestion and routing control overhead to achieve a higher service connection rate. The procedure was inefficient and resulted in a single point of failure. However, the existing methods have security issues such as privacy-preserving rate, information loss, etc. Also, the previous techniques have problems with attackers, which reduces the overall cloud performance. A

unique blockchain-based design is proposed to improve the anonymity in terms of data access and data privacy in the Internet of Things (IoT) and cloud computing to address the above-mentioned challenges.

3 Proposed Secure Data Access System

In this work, blockchain-based technique is proposed for improving anonymity in terms of data access and data privacy on the Internet of Things (IoT) and cloud computing. Academics and practitioners' interest in Internet-based Computing has shifted in recent years due to the rapid development of new and more efficient computing methods. The healthcare sector has seen great expansion as a result of the emergence of technology such as the Internet of Things (IoT) and the Cloud. It has also attained a significant level of automation. Modern healthcare, which is based on the Internet, faces substantial issues in terms of security and privacy. Managing access control and privacy protection for healthcare data is still one of the greatest barriers to cloud computing's more comprehensive implementation. These services are used by many people and are often accessed through a public network like the Internet. Data must be secured in terms of privacy and integrity, and access control in such an environment. Maintaining control over these security measures is one of the fundamental concerns. The traditional authentication method confirms and validates a user's identification before accessing the system with a password. However, because the client's sensitive information is accessible to attackers in the cloud, the verification performance of existing methodologies was not up to par. As a result, the user access authentication process must be as efficient as feasible, with minimal computational complexity. Therefore, in this paper, novel methodologies are proposed for enhancing data privacy and integrity using LSR-KL Anonymity and TSPP-SS authorization technique based on blockchain to ensure authenticity, thereby improving secured communication in the Internet of Things. The proposed work enhances the data privacy and communication between the user and device using the LSR-KLA techniques and Blockchain server. It performs efficient authentication using the TSPP-SS technique. The Block diagram of the proposed methodology is shown in Fig. 1.

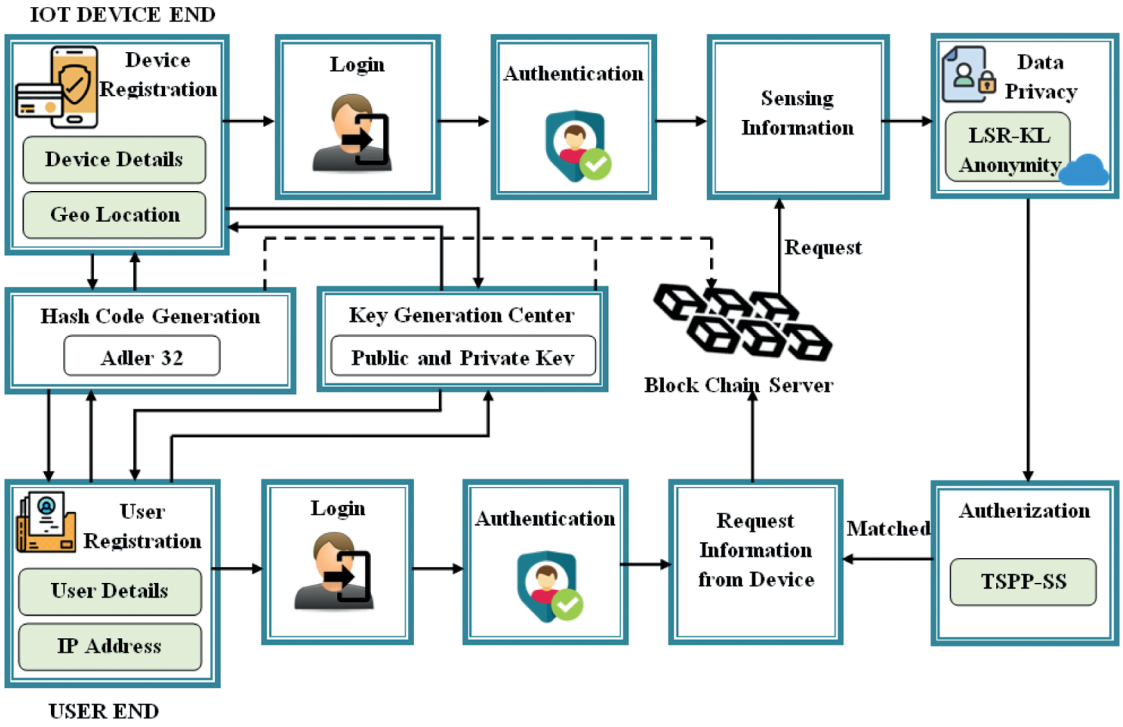


Figure 1: Block diagram of the proposed methodology

3.1 Registration

The proposed work starts with the registration phase, which has been done on both the user end (doctor) and the IoT device end (patient). During device registration, the device details with the corresponding Geo-location are collected. The user's IP address of the user's device and the user details are collected from the user. From the gathered details, the Geolocation of the IoT device and the IP address of the user's device are subjected to the hashcode generation phase. The shortened hash values are generated for the user and device individually. The registration phase contains hash code generation and key generation center which is discussed in the following subsection.

3.1.1 Hash Code Generation

In this section, the values returned by the registration phase are converted to the shortened hash values using the Adler-32 hashing algorithm. The Adler-32 algorithm is also known as the checksum algorithm, in which it computes two 16-bit checksums for the given input values and combines them into a 32-bit integer. Here, the input values for which the hash value is to be calculated are termed as Y_i. It represents the Geolocation on the device end and the IP address on the user end. Then, the checksums CS_1 and CS_2 are calculated as follows,

$$CS_1 = 1 + Y_1 + Y_2 + \ldots\ldots\ldots Y_n(\text{mod}65521) \tag{1}$$

where, CS_1 is the sum of all the bytes in Y_i and initialized to1. CS_2 is the 16-bit value with the initial value of 0 and is calculated as the sum of all individual values of CS_1,

$$\begin{aligned} CS_2 &= (1 + Y_1) + (1 + Y_1 + Y_2) + \ldots\ldots\ldots..(1 + Y_1 + Y_2 + \ldots\ldots Y_n)(\text{mod}65521) = \\ & n \times Y_1 + (n-1) \times Y_2 + (n-2) \times Y_3 + \ldots\ldots\ldots + Y_n + n(\text{mod}65521) \end{aligned} \tag{2}$$

where, Y is the string of the bytes with the length of n. Finally, the hash code is generated as follows,

$$HC(Y) = CS_2 \times 65536 + CS_1 \tag{3}$$

where, $HC(Y)$ does the Adler-32 algorithm produce the hash code. Then, the generated hash codes $HC(Y)$ of the registered user and the device are stored in its database.

Fig. 2a shows the hash code generated after user registration.

(a) (b)

Figure 2: (a) Hash code generation (b) device Login

3.1.2 Key Generation Centre

In addition to the hash codes, the respective public and private keys are generated for the user and deviced from the key generation centre. Key Generation Centre is responsible for generating keys for the

entire groups under the secret shared with each user during registration and sends them to each group member separately. Once the keys are generated they were stored in the database. This ends the registration phase and the Blockchain server monitored the entire registration process.

3.2 Login

After successful registration, the login phase takes place. Logging in is the process by which an individual can get access to a system by identifying and authenticating them. During login, the device and user input their username, password, and hash code on the system. Then, input data is sent to the authentication server to validate the stored values in the database. The process can be expressed as

$$\{Dev,\ Ur\} \xrightarrow{username,pw,HC(Y)} AS \tag{4}$$

where, $\{Dev, Ur\}$ denotes the device and user, AS denotes the authentication server. The device login page shown in Fig. 2b.

3.3 Authentication

Here, the authentication server compares the input data with the data stored in the database. When the match occurs, the system grants the user and device to perform the process such as information sensing, information storage, and data access, etc in a secured manner. It can be expressed as,

$$Username,\ pw,\ HC(Y) \xrightarrow{matched} \{Dev,\ Ur\}_{authenticated} \tag{5}$$

After the authentication process, the user requests the blockchain server to access the sensed IoT data. The blockchain server returns the request to the respective IoT device. Once the IoT device receives the request, the requested information from the IoT device is sent to the user. Before sending, the data anonymization is done to prevent data privacy breach while maintaining the integrity of the data gathered and shared.

3.4 Data Privacy

As the device contains personal and sensitive information, uploading these data causes major security risks such as unauthorized access to the collection, use, or disclosure of information. Therefore, to improve the security, the sensed data ψ_m is converted into anonymized data using Linear Scaling based Rider Optimization algorithm with integrated KL Anonymity (LSR-KLA) technique. K-anonymity is the well-established model used to anonymize data utilizing K-anonymization methods to allow data disclosure in a controlled manner while securing the value integrity of each tuple. In this method, the data is said to be K-anonymity when each attribute value is indistinguishable from one another. In this way, the privacy of the information is increased by increasing the value of K i.e., the greater the K-value, the higher the privacy protection. Suppose the sensitive values for a set of k records that share quasi-identifying variables are the same in the k-anonymity approach. In that case, the data is still open to many threats. To tackle such an issue, the sensitive values within the tuples that share similar values of their quasi-identifiers are further sanitized by diversifying the sensitive information (attributes) using the l-diversity method. And to select the number of quasi-identifiers to be anonymized, the Rider Optimization algorithm (ROA) is used. Thus, the algorithm is termed LSR-KLA. Initially, all the attribute values collected from the sensor ψ_m are represented in a table that contains the Quasi-identifiers. Quasi-identifiers are the set of non-sensitive information (attributes) used to recognize an individual and need to be anonymized. The table p_T that contains the finite set of attributes of each tuple is expressed as,

$$p_T = \{P_1,\ P_2,\ P_3,\ \ldots,\ p_n\} \tag{6}$$

Then, a combination of characteristic information called quasi-identifiers is selected by using the Linear Scaling based ROA algorithm LSA as follows,

Selection of Quasi-Identifiers

ROA is an optimization algorithm inspired based on the idea of rider groups who move toward a common target. The ROA comprises four groups called, Bypass riders, Followers, Overtakers, and Attackers. Bypass riders are the ones who bypass the leading path to reach the destination. Followers depend on the top riders where the overtakers follow their own position [25]. Attackers aim to take the position of the leading rider by following the followers. The improved success rate is attained in ROA using the linear scaling method.

Parameter Initialization: The algorithm is initialized with the four groups of riders which can be represented as, p_T (the table from which the quasi identifiers are to be identified). The total number of riders is determined based on the number of riders belonging to each group. It can be expressed as,

$$P = P_{bp} + P_{flw} + P_{ovt} + P_{att} where,\ P_{bp} = P_{flw} = P_{ovt} = P_{att} = \frac{P}{4} \tag{7}$$

Then, the random position of each rider is initialized as follows,

$$Z_r = \{Z_r(k,\ l)\},\ 1 \le k \le P,\ 1 \le l \le R \tag{8}$$

where, Z_r denotes the position of the k^{th} rider at a time r with the number of coordinates R. Followed by population initialization, the rider's parameters steering, gear, brake, and accelerator are also initialized. Therefore for the parameter steering, the steering angle of the rider's vehicle ($A_{str(r)}$), the Steering angle of the k^{th} rider ($A_{str(r)}(k, l)$), position angle of the k^{th} rider's vehicle (φ_k), and the coordinate angle used for finding the steering angle (Φ) are defined as,

$$A_{str(r)} = \{A_{str(r)}(k,\ l)\},\ 1 \le k \le P,\ 1 \le l \le R \tag{9}$$

$$A_{str(r)}(k,\ l) = \begin{cases} \varphi_k if (l = 1) \\ A_{str(r)}(k,\ l-1) + \Phi if (l \ne 1 \quad A_{str(r)}(k,\ l-1) + \Phi \le 360 \\ A_{str(r)}(k,\ l-1) + \Phi - 360^{\circ} otherwise \end{cases} \tag{10}$$

$$\varphi_k = k * 360^{\circ} * \frac{1}{P} \tag{11}$$

$$\Phi = 360^{\circ} * \frac{1}{R} \tag{12}$$

where, φ_k,Φ calculated based on the maximum angle of 360°. Then, the gear (Gr), accelerator (Acr), and brake (Brk) of the k^{th} rider's vehicle is initialized as,

$$Gr = \{Gr_k\},\ 1 \le k \le P \tag{13}$$

$$Acr = \{Acr_k\},\ 1 \le k \le P \tag{14}$$

$$Brk = \{Brk_k\},\ 1 \le k \le P \tag{15}$$

where, Gr is initially set to zero and it ranges between [0, 4], and the value of Acr and Brk range between [0, 1]. Finally, the maximum speed (Spd) of the rider at which the rider may drive is defined based on two boundary values as,

$$Spd = \frac{1}{r_{off}}\left(Z_{k(\max)} - Z_{k(\min)}\right) \tag{16}$$

where, $Z_{k(\max)}$ is the maximum value of the rider's position, $Z_{k(\max)}$ is the minimum value of the rider's position, r_{off} is the maximum time at the end of the iterations.

Success Rate Determination: The success rate of each rider is calculated when the population groups and parameters have been setup. The success rate is defined as the distance between the rider's current location and the destination. It can be estimated using linear scaling as,

$$sr_k = \frac{Z - Z_k}{Z_{tar} - Z_k} \tag{17}$$

where, sr_k denotes the success rate of the k^{th} rider, Z_{tar} is the position of the target.

Finding the Leading Rider: Once the success rate has been determined, the rider who is closest to the destination is considered to have the highest success rate. Because the leading rider is frequently changed, the success rate is crucial in defining the leading position. As a result, the leading rider can be any rider with the highest success rate.

Position Updation Process: In order to find the success rate of the riders, the leading rider as well as the winner, the position of each rider has to be updated for a certain period of time. The position updation process of each group is as follows,

- ***Position of Bypass Rider:*** Bypass Riders bypass their own way without following the leading riders. Therefore, the position of the bypass riders is updated as,

$$Z^{bp}_{r+1}(k,\ l) = \varpi[Z_r(\delta,\ l) * \gamma(l) + Z_r(\alpha,\ l) * (1 - \gamma(l))] \tag{18}$$

  where, ϖ, γ are the random values ranges between [0, 1], δ, α are the random values ranges between [0, P].

- ***Position of Followers:*** As the followers follow the path of the leading rider they reach the target quickly. The update process regarding the coordinate selector and the position updation is done to the selected values in R as,

$$Z^{flw}_{r+1}(k,\ L) = Z_{lr}(g,\ L) + [\cos(A_{str(r)}(k,\ l) * Z_{lr}(g,\ L) * D^r_k] \tag{19}$$

  where, Z_{lr} is the location of the leading rider g, L is the coordinate selector, D^r_k does the rider travel the k^{th} distance. The distance is calculated by multiplying the velocity of the rider with the rate of off time where the velocity is calculated based on the rider's parameters except steering.

$$D^r_k = \left(\frac{1}{3}[Gr^r_k * spd_{Gr} + Acr^r_k * Spd + (1 - Brk^r_k)Spd]\right) * \frac{1}{r_{off}} \tag{20}$$

  where, spd_{Gr} is the speed limit of the gear, and the term $(\bullet)$ denotes the velocity of the rider.

- ***Position of Overtaker:*** The position of the overtakers is updated based on the three important parameters called direction indicator, relative success rate, and coordinate selector. The position can be updated as,

$$Z^{ovt}_{r+1}(k,\ L) = [Z_r(k,\ L) + DI_r(k) * Z_{lr}(g,\ L)] \tag{21}$$

where, $DI_r(k)$ is the direction indicator of the k^{th} rider and it computes the direction based on the relative success rate.

$$DI_r(k) = \frac{2}{1 - \log\left(\frac{sr_k^r}{\max_{r=1}^{R}(sr_k^r)}\right)} - 1 \tag{22}$$

where, $\log(\bullet)$ denotes the relative success rate of the rider.

- ***Position of Attacker:*** The attackers are the riders who are trying to take the position of leading riders. Therefore the attackers follow the same way of position updation as followers except for the fact that the position updation is done for all the values in the coordinate rather than the selected values. It can be expressed as,

$$Z_{r+1}^{att}(k,\ l) = Z_{lr}(g,\ l) + [\cos(A_{str(r)}(k,\ l)) * Z_{lr}(g,\ l)] + D_k^r \tag{23}$$

After completion of position updation, the success rate of each rider is measured. Based on the success rate the new rider's position is determined i.e., the rider who has been leading the race since then has been replaced by a new rider at the position where the new rider's success rate is highest. The rider's parameters are then modified in order to discover the most effective and ideal solution.

In this way, the Quasi-identifiers associated with the table p_T are identified and are denoted as,

$$q_I = \{P_1,\ P_2,\ P_3,\ \ldots\ldots\ldots..,\ P_K\} \subseteq \{P_1,\ P_2,\ P_3,\ \ldots\ldots\ldots..,\ P_n\} \tag{24}$$

where, q_I is the Quasi-identifier of p_T containing a set of P_K attributes of M tuples selected by the LSR method. In order to anonymize the data, the algorithm uses the distinctive operation called generalization. Generalization is the process of substituting a specific value for a more general one without losing the truthfulness of the data. In generalization, the data becomes less informative with the help of domain generalization hierarchy. The original attribute is considered the ground domain and the domain value is increased with the increase in generalization. Thus, the attributes are mapped to generalize the values as,

$$P_0 \ \xrightarrow{T_0}\ P_1 \ \xrightarrow{T_1}\ \ldots\ldots\ \xrightarrow{T_{n-1}}\ P_n \tag{25}$$

where, P_0 is the minimal element called ground domain, P_n is the maximal element, and T is the function that puts a linear ordering on P. Generalization can be done when the attribute that belongs to a particular tuple is not equal to the attribute of another tuple. After Generalization the K-anonymous table contains the new maximal elements P_T* are obtained and it can be expressed as,

$$p_T* = \{P_1*,\ P_2*,\ P_3*,\ \ldots\ldots\ldots\ldots\ldots,\ p_n*\} \tag{26}$$

After anonymization, the sensitive information that shares similar values is diversified using the l-diversity method. A sensitive property is one whose value for any given individual must be kept hidden from those who do not have direct access to the source data. Thus, the anonymized table P_T* is said to be l-diverse when it contains at least l well-represented values for the sensitive information. It can be expressed as,

$$-\sum_{a\in A}\frac{K_{(P_T*,a)}}{\sum_{a'\in A}K_{(P_T*,a')}}\log\left(\frac{K_{(P_T*,a')}}{\sum_{a\in A}K_{(P_T*,a)}}\right) \geq \log(l) \tag{27}$$

where, $K_{(P_T*,a)}$ denotes the frequencies of each sensitive attribute $a \in A$. In order to achieve diversity, each block in P_T* should have at least $l \geq 2$ different sensitive values and also it needs to be ensured

that the most-frequent values have roughly the same frequency. Such a block is said to be well-represented by sensitive values. Thus, the anonymized data obtained is stored in the cloud server and it can be denoted as $\psi_{m(anonymized)}$.

3.4.1 Authorization

After anonymization, in order to achieve authorization, the digital signature has been created by using the Timestamp-based user Public and device Public Key Schnorr Signature (TSPP-SS) algorithm. Authorization is the process of giving the ability to access a resource. In the Schnorr signature algorithm, the key holder used their private key with a specific message as an input value in order to generate a verifiable digital signature on that specific message. In the proposed work, to improve the complexity of hash value, the combined public keys of both the device and user with the requested time stamp value from the user are considered the message for the schnorr signature algorithm. The algorithm steps to generate the digital signature are as follows,

Step 1: Select the primes x, y and an integer u in such a way that y is the factor of $x-1$ and $u^y = 1 \ mod \ x$. These values are global and can be common to all groups of users.

Step 2: Select a random integer c with $0 < c < y$ and c is known as the user's private key.

Step 3: Calculate the user's public key as,

$$d = u^{-c} \bmod x \tag{28}$$

where, d is the user's public key. Then, the digital signature is created for the user with the private and public keys.

Step 4: Select a random integer v with $0 < v < y$ and compute f as,

$$f = u^v \ mod \ x \tag{29}$$

Step 5: Concatenate the computed value f with the message and the value is hashed and stored in J as,

$$J = Hh(mes\{(\Theta_{Dev}, \ \Theta_{User}\}), \ tsp||f) \tag{30}$$

where, Hh denotes the hash function, mes denotes the message that contains the combined device public key Θ_{Dev}, user public key Θ_{User}, and the timestamp value tsp.

Step 4: Compute G as,

$$\boldsymbol{G = (v + cJ) mod \ y} \tag{31}$$

Thus the digital signature is created and it contains the pair (G, J).

Then, the created digital signatures (G, J), mes, and the requested anonymized data $\psi_{m(anonymized)}$ are sent to the user. On the user side, the digital signature is revised and the hash value is checked as

$$J' = Hh(mes\{(\Theta_{Dev}, \ \Theta_{User}), \ tsp\}||u\hat{G} * d\hat{J} \ mod \ x) \tag{32}$$

If both hash value J and J' is matched, then the system allows the user to access the sensed data. Otherwise, the access is denied. The pseudo-code of the TSPP-SS algorithm is shown in below Fig. 3.

Fig. 3 shows the fundamental steps involved in the TSPP-SS algorithm, which accepts the input message as the combined public key and timestamp and creates the digital signature pair for efficient authentication. During this process, all transactions are monitored by the Blockchain server.

Input: *Combined public key and timestamp*
Output: *Authorized user*

Begin
Initialize x , y , z , c , d
Select a random integer c
Compute $d = u^{-c} \bmod x$
Concatenate message with $f = u^{y} \bmod x$
Store the hashed value in $J = Hh(mes\{(\Theta_{Dev}, \Theta_{User}), tsp\} \| f)$
Compute $G = (v + cJ) \bmod y$
Create the signature pair (G, J)
//verification
Compute $J' = Hh(mes\{(\Theta_{Dev}, \Theta_{User}), tsp\} \| u^{\wedge} G * d^{\wedge} J \bmod x)$
Verify $J = Hh(mes\{(\Theta_{Dev}, \Theta_{User}), tsp\} \| f')$
If $J = Hh(mes\{(\Theta_{Dev}, \Theta_{User}), tsp\} \| f')$
Allows the user
Else
Denies the access
End if
End

Figure 3: Pseudo-code of the TSPP-SS algorithm

3.5 Blockchain Server

Blockchain is a distributed ledger system that enables the recording and distribution of digital data without changing it. In this approach, a blockchain serves as the foundation for immutable ledgers, or transaction records that cannot be changed, deleted, or destroyed [26]. Each transaction on the blockchain is time stamped and grouped in a block. The miners create blocks and they contain information as a block header and transactions. Blocks are data structures that are replicated to all nodes in the network and are used to bundle clusters of transactions. Block metadata includes the following fields: block version, timestamp, previous block hash, Merkel Root, and Nonce. The block version enables you to adhere to a set of block validation requirements. For each block, Timestamp provides a timestamp with the current time. Then in the previous block hash contains the value of 256-bit hash points to the previous block. When a transaction happens in Merkle root, hashing algorithms are employed to encrypt data, which is subsequently delivered to each node. The Merkle tree function generated a final hash value and Merkle tree root since it may include thousands of transaction records. The nonce is a four-byte field that starts at 0 and grows with every hash calculation. The general structure of a blockchain server is shown in Fig. 4,

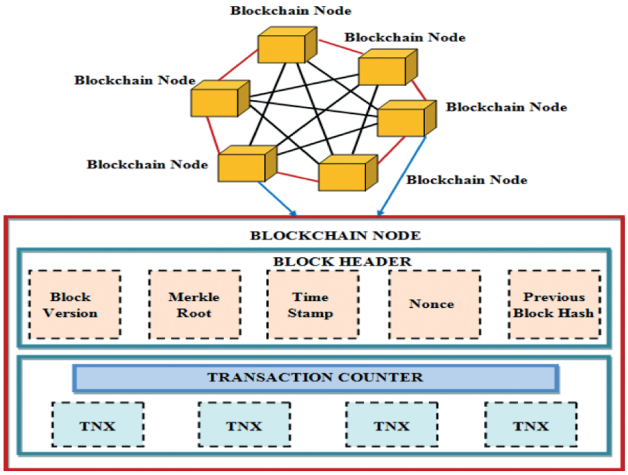


Figure 4: Architecture of blockchain server

From the above Fig. 4, the blocks are data structures that are replicated to all nodes in the network and are used to bundle clusters of transactions. The number of blockchain nodes is formed network, and it is focused on improving data privacy in a a better way. The blockchain server's architecture contains blockchain nodes that consist of the block header and transaction counter. Block header includes block version, merkle root, time stamp, nonce and previous block hash code. In the transaction counter, all transactions are monitored. In the proposed work, the blockchain is used to monitor all transactions. When a device sends some transaction data to the user in the blockchain, that transaction is represented as a block. When a new block is created, it is distributed to all other network nodes to add the block. Miners of the node need to approve the transaction. After the block has been verified and approved, it is inserted into the chain and linked with the previous blocks to complete the transaction. The next step is to decide which user published the next block. As a result, a chain of validated blocks forms the blockchain network. The privacy and integrity of healthcare data have therefore been safeguarded against external attackers and unauthorized access attempts within the network or ecosystem.

4 Results and Discussion

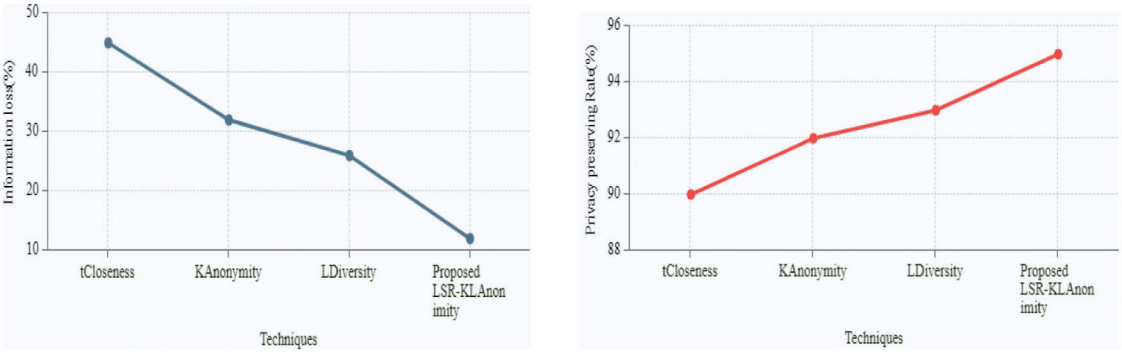
The Proposed Framework for Data Privacy and Access Anonymity in Cloud Computing Platform for the Healthcare sector is validated based on various performance metrics and thereafter compared with the existing methodologies to enhance the highly privileged method. The proposed work is implemented under the software configuration of JavaJDK8, Netbeans8, and under the hardware configuration of inteli5/corei7 processor, 3.20 GHz CPU Speed, Windows 7 operating system and 4GB of Ram.

4.1 Performance Analysis

The proposed LSR-KL Anonymity for privacy preservation are analyzed based on the metrics such as information loss, privacy-preserving rate, execution time, memory usage, feature selection time and fitness *vs.* Iteration and thereafter, compared with the existing techniques. The computational procedure for information loss, privacy preserving rate are calculated based on entropy value and efficiency [27]. The attained results are analyzed in the below section.

Fig. 5a discusses the information loss for the proposed and existing techniques. Information loss is the illicit transfer of data outside organizational boundaries. High information loss indicates insecure data privacy. According to that, the proposed LSR-KLAnonymity tends to achieve an information loss of 12%, whereas the existing methods such as tcloseness, KAnonymity, LDiversity tend to achieve an information loss of 45%, 32%, and 26% respectively. From the graphical observation, it can be stated that the proposed method preserves the data securely as compared to existing methods. Fig. 5b discusses the privacy-preserving rate for the proposed and existing methods. Privacy-preserving rate (PPR) defines the rate at which the data is transferred between the patients and Hospital in order to secure it from the third party who might be attackers or malicious users. PPR value should be maintained high in order to provide a robust securing of data.

According to that, the proposed LSR-KLAnonymity tends to achieve a PPR of 95%, representing a highly secure method compared to the existing method. Among existing methods, tcloseness tends to be the highly insecure method that obtains a PPR of 90% compared to KAnonymity and LDiversity which achieve a PPR of 92% and 93%, respectively. Tab. 1 shows a comparative analysis on values of execution time and CPU usage of tcloseness, KAnonymity, Ldiversity and the proposed LSR-KLAnonymity techniques. Based on number of iterations executed, execution time and CPU usage are calculated. For an instance, consider 20 iterations then,



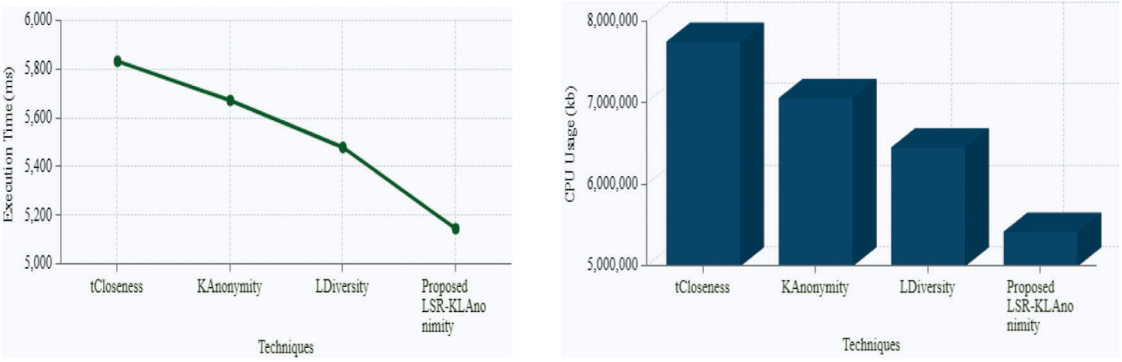
(a) (b)

Figure 5: Graphical demonstration of proposed privacy technique based on (a) information loss (b) privacy-preserving rate

Table 1: Comparative analysis on execution time and CPU usage

Metric/techniques	tcloseness	KAnonymity	Ldiversity	Proposed LSR-KL anonymity
Execution time (ms)	5880.12	5720.25	5490.23	5180.98
CPU usage (kb)	7789938.37	6999870.89	6345904.01	5520300.20

Fig. 6a discusses the execution time. Execution time is the time taken to transfer data from one end to other. High execution time illustrates a high probability of attacks happening. According to that, the proposed method achieves an execution time of 5147 ms which is relatively better as compared to the existing methods that range between 5481–5835 ms. Among the existing methods, tCloseness tends to take more execution time as compared to KAnonymity and LDiversity. The result concluded that the proposed LSR-KL anonymity method provides lower execution time effectively. Thus it proves that the proposed LSR-KL anonymity method gives improved data privacy performance rather than the existing tCloseness, KAnonymity and LDiversity methods. Fig. 6b elaborates Memory usage by the proposed and existing methods. CPU memory usage measures the available space being used. High utilization of memory slows down the execution time and may lead to data theft.



(a) (b)

Figure 6: Graphical demonstration of proposed privacy technique based on (a) execution time (b) CPU usage

The proposed method efficiently utilizes the memory without exceeding the limits and boosts the execution time. The proposed method utilizes a memory of 5421775 kb whereas the existing methods range between 6455783–7755284 kb which illustrates full-fledge utilization of memory that leads to insecure data privacy. The result concluded that the proposed LSR-KL anonymity method provides lower CPU usage efficiency. Thus it proves that the proposed blockchain based LSR-KL anonymity method gives improved data privacy performance rather than the existing tCloseness, KAnonymity and LDiversity methods.

4.1.1 Evaluation Based on Proposed Feature Selection in Data Privacy Technique

The proposed LSROA attribute selection technique in KLAnonymity is examined for best fitness value under different iterations along with selection time and compared with the existing methods such as Cockroach Swarm Optimization (CSO), Lion Optimization Algorithm (LOA), Greenfly Aphid Swarm Optimization Algorithm (GASOA) and Rider Optimization Algorithm (ROA).

Tab. 2 shows the fitness of the proposed LSROA and existing techniques for different iterations. From the table it can be said, the existing CSO, LOA, GASOA, and ROA have the fitness value ranging between 1905.39–8641.63 for the iteration ranging from 5–25 with a step size of 5, whereas, the Proposed LSROA tends to achieve a fitness value ranging between 5108.28–9619.73. The logic behind Fitness *vs.* Iteration is to obtain the best fitness value by keeping in mind the computation time. According to that, the proposed method achieves a better fitness value within a limited iteration that makes the data privacy technique more secure against hackers. The feature selection time in the anonymity technique is graphically represented in Fig. 7.

Table 2: Evaluation of proposed LSROA based on fitness *vs.* iteration

Techniques/iterations	5	10	15	20	25
CSO	1905.39	2215.82	3891.46	4620.37	5705.05
LOA	2805.89	3272.85	4322.49	5834.22	6418.14
GASOA	3022.47	4896.58	5635.01	6813.15	7201.50
ROA	4810.10	5434.06	6410.03	7525.50	8641.63
Proposed LSROA	5108.28	6853.10	7070.45	8397.48	9619.73

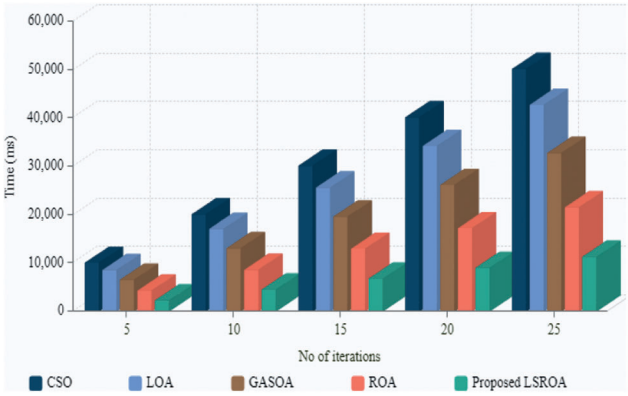


Figure 7: Graphical illustration of proposed LSROA based on feature selection time

Feature selection time depends upon the achieved fitness value at limited iterations. Obtaining the best fitness value within a low selection time is an efficient technique to boost data privacy. It is stated from the graph that the proposed LSROA tends to achieve a better solution within 11308 ms for the 25^{th} iterations. In contrast, the existing methods take more time to attain the best solution, which leads to a high selection time ranging between 21552–50182 ms for 25^{th} iterations. High selection time degrades the performance of data privacy techniques and leads to insecure data preservation. Hence, the proposed LSROA achieves the best attributes and leads the existing methods within a constrained time.

5 Conclusion

This paper proposes a novel blockchain-based framework for enhancing data privacy and access anonymity in cloud computing. The main goal of this paper is to design an efficient model for enhancing data privacy and integrity using LSR-KL Anonymity and TSPP-SS authorization technique to ensure authenticity, thereby improving secured communication in healthcare IoT. The proposed work comprises three important phases: registration, authentication, and authorization. In registration, the user and device details are collected, and the authentication level includes mapping of user id, password, and hash values. Final authorization allows the authorized user to access data through the blockchain server. In order to analyze the performance of the proposed method, the data collected from the MIMIC dataset was used. In an experimental evaluation, the performance of the LSR-KLA method is analyzed in terms of data loss and privacy preservation rate. The comparison result states that the proposed method achieves high privacy-preserving rate and data loss performance. Based on the performance analysis, the proposed method achieves the privacy-preserving rate of 95% and data loss of 12%, which are better compared to the existing methods. The experiment concluded that the proposed LSR-KLA and TSPP-SS techniques provide better performance through higher privacy-preserving rate, lower information loss, execution time, and CPU usage than the existing techniques. Thus, the proposed framework efficiently enhances data privacy and access to anonymity in cloud computing. In future work, hybrid machine learning and soft computing techniques can be proposed to progress novel and more suitable solutions to privacy problems that contain individuality disclosure that can lead to personal embarrassment and abuse.

Acknowledgement: The author with a deep sense of gratitude would thank the supervisor for his guidance and constant support rendered during this research.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] K. Miyachi and T. K. Mackey, "hOCBS: A privacy-preserving blockchain framework for healthcare data leveraging an on-chain and off-chain system design," *Information Processing & Management*, vol. 58, no. 3, pp. 1–24, 2021.

[2] M. T. Quasim, F. Algarni, A. A. E. Radwan and G. M. M. Alshmrani, "A blockchain based secured healthcare framework," in *Proc. IEEE Int. Conf. on Computational Performance Evaluation (ComPE)*, Shillong, Meghalaya, India, pp. 386–391, 2020.

[3] T. Veeramakali, R. Siva, B. Sivakumar, P. S. Mahesh and N. Krishnaraj, "An intelligent internet of things-based secure healthcare framework using blockchain technology with an optimal deep learning model," *The Journal of Supercomputing*, vol. 77, pp. 1–21, 2021.

[4] A. D. Dwivedi, L. Malina, P. Dzurenda and G. Srivastava, "Optimized blockchain model for internet of things based healthcare applications," in *Proc. Int. Conf. on Telecommunications and Signal Processing (TSP)*, Budapest, Hungary, pp. 135–139, 2019.

[5] K. M. Hossein, M. E. Esmaeili and T. Dargahi, "Blockchain-based privacy-preserving healthcare architecture," in *Proc. IEEE Canadian Conf. of Electrical and Computer Engineering (CCECE)*, Edmonton, AB, Canada, pp. 1–4, 2019.

[6] J. Kaur, R. Rani and N. Kalra, "Blockchain-based framework for secured storage, sharing, and querying of electronic healthcare records," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 20, pp. 1–24, 2021.

[7] B. S. Egala, A. K. Pradhan, V. R. Badarla and S. P. Mohanty, "Fortified-chain: A blockchain based framework for security and privacy assured internet of medical things with effective access control," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11717–11731, 2021.

[8] B. Shen, J. Guo and Y. Yang, "MedChain: Efficient healthcare data sharing via blockchain," *Applied Sciences*, vol. 9, no. 6, pp. 1–23, 2019.

[9] A. Murugan, T. Chechare, B. Muruganantham and S. G. Kumar, "Healthcare information exchange using blockchain technology," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 421–426, 2020.

[10] D. C. Nguyen, P. N. Pathirana, M. Ding and A. Seneviratne, "BEdgeHealth: A decentralized architecture for edge-based iomt networks using blockchain," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11743–11757, 2021.

[11] M. D. Cano and A. Cañavate-Sanchez, "Preserving data privacy in the internet of medical things using dual signature ECDSA," *Security and Communication Networks*, vol. 2020, no. 6, pp. 1–9, 2020.

[12] T. Frikha, A. Chaari, F. Chaabane, O. Cheikhrouhou and A. Zaguia, "Healthcare and fitness data management using the iot-based blockchain platform," *Journal of Healthcare Engineering*, vol. 2021, no. 7, pp. 1–12, 2021.

[13] M. Tahir, M. Sardaraz, S. Muhammad and M. Saud Khan, "A lightweight authentication and authorization framework for blockchain-enabled IoT network in health-informatics," *Sustainability*, vol. 12, no. 17, pp. 1–23, 2020.

[14] P. P. Ray, D. Dash, K. Salah and N. Kumar, "Blockchain for IoT-based healthcare: Background, consensus, platforms, and use cases," *IEEE Systems Journal*, vol. 15, no. 1, pp. 85–94, 2020.

[15] N. Al Asad, M. T. Elahi, A. Al Hasan and M. A. Yousuf, "Permission-based blockchain with proof of authority for secured healthcare data sharing," in *Proc. Int. Conf. on Advanced Information and Communication Technology (ICAICT)*, Dhaka, Bangladesh, pp. 35–40, 2020.

[16] P. Wei, D. Wang, Y. Zhao, S. K. S. Tyagi and N. Kumar, "Blockchain data-based cloud data integrity protection mechanism," *Future Generation Computer Systems*, vol. 102, no. 1, pp. 902–911, 2020.

[17] A. G. M. Alzahrani, A. Alenezi, A. Mershed, H. Atlam, F. Mousa *et al.,* "A framework for data sharing between healthcare providers using blockchain," in *Proc. 5th Int. Conf. on Internet of Things, Big Data and Security (IoTBDS 2020)*, United Kingdom, pp. 349–358, 2020.

[18] T. T. Thwin and S. Vasupongayya, "Blockchain-based access control model to preserve privacy for personal health record systems," *Security and Communication Networks*, vol. 2019, no. 6, pp. 1–15, 2019.

[19] A. Al Omar, M. Z. A. Bhuiyan, A. Basu, S. Kiyomoto and M. S. Rahman, "Privacy-friendly platform for healthcare data in cloud based on blockchain environment," *Future Generation Computer Systems*, vol. 95, no. 6, pp. 511–521, 2019.

[20] J. Xu, K. Xue, S. Li, H. Tian, J. Hong *et al.,* "Healthchain: A blockchain-based privacy preserving scheme for large-scale health data," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8770–8781, 2019.

[21] K. Arava and S. Lingamgunta, "Fine-grained k-anonymity for privacy preserving in cloud," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 23, no. 4, pp. 241–247, 2019.

[22] G. Nagasubramanian, R. K. Sakthivel, R. Patan, A. H. Gandomi, M. Sankayya *et al.,* "Securing e-health records using keyless signature infrastructure blockchain technology in the cloud," *Neural Computing and Applications*, vol. 32, no. 3, pp. 639–647, 2020.

[23] K. M. Prabha and P. V. Saraswathi, "Suppressed k-anonymity multi-factor authentication based schmidt-Samoa cryptography for privacy preserved data access in cloud computing," *Computer Communications*, vol. 158, no. 5, pp. 85–94, 2020.

[24] B. D. Deebak, F. Al-Turjman and L. Mostarda, "Seamless secure anonymous authentication for cloud-based mobile edge computing," *Computers & Electrical Engineering*, vol. 87, no. 10, pp. 1–13, 2020.

[25] K. Rahul, "Rider optimization algorithm (ROA): An optimization solution for engineering problem," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 12, pp. 3197–3201, 2021.

[26] S. Cao, G. Zhang, P. Liu, X. Zhang and F. Neri, "Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain," *Information Sciences*, vol. 485, no. 6, pp. 427–440, 2019.

[27] N. Li, T. Li and S. Venkatasubramanian, "T-Closeness: Privacy beyond k-anonymity and l-diversity," in *Proc. IEEE 23rd Int. Conf. on Data Engineering*, Istanbul, Turkey, pp. 106–115, 2007.