Received: 29 December 2021 | Revised: 23 June 2022 | Accepted: 15 August 2022

DOI: 10.1002/cpe.7485

RESEARCH ARTICLE

WILEY

IPFS based file storage access control and authentication model for secure data transfer using block chain technique

Mariyaprincy Antony Saviour[1] | **Dhandapani Samiappan**[2]

[1]Assistant professor, Department of Electronics and Communication Engineering, St. Joseph College of Engineering, Chennai, India

[2]Professor, Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, India

Correspondence
Mariyaprincy Antony Saviour, Department of Electronics and Communication Engineering, St. Joseph College of Engineering, Chennai, India.
Email: mariyaprincya05@gmail.com

Summary

The block chain is the emerging domain that offers the features like shared, immutable, and decentralized dataset that saves the asset's history. Security and storage are the most challenging aspect. The classical techniques' issues and challenges are an encouragement for developing a fresh block chain-assisted IPFS model with access control policies. Hence, this research paper introduces an effective access control and authentication based on an interplanetary file system (IPFS)-based block chain and access control system for attaining secured transactions. The developed method comprises five entities, that is smart contract, data owner, data requester, block chain and IPFS along with eight phases, to be exact setup, user registration, initialization, storage and data encryption, authentication, testing, access control, and decryption phase, to provide better access control and storage capacity. The proposed IPFS-based block chain and access control system offered effective storage and secure transactions in the block chain network. The highest detection rate (0.870), lowest communication overhead (0.766 MB), and highest privacy rate (0.719) are all achieved by the proposed IPFS-based block chain and access control system.

KEYWORDS

access control, authentication model, block chain technology, inter-planetary file system, smart contracts

1 | INTRODUCTION

Recently, the block chain technique is a decentralized dataset with storage mechanisms.[1] In addition, the block chain offers features like shared, immutable, and decentralized datasets, which accumulate records and transactions considering peer-to-peer (P2P) networks. The block chain is effectively used in bitcoin, and now it is well-known in different domains, like finance, healthcare, and supply chain.[2,3] A distributed immutable ledger where all transactions are recorded is provided by the block chain, which is referred to as a decentralized model. The block chain is a decentralized data store with prearranged records, such as events, also known as blocks.[4] Block chain represents an immutable and trustful distributed ledger that offers a stable, secure, auditable, transparent and effective recording of interactive data or information.[5] Thus, accumulating the data of insurance records using block chain facilitates data originality and addresses the security issues faced by central authorization techniques.[6,7] Here, the insurance business is attained amongst the client and insurance companies. However, to attain the client's privacy, the staff working on insurance should encrypt insurance data before the storage.[8] The profound computing functions minimize the effectiveness of staff and minimize the client's burden. Hence, the decryption and encryption of insurance data amongst the fog nodes are enforced.[9] The significance of block chain is facilitated to us by connecting the P2P crypto currency with bandwidth, storage space, and central processing unit (CPU) power.[10,11]

The hypertext transfer protocol (HTTP) is imperative in developing worldwide applications. However, the HTTP protocols work based on locations, which are not effective based on distributed file access. To overwhelm the issue of HTTP and BitTorrent, the IPFS[2,12] and distributed file storage

network is utilized to build the improved web for us. The IPFS utilizes a content-addressed procedure for storing with accessing the files from the web. The IPFS offers a distributed file storage model that provides connections using a P2P network. The IPFS evaluates a specific hash of files that is obtainable to all peers of networks. The hash is altered whenever the file is updated.[2,13] In the classical cloud storage model, the puzzle of single mark failure can address the technique of the decentralized storage model and enjoy several profits of the centralized storage model.[10,14] The distributed storage offers a structure wherein the data is divided into several nodes. Each node indicates distributed storage model. The IPFS is a crucial part of web 3.0 that offers a P2P decentralized file storage model and content-addressable technique for accessing the saved files.[10,15] IPFS offers a distributed hash table that is effective compared to the Git file storage system and BitTorrent. IPFS is also a version-controlled model that facilitates scalability, security, and reliability confronted by classical file storage and sharing systems.[2] Decentralized storage indicates a solution that independently permits data storage considering different network nodes with distributed ledger.[16,17]

To attain improved protection of privacy and data availability, one should enforce data storage and allocation using a central cloud storage model to decentralized storage systems with fewer prices in contrast to classical cloud storage. There exists a benefit of huge data throughput, but there is also a fear of failure.[10,18] Access control is a technique that limits genuine users' functions or tasks. The operations of access control maximize by incorporating a delegation module wherein the delegation is a procedure to assign provisional permissions amongst the users.[19,20] Mandatory access control (MAC), which is applied by a central administrator who examines the single-point failure problem, is another approach. According to the system's control mode, the IoT devices might here be a part of different management organizations. Attribute-based access control (ABAC) provides dynamic, trustworthy, adaptable, and fine-grained access control in this situation. Moreover, it abstracts the identities or the roles into a group of attributes provided by the attribute authorities. Access policies are defined using a Boolean formula considering a group of functions utilized to define the legitimate and certified Access. There is a huge requirement for assigning the roles for making access control lists amongst each system. The attribute authorities should control each attribute described in the system and split them amongst opposite users. Thus, the access management is effectually simplified with different attributes and is considered fewer users in the system.[21]

This goal is to offer an IPFS-based block chain and access control system for providing secured transactions. Furthermore, the goal is to offer common authentication between the data requester and data owner using the proposed IPFS-based block chain and access control system. Here, five entities are adapted that involve data requester, data owner, smart contract, block chain and IPFS. First, the setup phase produces security and public parameters. Then, the registration is performed to register the data requester by generating keys in the block chain network. Next, the initialization is done to record and secure the storage. Then the encryption of data is done and stored for further processing. Then, the authentication is done to the genuineness of user and testing and access control are done to provide access permissions. Finally, the decryption phase is executed for retrieving the data.

The paper's main contributions:

1. **Proposed IPFS-based block chain and access control system for secured transactions:** The proposed IPFS-based block chain and access control system is devised considering eight phases, that is setup, user registration, initialization, storage and data encryption, authentication, testing, and access control, and decryption phase.

The remaining information is listed as follows: The traditional access control methods based on block chain are described in Section 2. The block chain network's system model is covered in Section 3. The proposed IPFS-based block chain and access control system is described in Section 4; Section 5 illustrates the effectiveness of the proposed model by contrasting it with conventional methods. Section 6 provides a conclusion in the end.

2 | LITERATURE SURVEY

The eight classical techniques based on IPFS-based block chain for attaining secure transactions are detailed along with their issues. Randhir Kumar and Rakesh Tripathi[2] developed IPFS based block chain storage model for solving the storage issues of the transaction from a block considering the transaction of a specific block. The miners stored the transactions and returned the IPFS hash into a block chain block. The features in the IPFS network minimized the transaction size of the block. However, the time efficiency of this technique is very poor. Jin Sun et al.[22] devised a cipher text policy attribute-based encryption system and IPFS-based storage platform, which integrated block chain technique and devised an attribute-based encryption scheme to secure storage and provided effective sharing of health records in IPFS. This technique attained secured storage and searched for the medical data. The technique attained selective security for preventing keyword attacks. However, this technique failed to address functional issues of data accumulated in the block chain. Wang et al.[10] devised data storage and sharing scheme for a decentralized storage model and devised a model that combined the IPFS, block chain, and attribute-based encryption (ABE) technique. Here, the data owner poses the capability to split secret keys amongst data users and encrypt shared data by adapting access policies and attain fine-grained access amongst the data. However, this technique failed to execute the user's attribute revocation functions. Muqaddas Naz et al.[16] devised a block chain-based secure data sharing platform by combining the remuneration of IPFS. The metadata was uploaded to the IPFS server using the owner and then splitted into secret shares. The technique attained access control and security by adapting the access roles provided in the smart contract. Here, the encryption,

decentralized storage, and Ethereum block chain were integrated. However, this technique failed to concentrate on data monetization considering block chain.

Ammar Ayman Battah et al.[23] developed a fully decentralized block chain-based solution wherein the MPA was executed with proxy re-encryption techniques and Ethereum smart contracts and executed with different oracles to provide access to encrypted data a decentralized storage platform. The smart contract assists in verifying the outcomes using reputation mechanisms for determining suspicious behaviors. However, the technique failed to utilize front-end decentralized applications. Jin Sun et al.[9] developed a block chain technology and a cipher text-policy attribute-based encryption system for secured storage and update of insurance records amongst the IPFS storage platforms. Here, the fog nodes were utilized to outsource the encryption of records to enhance staff effectiveness. The security proof revealed that the technique attained selective security in opposition to the keyword attacks but failed to leverage smart contracts to deploy the model. Chao Lin et al.[24] developed a block chain-based system, namely BSeIn, to secure mutual authentication to provide fine-grained access control policies. The developed model utilized multi-receivers' encryption, attribute signature, and message authentication code to offer security and privacy. However, this technique failed to utilize hardware to optimize performance. Ding et al.[21] devised an attribute-based access control technique for the IoT model wherein the access management was performed with block chain. The block chain recorded the attribute distributions to avoid data tampering and single point failure. The access control was optimized for meeting the requirements of higher efficiencies and lightweight evaluations of IoT devices. The technique was suitable for resisting different attacks but unsuitable for fine-grained access control models. Wajde Baiod et al.[25] devised block chain technology and its applications across multiple domains. This study examines block chain-based use cases in a variety of industries, including finance, journalism, healthcare, and so forth. The supply chain systems are made more visible, transparent, and accountable thanks to this survey. However, the complete lack of block chain competence in many firms poses a significant barrier to the development of the block chain. Abdullah Al-Noman Patwary et al.[26] devised a secure location-based authentication scheme in fog computing environments using block chain. In this case, a Block chain can be used by fog devices to mutually authenticate one another at the fog layer. The proposed authentication mechanism was efficient and secure. However, due of the location validation operations carried out, the suggested system occasionally involves greater computing overhead than the previous approaches. The proposed IPFS-based block chain and access control system is devised considering eight phases: setup, user registration, initialization, storage and data encryption, authentication, testing, access control, and decryption. As a result, it offers effective storage. The comparison of existing works and proposed works is depicted in Table 1.

3 | CHALLENGES

The problems with traditional block chain-assisted access control methods are mentioned,

1. In reference 22, an attribute-based encryption technique is devised for secured storage and effective sharing of health records in IPFS. However, this technique failed to use the feature revocation utility to resolve the issues regarding expired users' access privileges.
2. Data storage and sharing techniques are devised to allocate the secret key amongst the data users and encrypt the shared data considering the access policies. However, this technique acquired fine-grained access control amongst the data but failed to execute the functions of access policy update and user's attribute revocation.[10]
3. A fully decentralized block chain technique is devised in reference 23 to access IPFS encrypted data. The block chain poses various entities that interact through software. However, the technique failed to use re-encryption oracles for integrating physical entities.
4. In reference 24, a secure mutual authentication technique is devised for enforcing fine-grained access control policies. This technique effectively offered security and privacy, like auditability, anonymous authentication, and confidentiality. However, this technique failed to induce underpinning block chain features for generating decentralized solutions.
5. The cloud uses a local dataset to keep the terminal access histories. As a result, there is a very real possibility that these datasets will be exposed to unauthorized access and modification. It is extremely challenging to determine the audibility and traceability of access records.

4 | SYSTEM MODEL

Block chain technology is essential for tackling the issue of data origin in the block chain network, since it offers a reliable and irrevocable public ledger for recording all transactions. Figure 1 depicts the system concept of the IPFS paradigm with block chain support for secure transactions.

The block chain-assisted IPFS model comprises three imperative functionalities: block chain technology, smart contracts, and IPFS.

TABLE 1 Comparison of existing works and proposed works

Authors	Methods	Advantages	Disadvantages
Randhir Kumar et al.[2]	IPFS based block chain storage model	The miners stored the transactions and returned the IPFS hash into the block chain block. The features in the IPFS network minimized the transaction size of the block.	The time efficiency of this technique is very poor.
Jin Sun et al.[22]	Ciphertext policy attribute-based encryption system and IPFS-based storage platform	This technique attained secured storage and searched for the medical data. The technique attained selective security for preventing keyword attacks.	This technique failed to address functional issues of data accumulated in the block chain.
Wang, S et al.[10]	Data storage and sharing scheme for a decentralized storage model	By adjusting access regulations and achieving fine-grained access amongst the data, the data owner can divide the secret key amongst data users and encrypt shared data.	This technique failed to execute functions of the user's attribute revocation.
Muqaddas Naz et al.[16]	Block chain-based secure data sharing platform	The encryption, decentralized storage, and Ethereum block chain were integrated.	This technique failed to concentrate on data monetization considering block chain.
Ammar Ayman Battah et al.[23]	A fully decentralized block chain-based solution	The smart contract assists in verifying the outcomes using reputation mechanisms for determining suspicious behaviors.	The technique failed to utilize front-end decentralized applications.
Jin Sun et al.[9]	Block chain technology and a cipher text-policy attribute-based encryption system	The fog nodes were utilized to outsource the encryption of records to enhance staff effectiveness.	This technique failed to leverage smart contracts to deploy the model.
Chao Lin et al.[24]	Block chain-based system.	The developed model utilized multi-receivers' encryption, attribute signature, and message authentication code for offerings security and privacy.	This technique failed to utilize hardware to optimize performance.
Ding S et al.[21]	Attribute-based access control technique for IoT model	The access control was optimized for meeting the requirements of higher efficiencies and lightweight evaluations of IoT devices.	The technique was unsuitable for fine-grained access control models.
Wajde Baiod et al.[25]	Applications of block chain technology in various fields.	The supply chain systems are made more visible, transparent, and accountable thanks to this survey.	A major barrier to the development of block chain technology is the complete lack of block chain knowledge in many enterprises.
Abdullah Al-Noman Patwary et al.[26]	Secure location-based authentication scheme in fog computing environments using block chain.	Block chain-based secure location-based authentication in fog computing environments.	Due to the location validation processes carried out, the suggested system occasionally involves greater computing overhead than the current approaches.
Proposed System	IPFS based File Storage Access control and Authentication model for Secure Data Transfer using block chain technique	The proposed IPFS-based block chain and access control system offered effective storage and secure transactions in the block chain network, and also it attains the highest detection rate.	The deployment of huge scale solutions can be a comprehensive study in the future.

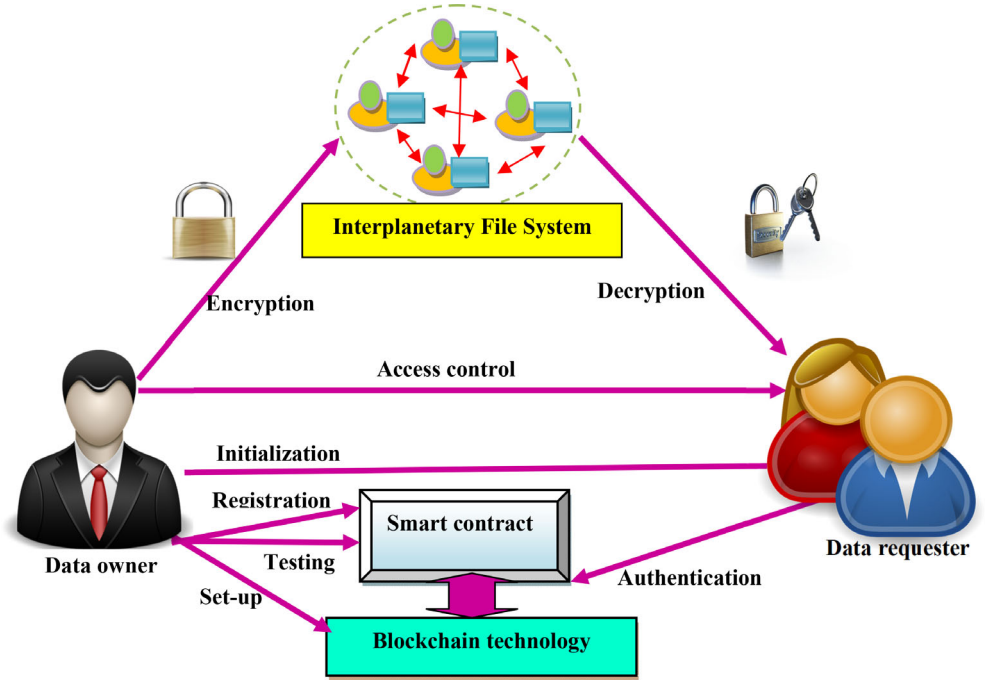

FIGURE 1 System model of block chain assisted IPFS model for secure transactions.

4.1 | Block chain technology

The block chain represents a distributed dataset, which records all transactions in theP2P network. Each user in the network holds a similar copy of the dataset. Here, central authority is absent, and no single node manages the complete network. The block chain is a sequence of associated data blocks wherein the blocks are added to the block chain with consensus amongst the nodes. Each block comprises a block header and a group of transactions, wherein each block header comprises link pointers and timestamps. Thus, the transaction is occurred in the block chain using a cryptography hash algorithm.

4.2 | Smart contracts

Smart contracts are a type of computer technology that can be executed and deployed without manual interference. A smart contract can be employed as a computer program, which can automatically perform all contracts-based functions and generate the evidence validated through the contract process. A smart contract is an account linked with some code in the block chain.

4.3 | Interplanetary file system

Interplanetary file system (IPFS) is a P2P disseminated file model that helps to link all computing devices using similar files. The IPFS offers a content-addressed block storage model with hyperlinks, and it integrates the technologies like incentivized block exchange, distributed hash tables (DHT), and self-certifying namespaces. The IPFS is useful and can be accessed from anywhere and anytime. Once uploaded to the IPFS system, the file generates a unique cryptographic hash string that helps retrieve the uploaded file. Table 2 represents attributes and their purposes.

5 | PROPOSED IPFS-BASED BLOCK CHAIN AND ACCESS CONTROL SYSTEM FOR SECURE TRANSACTIONS

The goal is to devise an IPFS-based block chain storage system and access control for secure transactions. Presenting a model for the system's mutual authentication of the requester and data owner is the main contribution. The data requester, data owner, smart contract, block chain, and

TABLE 2 Attributes and their purposes

Attributes	Purposes
IPFS based block chain storage	The Interplanetary File System is a code of behavior and peer-to-peer network for storing and allocating data in a disseminated folder system. IPFS uses content addressing to individually recognize each folder in a universal namespace linking all computing devices.
Attribute-based encryption	Attribute-based encryption is used to keep a record of encryption.
Data storage, sharing, and block chain-based secure data sharing scheme	Data storage records media to preserve statistics through computers or different devices. Data sharing is the capability to hand out the matching sets of statistics possessions to various users or applications while maintaining statistics trustworthiness crosswise all entities overwhelming the statistics.
Interplanetary file system (IPFS)	The interplanetary file system is a code of behavior and peer-to-peer network for storing and allocating data in a disseminated folder system. IPFS uses content addressing to individually recognize each folder in a universal namespace linking all computing devices.

IPFS are the additional five entities that make up this system. A data owner in this context is a person who has access to several files that contain data. Additionally, the data requester is a client of the data owner who has permission to see certain files. Additionally, there are eight processes in the created IPFS-based block chain storage system: setup, user registration, initialization, storage and data encryption, authentication, testing, access control, and decryption. The setup phase is the first thing the owner does. During this phase, input security attributes are collected, and the system master key and public parameter are generated. The owner is still working on the user registration phase. The system master key is obtained during the user registration step, and an output of a secret key is produced. Initialization is the following step, and it is carried out to document, distribute, and secure the entire storage process. The data is kept for later processing after the startup step is finished. The files are encrypted using an algorithm at the encryption step that accepts shared files as input. The participant's identity involved in communication before sharing or exchanging data is then authenticated during the authentication stage. The test phase then returned the pertinent transaction ID and the results of the success key matching. In this case, the testing phase serves to validate and verify the outcomes.

As a result, the access control phase is complete, which limits the actions and functions of a legitimate user. The encrypted file is then decrypted using the decryption algorithm at the final step of encryption. The proposed IPFS-based block chain and access control system's schematic perspective of access control and access authentication for safe transactions is shown in Figure 2. The symbol description of the proposed IPFS-based block chain and access control system is provided in Table 3.

5.1 | Set up phase

The setup phase is executed by the data owner such that the data owner generates security parameters s and p, public parameters c and d, and random numbers a and b that belong to [0, 1]. Then, the data owner deploys a smart contract SC on the block chain.[10] Figure 3 presents the setup phase of the proposed IPFS-based block chain and access control system.

5.2 | Registration phase

The data owner runs the registration phase. In the registration phase, data requester ID and password are generated by the data requester and fed to the data owner. The data owner and smart contract store the obtained data requester ID and data requester password R^*_{PW}. Furthermore, the data owner generates the security key K_R by considering the data requester ID R_{ID}, random number a and security parameter s. Here, the security key K_R is generated by concatenating data requester ID R_{ID}, and random number a and the hashing is performed on the obtained result and further XORed with the security parameter s, which is formulated as,

$$K_R = s \oplus h\left(R_{ID} \| a\right) \tag{1}$$

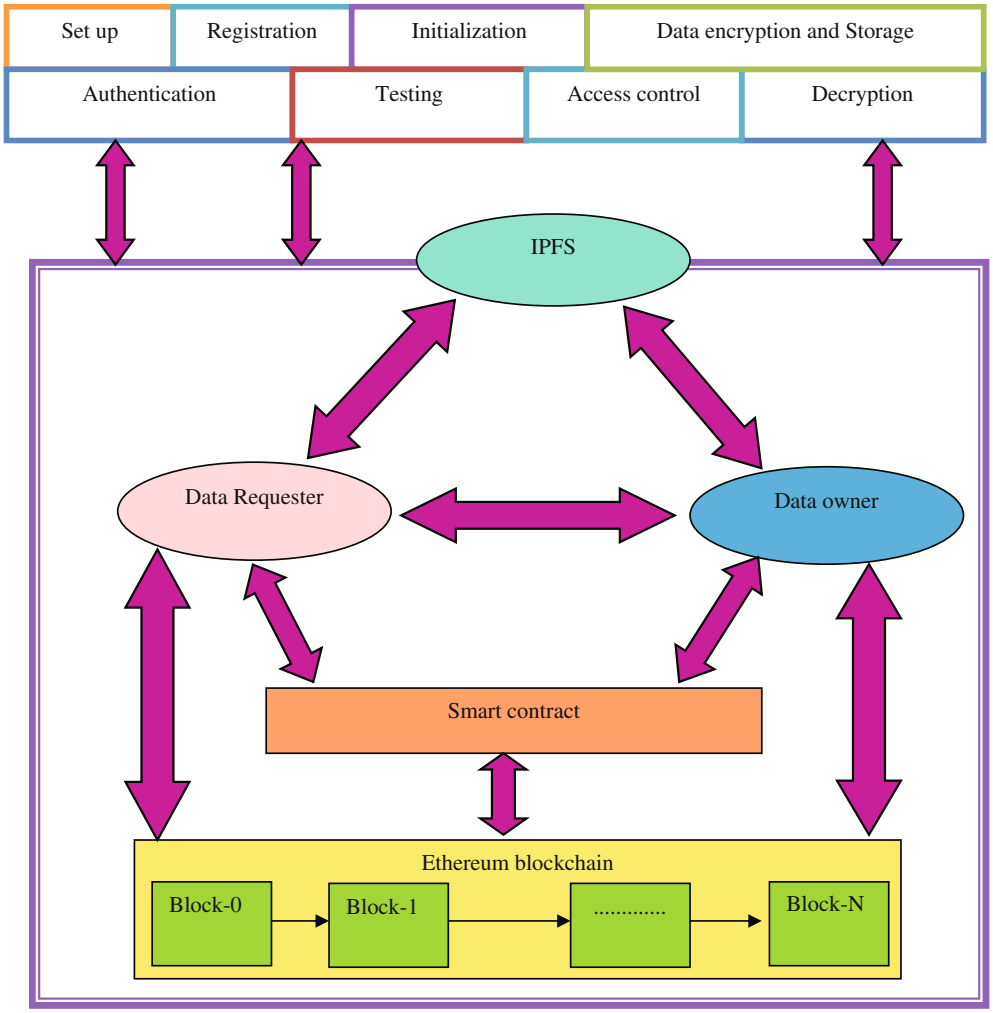

FIGURE 2 Schematic view of access control and access authentication for secure transaction using proposed IPFS based block chain and access control system

The security key K_R is stored by the data requester as K_R^*. In addition, it is also stored in a smart contract as K_R^*. After that, the data owner computes if $K_R = K_R^*$ the data requester is registered. Figure 4 presents the registration phase of the proposed IPFS-based block chain and access control system.

5.3 | Initialization phase

In the initialization phase, the registered requester sends a data access request to the data owner, and then the data owner computes and generates a message. The message is computed by encrypting the concatenated result obtained by concatenating data owner ID O_{ID} and data requester password R_{PW}. The obtained encrypted results are further XORed with the security parameter, which is expressed as,

$$m = s \oplus E\left(O_{ID} \| R_{PW}\right) \tag{2}$$

The obtained message is stored in a smart contract as m^*. Finally, the transaction ID and smart contract address are sent to the data requester. Figure 5 displays the initialization phase of the proposed IPFS-based block chain and access control system.

5.4 | Data encryption and storage phase

The data owner runs the encryption phase. Here, the data owner gets the data from the block chain and encrypts the data. Data encryption is done by concatenating input data D, file encryption key K_E and Chebyshev polynomial equation A. The encryption is done on the obtained concatenated

TABLE 3 Demonstrates the symbol description of the proposed IPFS-based block chain and access control system

Symbol	Description
s and p	Security parameter
c and d	Public parameter
a and b	Random numbers
R_{ID} and R_{PW}	ID and password of the data requester
K_R	Security key
T_{ID}	Transaction ID
SC_A	Smart contact address
$E(.)$	Encryption
$h(.)$	Hashing
A	Chebyshev polynomial
O_{ID}	ID of the data owner
D_{loc}	Data location
D^E	Encrypted data
M_1 and M_2	Authenticated messages
T_o	Timestamp generated by the owner
T_{SC}	Timestamp generated by smart contract
K_E	File encryption key
de	Decryption

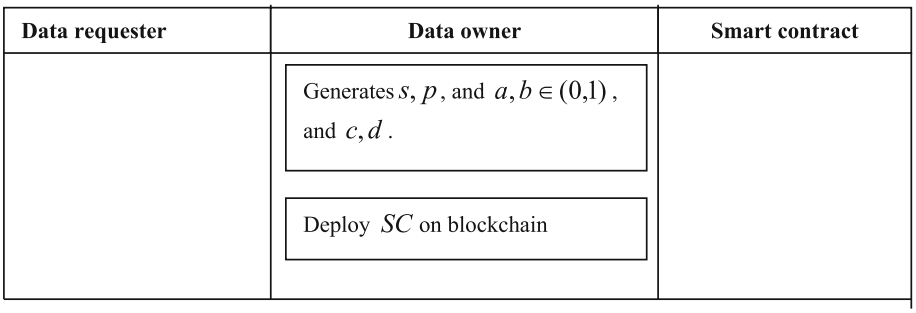

FIGURE 3 Set up phase of proposed IPFS-based block chain and access control system

result. Figure 6 depicts the data encryption and storage phase of the proposed IPFS-based block chain and access control system. The encrypted result is further XORed with a public parameter c and is given as,

$$D^E = E\left(D \| K_E \| A\right) \oplus c \tag{3}$$

The Chebyshev polynomials are two sequences of polynomials correlated to the cosine and sine functions used in the filter design. For the design of filters, Chebyshev polynomials are utilized. Plotting two cosine functions with time, one with a fixed frequency and the other with an increasing frequency, will yield them. Here, the Chebyshev polynomial equation A is expressed as,

$$A = 4x^3 - 3x \tag{4}$$

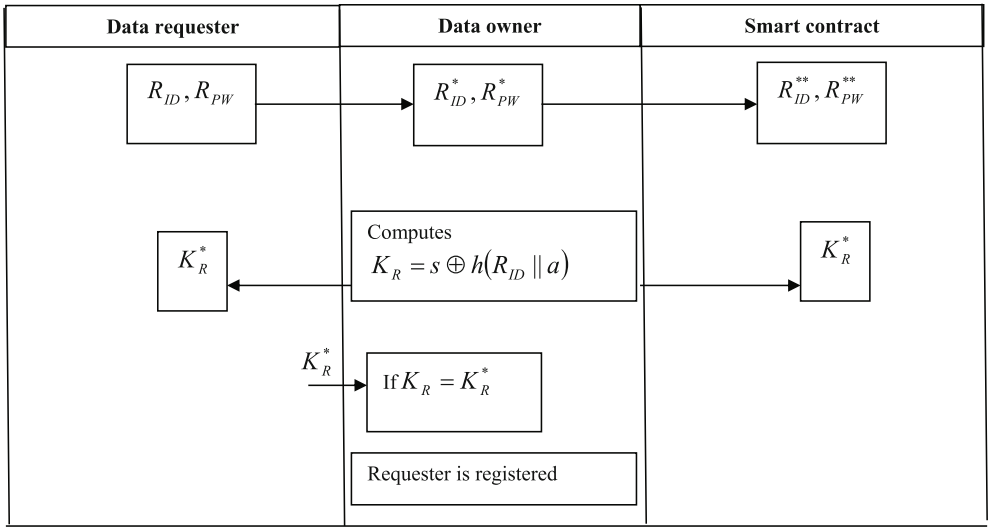

FIGURE 4 Registration phase of proposed IPFS-based block chain and access control system

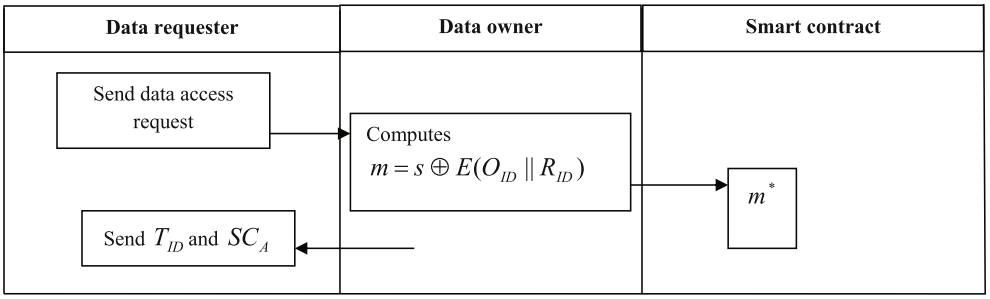

FIGURE 5 Initialization phase of proposed IPFS-based block chain and access control system

Here, the value of a variable *x* is obtained by concatenating the data owner ID O_{ID} and public parameter *d*. The concatenated result is further hashed and multiplied with a random number *a* and is formulated as,

$$x = a^* h\left(O_{ID} \| d\right) \tag{5}$$

The data owner records data location D_{loc} from IPFS. Finally, the data owner computes the file encryption key generated message *m* and random number *b*. The obtained result is encrypted. Likewise, the security parameter *p* and public parameter *d* are concatenated and the hashing is done on the concatenated result. The encrypted outcome and hashed outcome are multiplied to generate the file encryption key, which is formulated as,

$$K_E = E(m \| b) * h(p \| d) \tag{6}$$

5.5 | Authentication phase

The data requester generates two authenticated messages. The first authentication message is built by concatenating the stored security key K^*_R and public parameter *c*. The hashing is performed on obtained result and further XORed with ID of the data requester and is expressed as,

$$M_1 = R_{ID} \oplus h\left(K^*_R \| c\right) \tag{7}$$

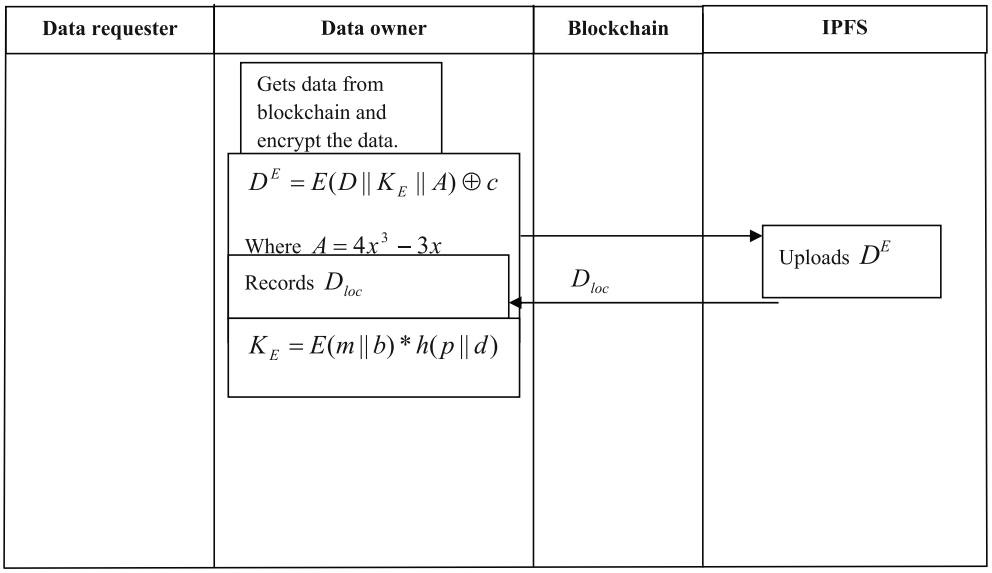

FIGURE 6 Data encryption and storage phase of proposed IPFS-based block chain and access control system

The second authentication message is obtained by multiplying the first authentication message with the data requester password and then encrypted. The modulus of the encrypted result and the random number *a* is performed to generate the second authentication message, which is formulated as,

$$M_2 = E\left(M_1^* R_{PW}\right) \text{mod} a \tag{8}$$

They obtained two authentication messages M_1 and M_2 are subjected to a smart contract to check correctness.

In a smart contract, the two authentication messages are computed. The first authentication message is obtained by concatenating the stored security key K_R^* and public parameter *c*. Then the hashing is performed on the concatenated result and further XORed with stored data requester ID, which is given as,

$$\widetilde{M}_1 = R_{ID}^{**} \oplus h\left(K_R^* \| c\right) \tag{9}$$

The second authentication message is obtained by multiplying the first authentication message $\widetilde{M}_1$ and stored data requester password R_{PW}^{**}. Then, the encryption is performed on the resultant product. After that, the modulus is performed between random number *a* and resultant product, which is formulated as,

$$\widetilde{M}_2 = E\left(\widetilde{M}_1^* R_{PW}^{**}\right) \text{mod} a \tag{10}$$

The smart contract computes if $M_1 = \widetilde{M}_1$ *and* $M_2 = \widetilde{M}_2$ the data requester is authenticated. Figure 7 presents the authentication phase of the proposed IPFS-based block chain and access control system.

5.6 | Testing phase

The data owner implements the testing phase. Here, the data owner generates a timestamp T_o and evaluates access permission. Here, the access control message is generated by concatenating the timestamp generated by the owner T_o, message *m* and public parameter *c* and then hashing is performed on the concatenated result, which is formulated as,

$$g = h\left(T_o \| m \| c\right) \tag{11}$$

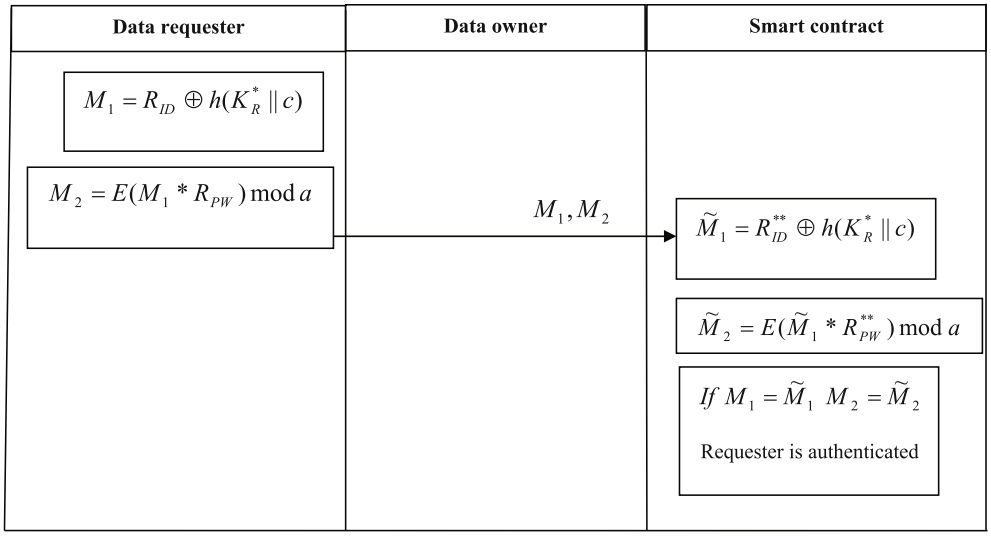

FIGURE 7 Authentication phase of proposed IPFS-based block chain and access control system

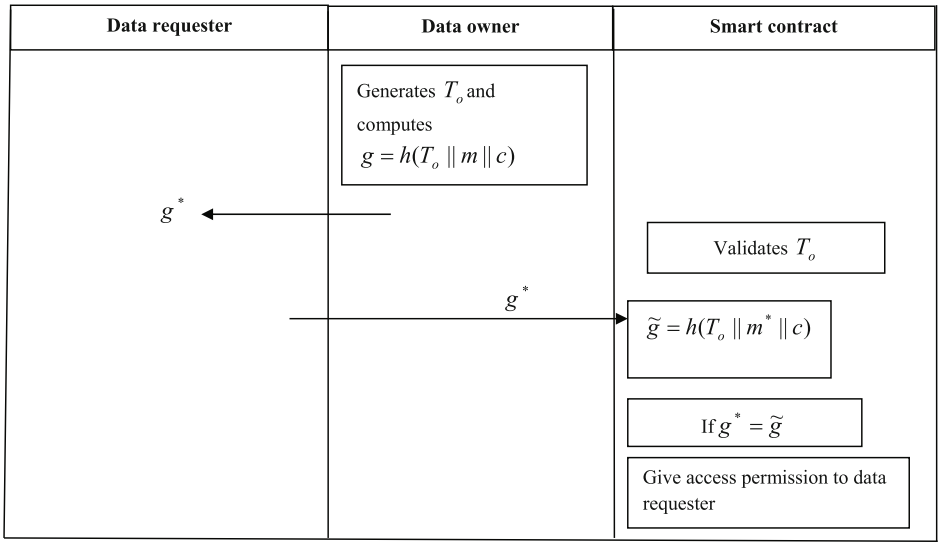

FIGURE 8 Testing phase of proposed IPFS-based block chain and access control system

The formed access control message is stored by the data requester as g^*. In a smart contract, the timestamp generated by the owner T_o is validated. After that, the stored access control message g^* is utilized to generate the smart contract access control message by concatenating the timestamp generated by the owner T_o, stored message m^* and public parameter c. The hashing is performed on a concatenated result, which is formulated as,

$$\tilde{g} = h\left(T_o \| m^* \| c\right) \tag{12}$$

If stored access control message g^* and smart contract access control message, $\tilde{g}$ that is if $g^* = \tilde{g}$ the access permission is given to the data requester. Figure 8 presents the testing phase of the proposed IPFS-based block chain and access control system.

5.7 | Access control phase

The data owner executes the access control phase by checking smart contracts. Initially, the smart contract generates T_{SC} and evaluates public parameters by concatenating the timestamp generated by the smart contract T_{SC} and random number a. The hashing is performed on the concatenated

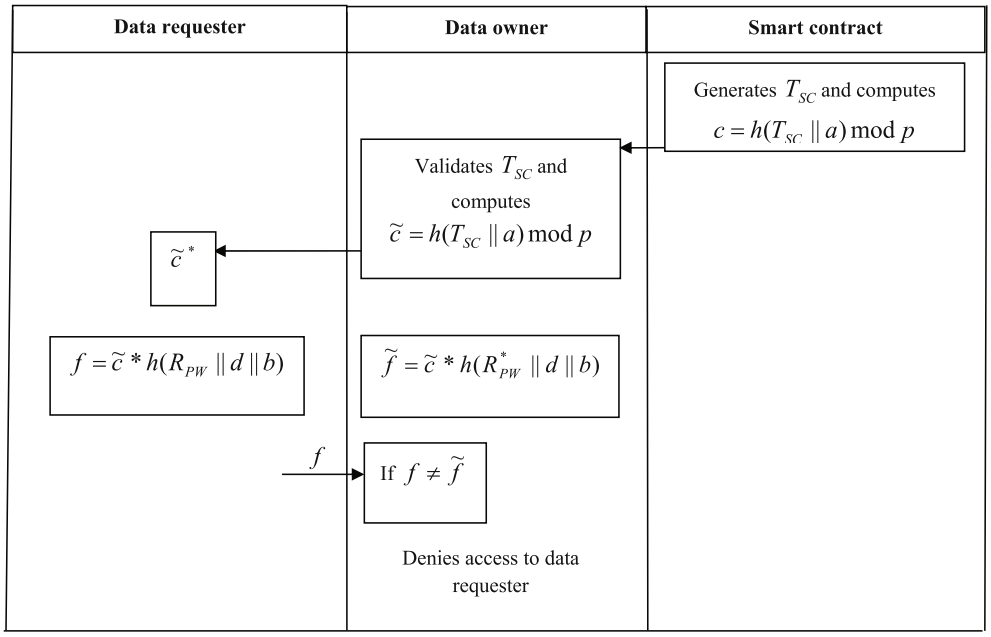

FIGURE 9 Access control phase of proposed IPFS-based block chain and access control system

result. The modulus of the security parameter p and the hashed result is done and is formulated as,

$$c = h\left(T_{SC} \| a\right) \bmod p \tag{13}$$

In a smart contract, the validation of the timestamp generated by the smart contract T_{SC} is evaluated, and the data owner's public parameter $\tilde{c}$ is computed by concatenating the timestamp generated by the smart contract T_{SC} and random number a. The hashing is performed on the concatenated result. The modulus of the security parameter p and the hashed result is done and is formulated as,

$$\tilde{c} = h\left(T_{SC} \| a\right) \bmod p \tag{14}$$

The obtained security parameter is stored by the data requester as $\tilde{c}^*$.

The data requester generates a message f by concatenating the data requester password, public parameter d and random number b. The hashing is performed on obtained concatenated result and further multiplied with the security parameter of the data requester $\tilde{c}^*$, which is formulated as,

$$f = \tilde{c}^* h\left(R_{PW} \| d \| b\right) \tag{15}$$

The data owner generates a message $\tilde{f}$ by concatenating stored data requester password R^*_{PW}, public parameter d and a random number b. The hashing is performed on obtained concatenated result and further multiplied with the security parameter of the data requester $\tilde{c}^*$, which is formulated as,

$$\tilde{f} = \tilde{c}^* h\left(R^*_{PW} \| d \| b\right) \tag{16}$$

The data owner computes if $f \neq \tilde{f}$ the access is denied to the data requester. Figure 9 displays the access control phase of the proposed IPFS-based block chain and access control system.

5.8 | Decryption phase

Figure 10 displays the decryption phase of the proposed IPFS-based block chain and access control system. Here, the data owner computes if $f = \tilde{f}$, the data requester stores data location and file encryption key as D^*_{loc} and K^*_E. The stored data location D^*_{loc} and file encryption key K^*_E are fed to IPFS,

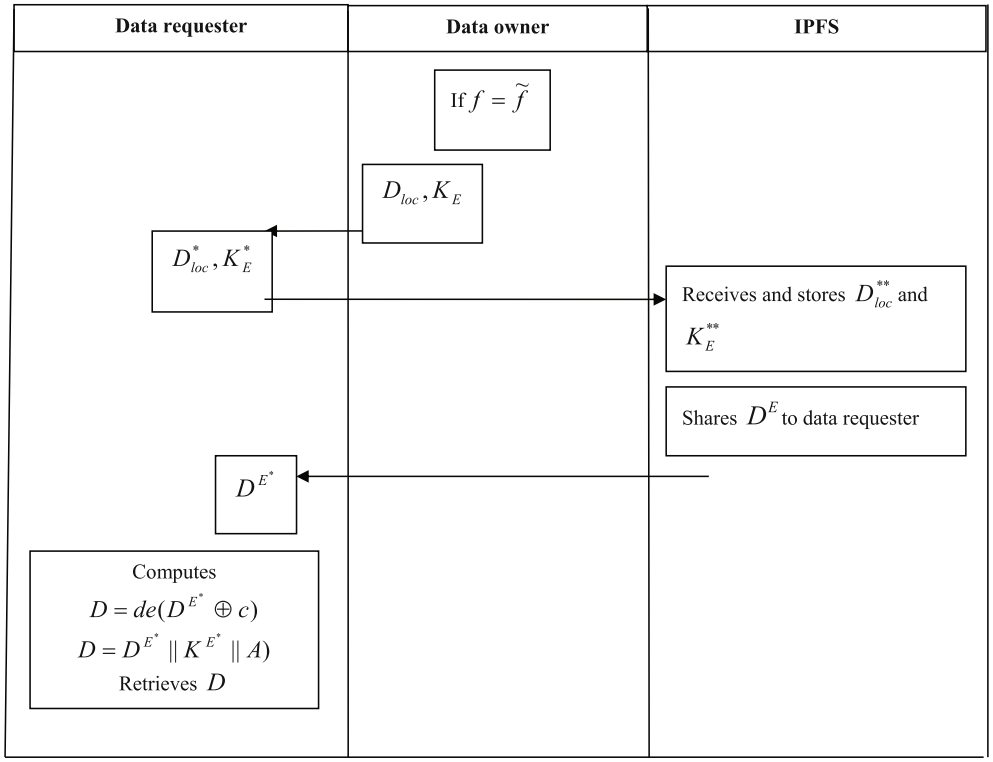

FIGURE 10 Decryption phase of proposed IPFS-based block chain and access control system

which is received and stored K_E^{**}. The IPFS shares encrypted data D^E to the data requester, which the data requester stores D^{E^*}. Finally, the requester computes the input data by XORing the encrypted data and public parameters *c*. The decryption is performed on the obtained XORed data, which is formulated as,

$$D = de\left(D^{E^*} \oplus c\right) \tag{17}$$

After that, the final input data *D* is retrieved by concatenating stored encrypted data D^{E^*}, file encryption key K_E and Chebyshev polynomial equation *A* and is formulated as,

$$D = D^{E^*} \left\| K^{E^*} \right\| A \tag{18}$$

Finally, the data *D* is retrieved.

6 | RESULTS AND DISCUSSION

The effectiveness of the proposed IPFS-based block chain and access control system is checked using detection rate, communication overhead, and private rate. Here, the research is conducted using 5-block chain networks of sizes 5, 10, 15, and 20-block chain networks. A variable number of users also participate in the assessment.

6.1 | Experimental set up

The suggested solution is implemented using Python on a computer running Windows 10 with 2GB of RAM and an Intel i3 core processor. The experimental setup for our suggested strategy is shown in Table 4.

TABLE 4 Experimental setup

Parameters	Values
Security parameters	$P = 0$ to 1 $A,B = 0$ to 1
Block size	20
Encryption attributes	Bit = 32
Input parameters for proposed method	Data, user_id, user_password, owner id, owner password, number of blocks

6.2 | Evaluation measures

The following performance measures—detection rate, communication overhead and private rate—are used to evaluate performance.

6.2.1 | Detection rate

The detection rate is described as a ratio of the number of suspicious users, which are determined correctly in contrast to the total number of users registered in a block chain network, and is formulated as,

$$DR = \frac{N_s}{N_T} \tag{19}$$

where N_s represent some suspicious users, which are determined correctly and signify the total number of users registered in the block chain network.

6.2.2 | Communication overhead

Long delays connected with accessing stored data are referred to as the communication overhead.

6.2.3 | Private rate

The indicator that guards against unwanted data access is the privacy rate.

6.3 | Comparative methods

The methods, like IPFS based block chain storage,[2] attribute-based encryption,[22] data storage and sharing,[10] block chain-based secure data sharing,[16] and proposed IPFS-based block chain and access control system are taken for the analysis.

6.4 | Comparative analysis

By changing the user count, evaluations of strategies for detection rate, communication overhead, and private rate are conducted. The block chain sizes of 5, 10, 15, and 20 are taken into consideration for the analysis in this case.

6.4.1 | Analysis with block chain size = 5

Figure 11 displays analysis of methods with block chain size = 5, considering detection rate, communication overhead, and private rate. The analysis of methods using detection rate is portrayed in Figure 11A. When number of user is 100, the corresponding detection rate measured by IPFS-based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block

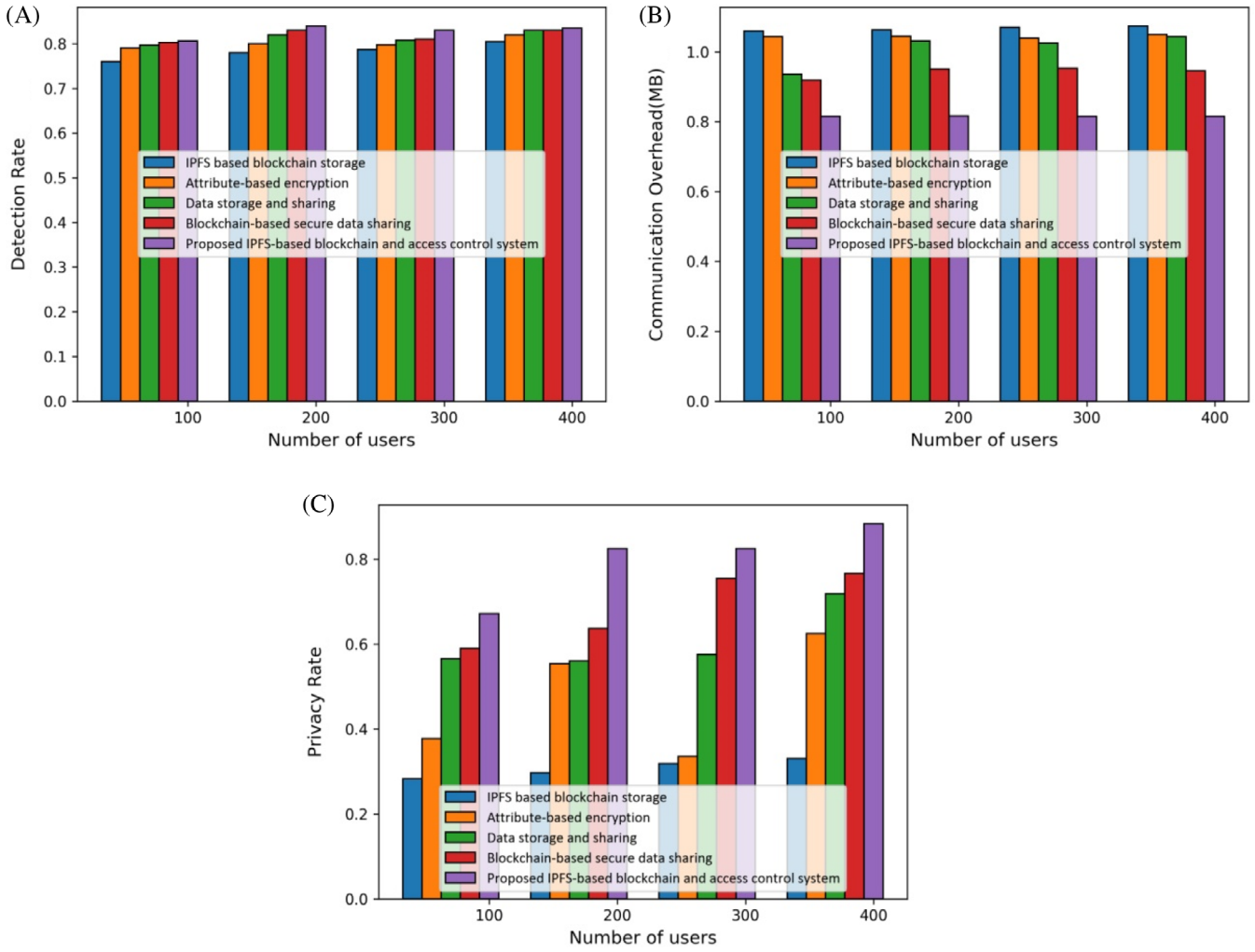

FIGURE 11 Analysis of methods with block chain size = 5 considering (A) detection rate, (B) communication overhead, (C) private rate

chain and access control system are 0.760, 0.790, 0.797, 0.803, and 0.807. Likewise, when number of user is 400, the corresponding detection rate measured by IPFS-based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.805, 0.820, 0.830, 0.830, and 0.835. The analysis of methods using Communication overhead is portrayed in Figure 11B. When number of user is 100, the corresponding communication overhead measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are1.060, 1.043, 0.936, 0.919, and 0.816 MB. Likewise, when number of user is 400, the corresponding Communication overhead measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are1.074, 1.050, 1.043, 0.945, and 0.816 MB. The analysis of methods using private rate is portrayed in Figure 11C. When number of user is 100, the corresponding private rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.283, 0.377, 0.566, 0.590, and 0.672. Likewise, when number of user is 400, the corresponding private rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.330, 0.625, 0.719, 0.766, and 0.884.

6.4.2 | Analysis with block chain size = 10

Analysis of techniques with a block chain size of 10 is shown in Figure 12 while taking into account the detection rate, communication overhead, and private rate. Figure 12A depicts the analysis of techniques utilizing detection rate. When there are 100 users, the corresponding detection rates

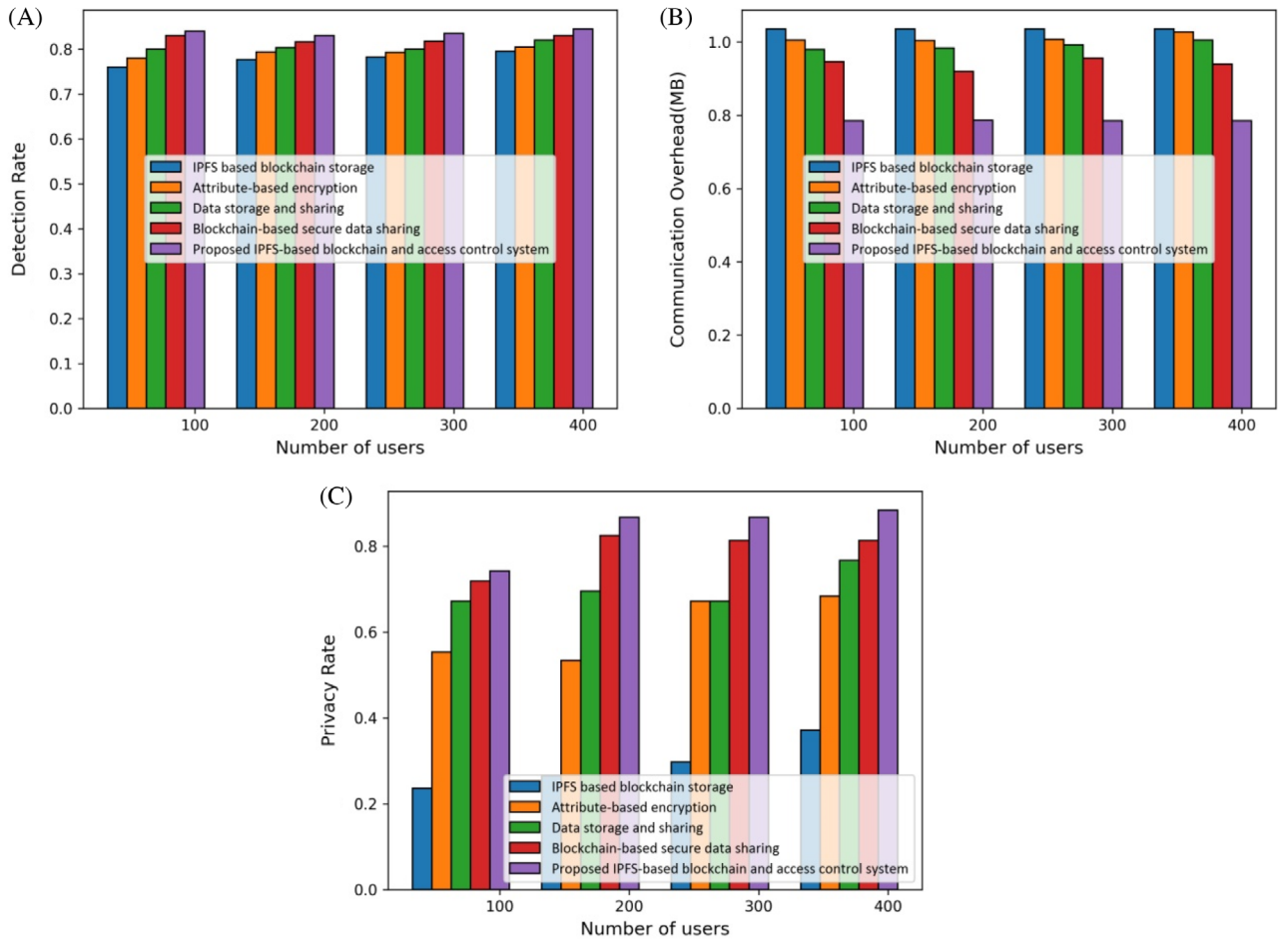

FIGURE 12 Analysis of methods with block chain size = 10 considering (A) detection rate, (B) communication overhead, (C) private rate

for attribute-based encryption, secure data sharing, and IPFS-based block chain storage are 0.760, 0.780, 0.800, and 0.830, respectively, while the detection rate for the proposed IPFS-based block chain and access control system is 0.840.Likewise, when number of user is 400, the corresponding detection rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing are 0.795, 0.805, 0.820, 0.830, while the detection rate evaluated by proposed IPFS-based block chain and access control system is 0.845. The analysis of methods using communication overhead is portrayed in Figure 12B. When number of user is 100, the corresponding communication overhead measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing are 1.036, 1.005, 0.980, 0.947 MB, while communication overhead measured by proposed IPFS-based block chain and access control system is 0.786 MB. Likewise, when number of user is 400, the corresponding communication overhead measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing are 1.036, 1.028, 1.005, 0.940, while communication overhead measured by proposed IPFS-based block chain and access control system is 0.785 MB. The analysis of methods using private rate is portrayed in Figure 12C. When number of user is 100, the corresponding private rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing are 0.236, 0.553, 0.672, 0.719, while privacy rate evaluated by proposed IPFS-based block chain and access control system is 0.743. Likewise, when number of user is 400, the corresponding private rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing are 0.371, 0.684, 0.766, 0.813, while privacy rate evaluated by proposed IPFS-based block chain and access control system is 0.884.

6.4.3 | Analysis with block chain size = 15

Figure 13 displays analysis of methods with block chain size = 15, considering detection rate, communication overhead, and private rate. The analysis of methods using detection rate is portrayed in Figure 13A. When number of user is 100, the corresponding detection rate measured by IPFS based block chain storage is 0.733, attribute-based encryption, is 0.755, data storage and sharing is 0.790, block chain-based secure data sharing is

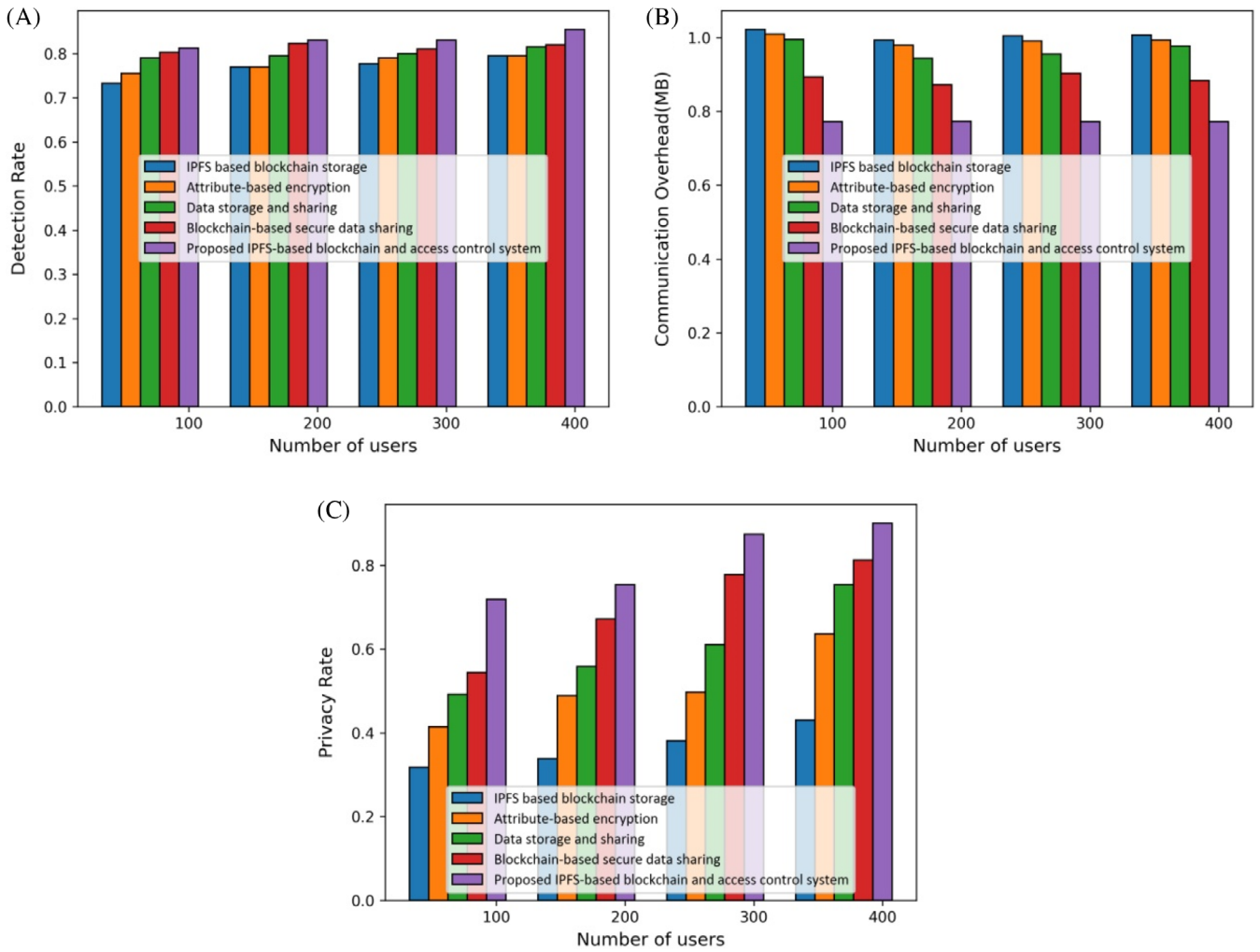

FIGURE 13 Analysis of methods with block chain size = 15 considering (A) detection rate, (B) communication overhead, (C) private rate

0.803, and proposed IPFS-based block chain and access control system is 0.813. Likewise, when number of user is 400, the corresponding detection rate measured by IPFS based block chain storage is 0.795, attribute-based encryption is 0.795, data storage and sharing is 0.815, block chain-based secure data sharing is 0.820, and proposed IPFS-based block chain and access control system is 0.855. The analysis of methods using communication overhead is portrayed in Figure 13B. When number of user is 100, the corresponding communication overhead measured by IPFS based block chain storage is 1.023 MB, attribute-based encryption is 1.010 MB, data storage and sharing is 0.995 MB, block chain-based secure data sharing is 0.893 MB, and proposed IPFS-based block chain and access control system is 0.773 MB. Likewise, when number of user is 400, the corresponding communication overhead measured by IPFS based block chain storage is 1.007 MB, attribute-based encryption is 0.993 MB, data storage and sharing is 0.978 MB, block chain-based secure data sharing is 0.883 MB, and proposed IPFS-based block chain and access control system is 0.772 MB. The analysis of methods using private rate is portrayed in Figure 13C. When number of user is 100, the corresponding private rate measured by IPFS based block chain storage is 0.318, attribute-based encryption is 0.414, data storage and sharing is 0.492, block chain-based secure data sharing is 0.544, and proposed IPFS-based block chain and access control system is 0.719. Likewise, when number of user is 400, the corresponding private rate measured by IPFS based block chain storage is 0.431, attribute-based encryption is 0.636, data storage and sharing is 0.754, block chain-based secure data sharing is 0.813, and proposed IPFS-based block chain and access control system is 0.901.

6.4.4 | Analysis with block chain size = 20

Figure 14 displays analysis of methods with block chain size = 20 considering detection rate, communication overhead, and private rate. The analysis of methods using detection rate is portrayed in Figure 14A. When number of user is 100, the corresponding detection rate measured by IPFS

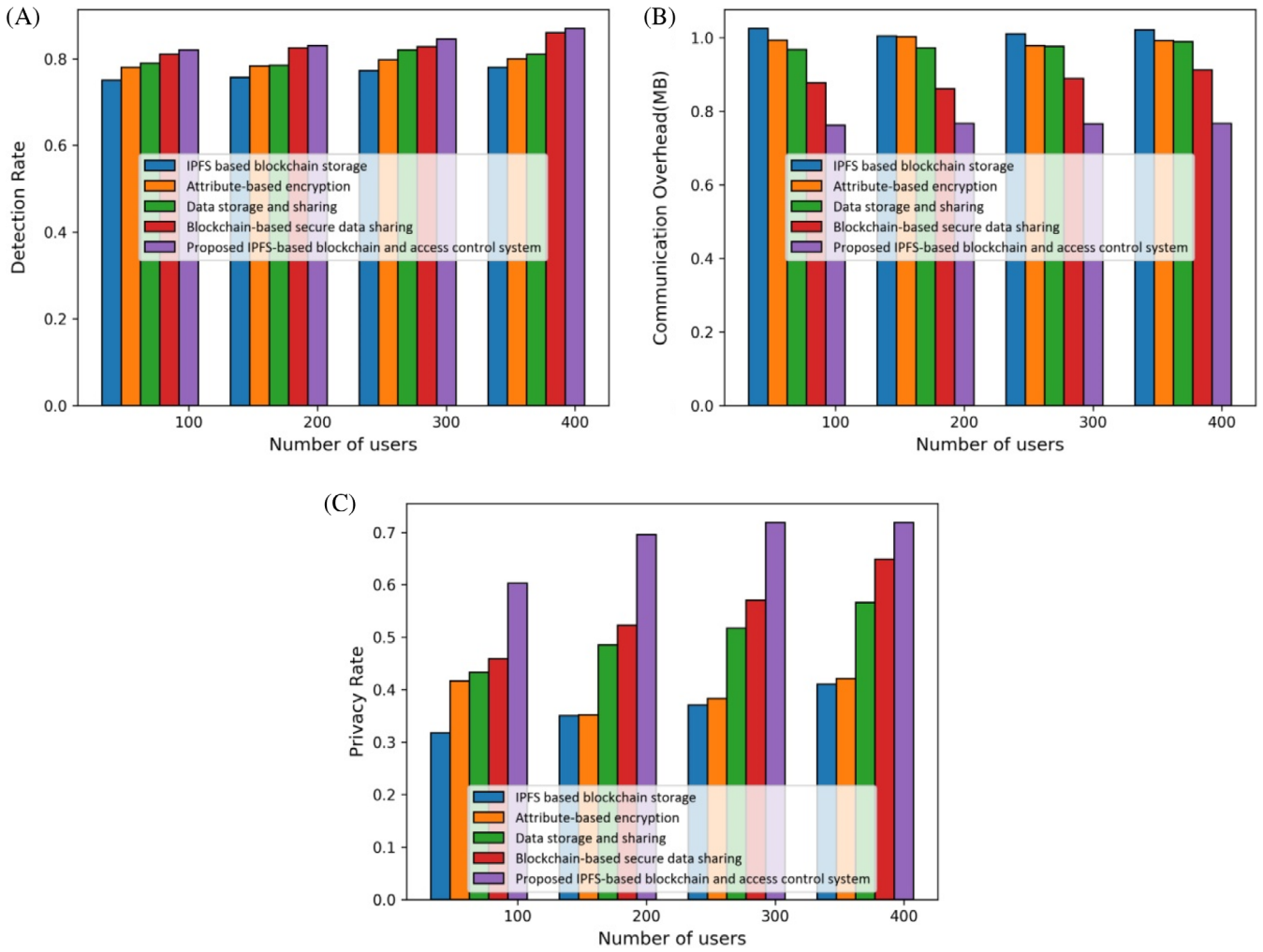

FIGURE 14 Analysis of methods with block chain size = 20 considering (A) detection rate, (B) communication overhead, (C) private rate

based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.750, 0.780, 0.790, 0.810, and 0.820. Likewise, when number of user is 400, the corresponding detection rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.780, 0.800, 0.810, 0.860, and 0.870. The analysis of methods using communication overhead is portrayed in Figure 14B. When number of user is 100, the corresponding communication overhead measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 1.026, 0.993, 0.967, 0.878, and 0.762 MB. Likewise, when number of user is 400, the corresponding communication overhead measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 1.022, 0.992, 0.989, 0.912, and 0.766 MB. The analysis of methods using private rate is portrayed in Figure 14C. When number of user is 100, the corresponding private rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.318, 0.417, 0.433, 0.459, and 0.603. Likewise, when number of user is 400, the corresponding private rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.411, 0.421, 0.566, 0.648, and 0.719.

6.5 | Comparative discussion

With block chain sizes of 5, 10, 15, and 20, Table 5 shows the study of several ways utilizing the detection rate, communication overhead, and private rate. Using block chain size = 5, the detection rate measured by IPFS based block chain storage, attribute-based encryption, data storage and

TABLE 5 Comparative analysis

Block chain size	Metrics	IPFS based block chain storage	Attribute-based encryption	Data storage and sharing	Block chain-based secure data sharing	Proposed IPFS-based block chain and access control system
Block chain size = 5	Detection rate	0.805	0.820	0.830	0.830	**0.835**
	Communication overhead (MB)	1.074	1.050	1.043	0.945	**0.816**
	Private rate	0.330	0.625	0.719	0.766	**0.884**
Block chain size = 10	Detection rate	0.795	0.805	0.820	0.830	**0.845**
	Communication overhead (MB)	1.036	1.028	1.005	0.940	**0.785**
	Private rate	0.371	0.684	0.766	0.813	**0.884**
Block chain size = 15	Detection rate	0.795	0.795	0.815	0.820	**0.855**
	Communication overhead (MB)	1.007	0.993	0.978	0.883	**0.772**
	Private rate	0.431	0.636	0.754	0.813	**0.901**
Block chain size = 20	Detection rate	0.780	0.800	0.810	0.860	**0.870**
	Communication overhead (MB)	1.022	0.992	0.989	0.912	**0.766**
	Private rate	0.411	0.421	0.566	0.648	**0.719**

Bold value indicates the proposed method results.

sharing, block chain-based secure data sharing, and proposed IPFS-based block chain and access control system are 0.805, 0.820, 0.830, 0.830, and 0.835. The minimal communication overhead of 0.816 MB is computed by the proposed IPFS-based block chain and access control system, whereas the detection rate measured by IPFS based block chain storage, attribute-based encryption, data storage and sharing, block chain-based secure data sharing are 1.074, 1.050, 1.043, and 0.945 MB. The maximal private rate of 0.884 is computed by the proposed IPFS-based block chain and access control system, while the privacy rate evaluated by IPFS based block chain storage, attribute-based encryption, data storage, and sharing, and block chain-based secure data sharing are 0.330, 0.625, 0.719, and 0.766. Using block chain size = 10, the detection rate of 0.845, minimal communication overhead of 0.785 MB and maximal private rate of 0.884 are computed by proposed IPFS-based block chain and access control system. Using block chain size = 15, the maximal detection rate of 0.855, minimal communication overhead of 0.772 MB and maximal private rate of 0.901 is computed by proposed IPFS-based block chain and access control system. Using block chain size = 20, the maximal detection rate of 0.870, minimal communication overhead of 0.766 MB and maximal private rate of 0.719 is computed by proposed IPFS-based block chain and access control system.

7 | CONCLUSION

In this paper, an IPFS-based block chain and access control system is devised for performing secure transactions. In addition, this technique comprises five different entities, that is, data owner, smart contract, block chain, data requester, and IPFS. The data owner controls various files over the block chain network and shares data with the requester after proper authentication. The data requester is a user authorized to view some files. Moreover, the developed IPFS based block chain and access control system consists of eight stages: setup, user registration, initialization, data encryption, and storage, authentication, testing, access control, and decryption. The setup phase generates security and public parameters and deploys smart contracts on a block chain. Then, the registration is performed to register the data requester by generating keys. Next, the initialization is performed to record and secure the storage. Then, the encryption of data is done and stored. Then, the authentication is done to the genuineness of user and testing and access control are done to provide access permissions. Finally, the decryption phase is executed for retrieving the data. The proposed IPFS-based block chain and access control system offered effective storage and secure transactions in the block chain network and attained the highest detection rate. The proposed IPFS-based block chain and access control system enhanced performance with the highest detection rate of 0.870, smallest statement overhead of 0.766 MB, and privacy rate of 0.719. In future, the model can be extended to deploy in real networks. In addition, the deployment of huge scale solutions can be a comprehensive study in the future.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1. Rajput AR, Li Q, Ahvanooey MT, Masood I. EACMS: emergency access control management system for personal health record based on block chain. *IEEE Access*. 2019;7:84304-84317.
2. Kumar R, Tripathi R, Implementation of distributed file storage and access framework using IPFS and block chain. Proceedings of 2019 Fifth International Conference on Image Information Processing (ICIIP); 2019; pp.246-251.
3. Zheng Z, Xie S, Dai H, Chen X, Wang H, An overview of block chain technology: architecture, consensus, and future trends. Proceedings of 2017 IEEE International Congress on Big Data (BigData Congress); 2017; pp.557-564.
4. Benchoufi M, Ravaud P. Block chain technology for improving clinical research quality. *Trials*. 2017;18(1):335.
5. Iansiti M, Lakhani KR. The truth about block chain. *Harv Bus Rev*. 2017;95(1):118-127.
6. Toapanta SMT, Gallegos LEM, Baldeon PO, Triviño FDT. Block chain analysis applied to a process for the national public data system for Ecuador. Paper presented at: 2020 3rd International Conference on Information and Computer Technologies (ICICT); 2020.
7. Gursoy ME, Inan A, Nergiz ME, Saygin Y. Privacy-preserving learning analytics: challenges and techniques. *IEEE Trans Learn Technol*. 2016;10(1):68-81.
8. Zhang X, Chen X. Data security sharing and storage based on a consortium block chain in a vehicular ad-hoc network. *IEEE Access*. 2019;7:58241-58254.
9. Sun J, Yao X, Wang S, Wu Y. Non-repudiation storage and access control scheme of insurance data based on block chain in IPFS. *IEEE Access*. 2020;8:155145-155155.
10. Wang S, Zhang Y, Zhang Y. A block chain-based framework for data sharing with fine-grained access control in decentralized storage systems. *IEEE Access*. 2018;6:38437-38450.
11. Ahila SS, Shunmuganathan KL. Role of agent technology in web usage mining: homomorphic encryption based recommendation for e-commerce applications. *Wirel Pers Commun*. 2016;87(2):499-512.
12. Benet J. Ipfs-Content Addressed, Versioned, p2p File System. 2014; ArXiv preprint arXiv:1407.3561.
13. Alessi M, Camillo A, Giangreco E, Matera M, Pino S, Storelli D. Make users own their data: a decentralized personal data store prototype based on ethereum and IPFS. Proceedings of 2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech); 2018; pp.1-7.
14. Zhu J, He P, Zheng Z, Lyu MR. A privacy-preserving QoS prediction framework for web service recommendation. Proceedings of IEEE International Conference on Web Services; 2015; pp. 241-248.
15. Kuang L. A privacy protection model of data publication based on game theory. *Secur Commun Netw*. 2018;2:1-13.
16. Naz M, Al-zahrani FA, Khalid R, et al. A secure data sharing platform using block chain and interplanetary file system. *Sustainability*. 2019;11(24):7054.
17. Mandala J, Dr MVP, SekharaRao C. HDAPSO: enhanced privacy preservation for health care data. *J Netw Commun Syst*. 2019;2(2):10-19.
18. Memon I. Authentication user's privacy: an integrating location privacy protection algorithm for secure moving objects in location based services. *Wirel Pers Commun*. 2015;82(3):1585-1600.
19. Ali G, Ahmad N, Cao Y, Asif M, Cruickshank H, Ali QE. Block chain based permission delegation and access control in internet of things (BACI). *Comput Secur*. 2019;86:318-334.
20. Veeramanickam MRM, Mohanapriya M, Pandey BK, et al. Map-reduce framework based cluster architecture for academic student's performance prediction using cumulative dragonfly based neural network. *Clust Comput*. 2019;22(1):1259-1275.
21. Ding S, Cao J, Li C, Fan K, Li H. A novel attribute-based access control scheme using block chain for IoT. *IEEE Access*. 2019;7:38431-38441.
22. Sun J, Yao X, Wang S, Wu Y. Block chain-based secure storage and access scheme for electronic medical records in IPFS. *IEEE Access*. 2020;8:59389-59401.
23. Battah AA, Madine MM, Alzaabi H, Yaqoob I, Salah K, Jayaraman R. Block chain-based multi-party authorization for accessing IPFS encrypted data. *IEEE Access*. 2020;8:196813-196825.
24. Lin C, He D, Huang X, Choo KKR, Vasilakos AV. BSeIn: a block chain-based secure mutual authentication with fine-grained access control system for industry 4.0. *J Netw Comput Appl*. 2018;116:42-52.
25. Baiod W, Light J, Mahan A. Block chain technology and its applications across multiple domains: a survey. *J Int Technol Inform Manag*. 2021;29(4):78-119.
26. Patwary AA-N, Sudheer AF, Battul K, Naha RK, Garg S, Mahanti A. Fog Auth chain: a secure location-based authentication scheme in fog computing environments using block chain. *Comput Commun*. 2020;162:212-224.

How to cite this article: Antony Saviour M, Samiappan D. IPFS based file storage access control and authentication model for secure data transfer using block chain technique. *Concurrency Computat Pract Exper*. 2023;35(2):e7485. doi: 10.1002/cpe.7485