Blockchain based hierarchical semi-decentralized approach using IPFS for secure and efficient data sharing

Smita Athanere\textsuperscript{a,}\textsuperscript{*}, Ramesh Thakur\textsuperscript{b}

\textsuperscript{a}Computer Engineering, IET, DAVV, Indore, MP, India
\textsuperscript{b}International Institute of Professional Studies, DAVV, Indore, MP, India

\section*{ABSTRACT}

Nowadays, cloud servers are gathering an increasing amount of data. Data is commonly stored on cloud servers in the form of ciphertext to protect security and concealment of data. When a consumer requests access to encrypted data, a third party must provide an access key. The system’s security, however, will be compromised if the third party or internal personnel are dishonest. To address this issue, a novel blockchain-based secure decentralized system using IPFS is proposed in this research for secure data transfer. Because all participants of system model are recorded on the chain, and the continuously extending chain makes it conditionally difficult to modify any block without being detected, a blockchain-based system is often regarded as a safe platform. In the proposed approach, the data owner uploads an encrypted file to IPFS, which is subsequently separated into \( n \) secret sections called hash codes for data security. The data owner must additionally write the access permissions in order to achieve access to this secure data. For security, the system uses two-level key management: first, the data owner encrypts the file, and then the IPFS server makes a hash code of that encrypted file. The proposed solution, which employs blockchain technology, enables consumers to be handled across several domains, erases the single-point failure in traditional centralized systems, and overhead related to communication and computation are decreased at the consumer level. According to the security analysis, the proposed system might effectively resist single and collaboratively malicious persons, as well as untrustworthy cloud servers.

\textcopyright 2022 The Author(s). Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

\section*{1. Introduction}

While cloud companies put a lot of endeavor into making data sharing as convenient as possible, there are still a lot of security concerns to be addressed. One of the most serious things is that cloud service providers aren’t always reliable. They will be vulnerable to not just external attacks by adversaries, but also inside attacks by hostile employees, providing serious data security risks for cloud data sharing (see Table 1).

Encryption technology can be considered a security assurance in order to accomplish data access authority. On the other hand, access authority for encrypted data is a major challenge. Numerous access patterns with hierarchical links are widely employed for data sharing in the cloud. ‘Ciphertext-policy attribute-based encryption (CP-ABE),’ on the other hand, is one of the best solutions for delivering safe data access from cloud storage (Xue et al., 2019). In the majority of existing ‘CP-ABE’ schemes, all consumers must have faith in a single authority (Hao et al., 2019; Li et al., 2019). It is easier to build a failure on single point, and in practical implementations, attributes are often dispersed among numerous trust domains and organizations. Chase provides the first multi-authority CP-ABE solution, in which different authorities govern many discontinuous domains to recognize that consumer’s attributes are issued across several authorities (Chase, 2007). Despite the fact that several extensions of centralized systems have been proposed, the single-point of failure issue still exists (Lewko and Waters, 2011; Yang et al., 2013; Sandor et al., 2019). The entire attribute collection is divided into many distinct subsets, each governed by a separate authority, enabling these schemes to refer to centralized systems.
Data is increasingly being stored on cloud services. To preserve privacy and security of data, it is usually stored on the cloud server in the form of ciphertext. When a user requests for access to encrypted data, a third-party access key is required. The system's security will be affected if the third party is untrustworthy. To solve this problem, we may use decentralized system using Blockchain to create a multi-authority basis access control scheme for data sharing (Lyu et al., 2020). Blockchain technology is characterized by decentralization, openness, autonomy, and independence from a trusted third party. It makes use of encryption technology to preserve anonymity, and it is generally safer than standard data storing methods. A lot of promise exists when blockchain technology is combined with cloud data sharing systems (Gai et al., 2020). Currently, blockchain technology may be utilized to address issues such as secure data access via a decentralized multi-authority framework. Almost all applications require access control as a security feature. Because of its unique properties such as immutability, durability, auditability, and dependability, blockchain is being considered as a supplement to access control solutions. One of the key reasons for the rise in popularity of blockchain was its ability to provide transactional transparency as well as support transactional privacy and secure data. On all peers, the blockchain retains all recorded transactions and data. Block chain is being explored as a supplement to safe access control systems due to its unique qualities such as immutability, stability, traceability, and reliability. It keeps track of all transactions and data that have been logged across all peers. The autonomy and security of the blockchain, as well as its new application in data transfer and sharing, inspired this research (Yang et al., 2018). We could leverage blockchain and IPFS to create a decentralized multi-authority system, encode collective computing processes among various authorities from diverse attribute domains into contracts, and provide decryption tokens to consumers (Zhu et al., 2019, Chen et al., 2019). Because blockchain transactions are public, it’s necessary to think about how to include chain algorithms into the privacy of a blockchain basis access control strategy (Wei et al., 2020).

1.2. Challenges

Despite the potential benefits of blockchain technology, it still faces significant difficulties that limit its utility applicability in security. In this part, we’ll go over a few of these issues and how they relate to data sharing in the cloud security applications (Rouhani and Deters, 2019).

i) Communication Overhead: There will be a serious increase in network traffic and system processing capacities because blockchain has connection between peer to peer. Transactions and blocks must be broadcast rather than unicast in traditional systems. As a result, the network's overhead has increased significantly, posing a substantial difficulty.

ii) Scalability: It is thought that blockchain technology will scale better than old centralized methods. As the number of consumers and networking nodes grows, the technology performs badly.

iii) Single point of failure: The majority of relevant studies focused on this problem, which is inherent in centralized systems, because classic access control solutions are all cen-

<table>
<thead>
<tr>
<th>Reference</th>
<th>Work</th>
<th>Access control scheme</th>
<th>Blockchain Technology/platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>(He et al., 2020)</td>
<td>Access Control for sharing the data</td>
<td>Attribute basis Hierarchical scheme</td>
<td>–</td>
</tr>
<tr>
<td>(Wang et al., 2019)</td>
<td>Access Control framework for sharing the data</td>
<td>Attribute basis Encryption scheme</td>
<td>Ethereum platform</td>
</tr>
<tr>
<td>(Zhang et al., 2018.)</td>
<td>Privacy guaranteeing and user controlled data exchange in IoT devices</td>
<td>Fine grained and attribute basis cryptosystem</td>
<td>Smart Contract based on Ethereum platform</td>
</tr>
<tr>
<td>(Qin et al., 2021)</td>
<td>Access control strategy for safe data exchange in cloud</td>
<td>Multi-authority basis and Shamir secret sharing scheme</td>
<td>Hyper ledger Fabric platform</td>
</tr>
<tr>
<td>(Zuo et al., 2021)</td>
<td>Sharing secure data in cloud without intervention of any trustworthy parties</td>
<td>Cipher text policy attribute basis encryption scheme</td>
<td>Block secure technology</td>
</tr>
<tr>
<td>(Zhu et al., 2018)</td>
<td>Management of digital asset using transaction basis access control for data sharing</td>
<td>Secure attribute basis scheme</td>
<td>Bit coin</td>
</tr>
<tr>
<td>(Maesa et al., 2019)</td>
<td>Auditable access control system using blockchain</td>
<td>General access control scheme</td>
<td>Ethereum platform</td>
</tr>
<tr>
<td>(Gao et al., 2020)</td>
<td>Trustworthy secure data exchange</td>
<td>Fine grained basis scheme</td>
<td>–</td>
</tr>
<tr>
<td>(Zhu et al., 2019)</td>
<td>Model for cloud data management</td>
<td>General access control scheme</td>
<td>Ethereum platform</td>
</tr>
<tr>
<td>(Paillisse et al., 2019)</td>
<td>Access control with Multi-administrative domain</td>
<td>Distributed Access control</td>
<td>Hyper ledger Fabric platform</td>
</tr>
<tr>
<td>(Guo et al., 2019)</td>
<td>Multi-authority access control</td>
<td>General attribute basis scheme</td>
<td>Blockchain with Smart contracts</td>
</tr>
<tr>
<td>(Sandor et al., 2019)</td>
<td>Effective decentralized system for mobile cloud data storage</td>
<td>Multi-authority basis scheme</td>
<td>–</td>
</tr>
<tr>
<td>(Stanciu, 2017)</td>
<td>Edge Computing</td>
<td>–</td>
<td>Hyper ledger Fabric platform</td>
</tr>
<tr>
<td>(Jemel and Serrouchni, 2017)</td>
<td>Data Exchange</td>
<td>General attribute basis encryption</td>
<td>Multi-chain technology</td>
</tr>
<tr>
<td>(Steichen et al., 2018)</td>
<td>To provide access control for file sharing</td>
<td>Modified Interplanetary File System (IPFS)</td>
<td>Ethereum Platform</td>
</tr>
<tr>
<td>(Wang et al., 2018)</td>
<td>A decentralized access control mechanism for data sharing</td>
<td>Attribute basis encryption scheme and Interplanetary File System (IPFS)</td>
<td>Ethereum Platform</td>
</tr>
<tr>
<td>(Sun et al., 2020)</td>
<td>Ciphertext basis encryption scheme for data access</td>
<td>IPFS based scheme</td>
<td>Generic blockchain technology</td>
</tr>
</tbody>
</table>
tralized. To solve this difficulty, relevant researches have used distributed access control using Blockchain technology and a variety of other strategies.

iv) Security: Another key difficulty that every access control system should face is security. However, throughout history, advancements in attack vectors and structure have occurred. However, whichever technology it is combined with, Blockchain technology provides security as an inherent quality. Although, to reach the maximum levels of security in a system, encryption procedures are used.

v) Management & Delegation of Access rights: Delegation of access rights, their administration, and authorization are also significant aspects of access control systems. It is critical to underline that authorized entities’ access to a certain resource is at the heart of access control systems. Despite the fact that every access control system is expected to address this issue, relevant research have focused on it.

1.3. Problem statement

Cloud servers, which are centralized authorities, keep an enormous amount of data. A central authority is associated with a variety of dangers, including single point failure. To avoid such failures, a multi-authority hierarchical architecture is built, with central authorities divided into domain-based authorities. As a result, we can partially address the single point failure issue. Because all user characteristics are handled by a single central authority, it’s trivial to construct a single point of failure in most existing attribute-based encryption solutions. As a result, a number of multi-authority CP-ABE methods for handling user attributes have been presented. However, these systems do not eliminate the particular failure point in the sense that they do not eliminate the significant processing and communication overhead that data users confront. To address these issues, we offer a decentralized system based on Blockchain and IPFS, which allows many authorities to access data from cloud storage in an efficient and secure manner.

Blockchain technology is used to keep a permanent record of data owner-user agreements. A user requests data from the owner after agreeing to the blockchain’s terms and conditions. Using blockchain, the suggested method offers research data transparency, accountability, and fine-grained granularity.

1.4. Our contribution

In this research, we present a centralized, decentralized and semi-decentralized approach for multi-authority cloud access. The traditional approach is used to create the centralized system, which does not include blockchain. Some disadvantages of a centralized system exist, such as single point failure and data that can be modified by lower level authorities in some instances. As a result, we propose adopting a blockchain and IPFS based decentralized and semi-decentralized multi-authority architecture to provide secure and efficient cloud access. Then compare the centralized, decentralized and semi-decentralized systems. Because there is no trustworthy third party and the scheme is constructed in a decentralized fashion, users can securely and efficiently retrieve data from cloud servers using blockchain and IPFS. Here are our contributions to the development of this scheme:

i) Initially, a centralized system for cloud data access via multi-authorities in a hierarchical way is designed.

ii) A decentralized access control mechanism is developed that combines blockchain technology with an interplanetary file system. This technique can solve the concerns of a single point failure, expensive processing and communication overhead for data users. Because blockchain has the properties of traceability and auditability, it can also track data access transactions and keep track of log files at all times.

iii) We have also designed the Semi-decentralized system. In our decentralized blockchain-based system, we expand the traditional hierarchical multi-authority attribute-based encryption approach and build encryption decryption techniques to realize multi-authority based cross-domain collaboration.

iv) The suggested scheme was tested for performance and security, as well as the efficacy of the proposed scheme in terms of computation & communication overhead on the user side.

The rest of the research paper is structured as follows. Section 2 summarizes the numerous studies on blockchain and IPFS basis access control for cloud data sharing. Section 3 introduces basic preliminaries that must be included in any proposed project. The suggested blockchain and IPFS based network model and algorithm for sharing the data in the cloud using blockchain and IPFS are detailed in Section 4. Section 5 contains the proposed work and performance evaluation. Finally, in Section 6, we bring our research to a conclusion.

2. Literature survey

This section discusses the possibility of using blockchain technology to control data access in the cloud server. Several studies on multi-authority attribute based encryption mechanisms and blockchain-based data sharing access control mechanisms have been already published. The tamper-proof and auditable nature of blockchain access records is due to its security features. As a result, the following literature is offered on existing and recent blockchain and IPFS protocol based systems for access control.

For sharing the data in cloud, He et al. devised an ‘attribute based hierarchical access control scheme (AHAC)’. In this scheme, if any attribute as a data visitor is matched with the access control structure, the data associated with this structure can be decrypted by that attribute. The authors claim that this approach is more competent and secure than ‘CP-ABE (Ciphertext-Policy Attribute-Based Encryption)’ scheme after conducting numerous experiments. In addition, when the number of encrypted files of data was increased, AHAC’s speed improved significantly (He et al., 2020). Wang et al. suggested access control architecture for sharing the data in secure cloud server using blockchain. In this framework, they merged two techniques, ‘ciphertext-policy attribute-based encryption (CP-ABE)’ and Ethereum based blockchain. The proposed cloud access control framework is developed in a decentralized manner, with no trusted third party involved. It has three primary characteristics. Initially, smart contracts can be used by data owners on the Ethereum blockchain to store ciphertext. Second, the data owner can specify permissible access times for decrypting ciphertext. Finally, the trace function is used to track each smart contract’s invocation (Wang et al., 2019). With ‘ABS (Attribute based scheme)’ and ‘CP-ABE (ciphertext-policy attribute-based encryption)’ schemes, Zhang et al. suggested a blockchain-based architecture for privacy preservation and user controlled base data exchange with fine grained access control in the Internet of things (IoT). It gathers data from IoT devices and sends back to the cloud for processing or other operational purposes. However, due to data leaks and privacy concerns, the cloud may not be a completely trustworthy entity. For security and fault tolerance, the suggested privacy-preserving architecture used a blockchain and an attribute-based cryptosystem (Zhang et al., 2018; Cha et al., 2018).
For secure data sharing, Qin et al. presented a multi-authority data access control mechanism on the basis of blockchain. The multi-authority concept is used to avoid single point failure, and secure data access is possible owing to blockchain technology. Shamir secret sharing and Hyper ledger Fabric blockchain technologies are used to implement this method. Furthermore, they leverage blockchain based technology to build faith among different management authorities, as well as share the smart contracts using tokens among the attributes and management authorities. It will decrease user’s communication and computation overhead. (Qin et al., 2021).

Zuo et al. suggested a blockchain-based cipher text policy attribute based encryption technique for exchanging secure data in the cloud beyond the involvement of any trusted third parties. They can secure the rights and security of data owners in this scheme by using blockchain technology. The authors compared the proposed scheme to existing cloud security schemes and discovered several advantages, including: i) Only the data owners have the authority to determine who has access to the data; ii) Trace every record; iii) Scheme has ‘one-to-many’ encryption; iv) There are no third-parties involved in the scheme; v) They used a discrete logarithm problem to show that the suggested system is secure (Zuo et al., 2021). For data sharing, Zhu et al. suggested transaction basis access control architecture. It is a hybrid combination of the ‘attribute based access control (ABAC)’ scheme and blockchain technologies. In this approach, four types of transactions are used with respect to four phases as an access control mechanism, these four transactions are object escrow, subject registration, access request and grant. These are combined into an algorithm that allows for flexible and differing permission management as well as the processing of verified transparent access in an open decentralized environment (Zhu et al., 2018).

Maesa et al. developed a blockchain based strategy for an auditable access control system. The authors used block chain’s transactions to build, edit, and cancel policies in this manner. Furthermore, A transaction can be used to transfer a resource owner’s access entitlement from one user to another. Then, any user can review the access control history by auditing these transactions. Transaction auditability is possible because of the immutability and transparency aspects of blockchain. To test the scheme’s performance, the authors provided a reference implementation based on XACML rules and Solidity-authored smart contracts deployed on the Ethereum network (Maesa et al., 2019).

Gao et al. devised the ‘Trust access’ approach, which is based on a secure ciphertext policy and blockchain-based access management. Blockchain is employed for traceability and accountability in this system. With the use of smart contracts, a data user proves its authority. The data owner then transmits the access transaction record to the block chain’s access control mechanism and offers the user a secret key to decrypt the ciphertext. Rather than writing an unique contract for each combination of data owners and users, the proposed solution allows them to implement access control between different consumers and data owners using four smart contracts (Gao et al., 2020). Zhu et al. proposed a cloud data management strategy based on controllable and trustworthy blockchain technology. It can handle any data storage in the cloud situation including a lack of control. In this technique, the trust authority node allows one to terminate any powerful damaging actions even under a majority attack. The proposed approach has been tested for security and performance, as well as its utility (Zhu et al., 2019). Paillisse et al. exploited blockchain to implement multi authority based distributed access control mechanism. They employ a permission blockchain implementation to broadcast access control policies in a secure manner. Network administrators define group-based policies here, which are then stored in the blockchain via transactions. Routers utilize blockchain to analyze if user is authorized to access a system before establishing a safe connection with the resource consumer (Paillisse et al., 2019).

Sandor et al. proposed a decentralized multi-authority attribute based technique for mobile cloud data storage. It eliminates the central authority problem with CP-ABE, eliminating the need for worldwide user identification, and reduces communication overheads on the user side with cloud user assistance. During interactions with numerous attribute authorities, it acts as a gateway for the mobile user (Sandor et al., 2019).

Alizadeh et al. introduced a decentralized system based on IPFS and blockchain that employs ‘distributed hash table (DHT)’ technology to store information with high reliability and minimal storage costs. By lowering overall processing time on the blockchain and establishing an agreement service that can communicate with the blockchain via a Restful API, efficiency was increased with the usage of DHT technology. They demonstrated the usefulness of the proposed system, concluding that integrating IPFS and blockchain allows for efficient encrypted storage, a stable record, and overall better efficacy in a decentralized manner (Alizadeh et al., 2020). Using Blockchain and IPFS, Naz et al. established a decentralized architecture for secure data sharing. The owner uploads meta data to an IPFS server, which is subsequently partitioned into secret shares. The proposed approach achieves security and access control by applying the owner’s access roles defined in smart contracts. Ethereum blockchain, encryption, decryption, smart contracts, and various incentive systems make up the proposed secure and decentralized framework. Transparency, confidentiality, access control, owner legitimacy, and data sharing quality are all achieved (Naz et al., 2019). Gao et al. introduced a safe data sharing system for personal data with fine-grained data access. Blockchain technology, ‘ciphertext-policy attribute-based encryption (CP-ABE)’, and the ‘Inter Planetary File System (IPFS)’ were combined in the proposed approach. The data owner encrypted the sharable material and stored it on IPFS, emphasizing the system’s decentralization. Only the data user whose attributes satisfy the access policy can download and decrypt the data. Finally, the authors evaluated the suggested scheme’s security and performance and concluded that it was practical (Gao et al., 2021).

Javed et al. introduced a safe data exchange approach for vehicle networks using blockchain technology. In this technology, the large set of data produced by smart vehicles which is stored in a distributed file storage system known as the Interplanetary File System (IPFS). It is utilized to solve problems with data storage in centralized systems, such as data tampering, privacy breaches, hacker vulnerabilities, and so forth. In this scheme, smart contracts are used to automate system procedures without involving a third party and to validate the reviews provided to edge nodes (Javed et al., 2020).

Daniel and Tschorsch researched several papers and compiled a survey on IPFS and its Friends. They concentrated on the next-generation data network’s overview. They employed a variety of data networks to introduce general concepts and highlight fresh advances. The authors went over the Interplanetary File System in greater detail, as well as Swarm, the Hypercore Protocol, Storj, and SAFE in general (Daniel and Tschorsch, 2021).

Gutub et al. discussed the secret sharing where more than one person is interested in single resource at a time. They proposed a work that based on parallel counting of ones, in the contribution which is responsible for secret generation. In this they suggested two models for secret sharing and their advantages and disadvantages (Gutub et al., 2019).

Qurashi et al. proposed a promising approach i.e. based on counting based secret sharing for multi user authentication. Generation of secret key is done by simple flipping operation of one or two 0-bits within that key at different locations. Reconstruction of secret key is done by (n, k) threshold in parallel. Security Enhancement is achieved by modification of share generation is
done by dividing the share into \( m \) different blocks (Gutub and Al-Qurashi, 2020).

Gutub et al. watermarking images via counting based secret sharing for lightweight semi complete authentication. In this paper utilization of counting-based secret sharing strategy to permit validation of ownership RGB watermarking even if some of the image-file is interfered. We validate image watermarking partially as lightweight semi-complete verification, which is not possible in the normal random-stream watermarking schemes. This work is tested and compared with other state of art methods (Gutub, 2022).

Almehmadi et al. proposed a novel Arabic e-text watermarking supporting partial dishonesty based on counting based secret sharing. In this paper, they are using the extension Arabic character "‘Kashida’" for hiding the watermarking data. The process of hiding watermarking bits is responsible for full security, which is not responsible to hide the secrets directly as previous references. Password is required to initiate counting-based secret sharing to generate the stream shares of watermarking hidden bits to be embedded (Almehmadi and Gutub, 2021; Gutub, 2021).

Al-Shaarani et al. suggested a scheme named Securing matrix counting-based secret-sharing involving crypto steganography. In standard secret sharing method, the key is distributed among several authorized participants in a way that only the intended groups of them are needed to reconstruct the original key. These schemes became popular in security of both cryptography and steganography and accomplished extraordinary results combined with each. This focuses on two particular secret sharing techniques known as counting-based secret sharing and matrix-based secret sharing, which is based on the former. These methods are simple and intuitive (Al-Shaarani and Gutub, 2021a, 2021b).

3. Preliminaries

We will lay out some background information in this section, which will serve as the foundation for the proposed model. These fundamentals are based on bilinear maps, the integration of blockchain into a multi-authority access system and interplanetary file system (IPFS).

3.1. Bilinear pairing

The bilinear pairing creates three multiplicative cyclic groups of the same prime order \( P: G_1, G_2, \text{and } GT \). Let \( g_1 \in G_1 \) and \( g_2 \in G_2 \) two generators for \( G_1 \) and \( G_2 \) respectively. A bilinear pairing relationship \( e: G_1 \times G_2 \rightarrow GT \) is defined on the three groups \( G_1, G_2, GT \) and it satisfies three properties as follows:

(i) **Bilinearity:** \( \forall g_1 \in G_1, g_2 \in G_2, a, b \in \mathbb{Z}_p, e(g_1^a, g_2^b) = e(g_1, g_2)^{ab} \).

(ii) **Nondegenerate:** \( \exists g_1 \in G_1, g_2 \in G_2, e(g_1, g_2) \neq 1 \), which signifies that no pair in \( G_1 \times G_2 \) is sent to the identity in \( GT \) group.

(iii) **Efficient Computability:** For the given two items \( \forall g_1 \in G_1, g_2 \in G_2 \), at least one existing algorithm can effectively compute \( e(g_1, g_2) \).

Along with bilinear pairing, bilinear pairing generator (BPG) is defined as a probabilistic algorithm which uses a security parameter \( \lambda \) as its input, and some parameters in output as \( [g_1, g_2, G_1, G_2, e, q] \), where \( g_1 \) and \( g_2 \) are two generators, \( e \) is a bilinear map, \( q \) is a \( \lambda \)-bit prime number, \( G_1, G_2 \) are two cycle groups of the same order \( q \), \( G_1 \) is multiplicative group, \( G_2 \) is additive group, \( G_1 \) is multiplicative group.

3.2. Blockchain

Blockchain technology allows for the establishment of an unchangeable, decentralized, always accessible, safe, and publicly accessible data repository. A distributed consensus protocol is used to manage this repository in a distributed manner. The level of trust connected with write and read operations varies depending on the type of blockchain. A write operation is the capacity to update the ledger, that is, to add new content to it, whilst the ability to access the existing contents is referred to as a read operation. Because any trustless entity can read a blockchain, it is referred to as public. They’re called permission less because they allow any untrustworthy entity to write (Sankar et al., 2017; Stanciu, 2017; Xu et al., 2017). Blockchain technology was initially introduced to support crypto-currencies in the past (Cachin, 2016). The blockchain is utilized as a public ledger to store transactions that transfer value between entities in this situation. Despite the fact that the bit coin crypto currency system was the first to use blockchain, a slew of new ideas have developed since then. In a blockchain-based access control system, access logs are stored as a sequence of transaction blocks, with each block linked to the one before it by the hash value. (Nakamoto, 2008; Yli-Huumo et al., 2016).

3.3. Interplanetary file system (IPFS)

The Interplanetary File System (IPFS; Benet, 2014) is a set of sub protocols that make up a peer-to-peer distributed file system that uses a DHT (distributed hash table) to keep track of everything, including who owns what data. Furthermore, it offers a novel approach for file sharing in a decentralized environment. Hash tables are used by the IPFS to store data packages. Kademlia is used by the IPFS to figure out which nodes contain which data. Petar Maymounkov and David Mazieres created Kademlia in 2002 as a DHT for decentralized peer-to-peer computer networks (Maymounkov and Mazieres, 2002). In the IPFS, a unique hash is the consequence of preserving data without regard for its size. IPFS can save the hash, which can later be used by parties to retrieve data. The data will be broken into numerous little bits when it is ready to be added to the IPFS network. Each piece has its unique hash to identify it. The chunks will then be dispersed over the network to nodes with hashes that are closest to peer Id. The DHT is used to travel to nodes where the hash is existent when a user requests a chunk. After inspecting all of the existing chunks, the main object is simply concatenated from all of the chunks. The distributed part of DHT, on the other hand, means that the complete table is shared over multiple places. DHT-based distributed systems are completely connected systems with no substantial differences between participants that are located in various locations. As a result, everyone has access to the most up-to-date hash table.

4. Proposed work

The proposed work is implemented and experimented in three phases in which first phase is centralized server system, second is decentralized system and third one is semi-decentralized system. In recent scenario, cloud providers are offering lot of services for data storage and access using centralized server system. In centralized system, data owner uploads the data on main server and data user access the data from server. However, one important concern is that cloud service vendors are not completely trustworthy. They are vulnerable to both external and internal attacks from malevolent personnel, providing substantial data security issues for cloud data sharing. Amazon employees, for
example, sell consumers’ sensitive information to third parties for selfish enrichment, and cloud storage companies routinely eavesdrop on users’ personal information. To address the aforementioned issue, cryptographic access control techniques are presented in centralized system to provide safe data sharing by performing encryption using secret key that can merely be decrypted by individuals who have that secret key. However, in this model, all members are required to have faith in a single authority which might lead to a single point of failure, and important parameters are typically scattered among several trust domains and organizations in actual implementations. Fig. 1 depicts the centralized system model. The data owner, data user, central authority, and domain authority are all depicted in this figure. Central authority is the main server where data owner uploads the data and domain authorities are the domain wise partitions of server in hierarchical manner which have complete responsibility of users of that domain. Data user of specific domain can fetch the data from its domain.

In centralized system two major issues are identified, first is data manipulation by internal employees and second is single point failure. Block chain based decentralized system is proposed to overcome the issues such as single authority failure, data manipulation by internal employees, communication overhead and excessive processing overhead at consumer side. The block chain can keep track of data access logs which allows for auditable access control management. In this paper secure and proficient data transfer mechanism is proposed using block-chain based decentralized and semi-decentralized multi-authority access concept. The concept of decentralized block chain is implemented through IPFS protocol. The system model of decentralized system is mentioned in Fig. 2. In decentralized system, data owner uploads the data on cloud server and block chain creates the hash code of that encrypted data. The hierarchical structure is created on block chain in which domain authorities are created with its user distribution. The users of one domain can access and keep each other’s copy of data and other domain users can not access and share the data. When data user wants to access the data uploaded by owner, user can access from its domain authority and user can decrypt the original content on its local machine using decryption engine. In this system, two level key management techniques have been proposed, first at the time of encryption on data owner side then second is at block chain by creating hash code. The users who have decryption key can only access the data, so the communication overhead at user side is minimized or stable. The domain users only keep the copy of same domain data which also reduces the storage overhead of replicas.
Fig. 2. IPFS based decentralized system.

Fig. 3. Hierarchical structure and IPFS based semi-decentralized system.
4.1. The system model

This section explains the system model, which consists of six different entities. In order to comprehend the next theoretical algorithms, the syntax of the procedure flow is also introduced.

**Administrator**: The administrator is in charge of setting up system parameters and initializing the system.

**Domain authority (DA)**: DA is in charge of granting credentials to data users through IPFS. Each DA consists of several users in one domain.

**Cloud service provider (CSP)**: Data owners can use the CSP to store their data.

**Data owner (DO)**: CSP is not trustworthy to upload files, so that DO uses encryption to regulate data access. DO creates access policies, encrypt data files prior to submitting them to the vendor of cloud, and submit the key and data ciphertext to IPFS individually.

**IPFS**: All of the above entities are part of an IPFS network. IPFS generates the separate hash code for data and key ciphertext uploaded by data owner. To preserve its integrity and immutability, IPFS saves public parameters and access metadata. It also helps businesses do partial trusted computing and allows many DA to manage user credentials collectively.

**Data user (DU)**: The CSP provides free ciphertext downloads to all DUs. User must demand for key information from the IPFS for decryption, if he or she wishes for data accessibility. The decryption permissions of users from various domains are different. Only if the customer’s credentials match the authentication scheme encoded in the ciphertext can the decryption be completed successfully. For this reason, the decryption engine is linked to authorized users.

4.2. Algorithm

The block chain based decentralized hierarchical system is composed of the following algorithm. The section provides an overview of the procedures in Fig. 4. At the end of this section script based algorithm is also shown.

① The administrator creates and uploads system initialization parameters to the IPFS server initially.

② Users build private and public key combinations and submit public keys to IPFS when they enter into the system.

③ For encryption, the data owner utilizes IPFS public parameters that have been initialized. Data F is encrypted by the DO, who acquires public keys from IPFS in order to do so.

④ DO encrypts the file and transfers it to the IPFS server using the encrypted file Evk (F). IPFS generates the hash code HC (ID) of the file F before submitting it to the CSP.

⑤ Data Owner “DO” saves the ciphertext of the encrypted article or file as well as the file address provided by the cloud server.

⑥ The DU makes an access request to its DA, who then uploads it to the IPFS server.

⑦ IPFS finds the credentials of user and authenticate for accessibility.

⑧ DO keeps authenticated DU’s secret key in the IPFS server after encrypting it.

⑨ The DU obtains the key information and ciphertext of the key from IPFS for final decryption.

⑩ DU retrieves his secret key ciphertext by downloading an encrypted file from an IPFS server. DU decrypts the ciphertext off-chain using decryption engine and gets the original file on its local machine.

**Setup** (1 k, AT) → (PK, MK): To conduct the setup procedure, DO employ the security parameters defined by k and the universal set designated by U of features as inputs. As a result of implementation of algorithm, the public key denoted as “PK” and the master key denoted as “MK” are produced. Fig. 4 shows step ①② of the process. The file F with the file ID is encrypted using the SHA256 encryption method and logged as Evk (F) when DO transmits the encrypted file to the cloud server (where vk denotes for “encryption key”). The SHA256 hash technique is used to hash the provided file name ID to HC (ID). The IPFS node’s address, file ID, and encrypted file Evk (F), as well as the file ID hash HC, are subsequently packaged and delivered to the IPFS server (ID). DO keep note of the file path provided by the IPFS server, as seen in Fig. 4 step ③④.

**Encrypt** (PK, vk, 0) → CT: The public key PK, access structure level 0, and symmetric encryption key vk are all inputs to the encryption method, which returns the ciphertext CT as output. DO maintain the ciphertext CT. Step ⑤ in Fig. 4 shows how to use it.

---

Fig. 4. Communication diagram.

1530
\[ F(m) = CT \]

\[ P_n = k^m \mod p \]

In Eq. (1), \( F \) denotes the function of SHA-256, \( m \) denotes the file or document, \( CT \) denotes the created output.

In Eq. (2), \( P_n \) denotes the public key, \( K_g \) denotes the the key generator, \( x \) denotes the private key members, \( p \) denotes prime number.

**Key Gen (MSK, KS) → SK**: The essential creation procedure is still being carried out by the “DO”. “DU” presents an access demand to DA, who allocates DU a set of attributes and the effective access term. The method accepts the DU attributes set \( X \) as well as the master key “MSK” as contribution and returns the DU private key “SK”. “SK” is symmetrically encrypted using the common key as the encryption keys, once “DO” and “DU” shares the frequent key. The ciphertext SK0, which is the encrypted private key, is utilized to maintain anonymity. As seen in Fig. 4 at step ⑥⑦⑧.

**Decrypt (PK, SK, and CT) → vk**: DU performs the decryption process. DA grants access authority to the DU. The decryption procedure is then performed by DU if and only if proper access credentials are provided. “DU” acquires “CT” and the ciphertext of private key “SK0” from the IPFS server. The encryption method SHA256 decrypts the SK0 as the private key with the regular key as a “decryption key”. The process takes as input the public key “PK”, private key “SK” and ciphertext “CT”. DU can retrieve the encrypted document’s key \( ck \), allowing it to be decrypted, if and only if “SK” complies with the access strategy; if not, then decryption will be unsuccessful. Before DO encrypts the document \( F \), DU retrieves the encrypted document Evk (F) from the CS (cloud server), decrypts the encrypted article “F0” with key “vk”, and output the article \( F \). As illustrated in Fig. 4 at step ⑨⑩.

**Algorithm: Data Owner upload the file, key generation, Data User Registration**

**Input**: Encrypted File by owner to IPFS

**Output**: User get the decrypted file

1. **Encrypt** (PK, vk, 0) → CT
2. If (request if from DO):
   - accept the encrypted file by IPFS
   - verifiedDataOwner()
   else:
   - discard the request
3. **Setup** (1 k, AT) → (PK, MK)
4. (MSK, XS) → SK → generate_keys()
5. Addr → createIPFSaddress() + PK, SK
6. keepPrivatekey()
7. **register**!DataUser()
8. public_key, private_key = generate_keys()
9. authenticateUserRequest()
10. **Decrypt** (PK, SK, CT) → vk → getDecryptedFile()
11. extractOriginalFile()
5. Security and performance analysis

This section examines the security of the proposed blockchain-based decentralized data exchange system. It is identified that, the proposed system is secure under the not-completely-trusted cloud environment. In the following subsections few points are discussed regarding security and performance of proposed access control method.

5.1. Security analysis

Avoid single point of failure: By providing semi-decentralized block chain-based technology, this work modifies the classic “ciphertext-policy-attribute-based encryption” approach. The difficulty of a single central authority distributing keys in initial method can be overwhelmed by more user access. Instead of the central authority, the domain authority assigns the data user’s key, which is produced by the data owner.

Security achieved by Auditable Property of Blockchain: The system’s confidentiality is additionally ensured by the use of the best encryption and decryption engine. The cloud service provider is accountable for keeping ciphertext of file; nevertheless, while “CS” cloud server has unfettered access to ciphertext, it has no advantage over malicious users. Users’ access credentials are retrieved from domain servers, and they are only available for a specific DA. IPFS is a permission adopted network that supports transaction privacy and private data, and it is based on the blockchain concept. Because these DA credentials are linked to the user’s public key, even if an attacker hacks several DAs and obtains these credentials, it will be unable to decode without the private key.

Security against Collusion Attacks: This work also achieves security against collusion attacks. When malevolent users work together, they may exchange their secret attribute keys, which are produced after decryption, in order to obtain extra privilege. Because the decryption engine is linked to a precise user, if two users with diverse individuality try to collaborate and merge their secret attribute keys, the original factor will be lost. As a result, malevolent users will be unable to collaborate and gain more privilege.

Security achieving user authentication: While a user makes an access demand during blockchain operations, the proposed blockchain-based system provides auditable access control by recording a trustable and absolute access log on the chain. The user and the data owner may both see how the access policy was implemented to each access request. As a result, the system inherits the auditability feature of blockchain technology.

5.2. Performance analysis

The experimental examination of the suggested system is discussed in this section. The experimental platform and environment are configured as follows: 4 GB RAM, Intel(R) Core(TM) 2 Duo CPU E8400@3.00 GHz processor, and Windows10 operating system. Java and nodejs are the programming languages used. Python is used as an external assist for encryption and decryption. IPFS is a distributed decentralized block chain idea. IPFS is a peer-to-peer hypermedia protocol that makes the web more upgradeable, durable, and open so that humanity’s information can be preserved and grown.

5.2.1. Communication overhead

Let |G1| represent the size of an element in G1 and |G2| represent the size of an element in G2. The total number of people who have logged into the system is indicated by the symbol Nm. The number of DAs in a system is represented by NDA. The average number of users controlled by DAs is indicated by the symbol NU. The average number of credentials held by users is shown by the symbol Nw. The number of DAs required recovering domain data is represented by t threshold. Table 2 depicts the communication overhead of the centralized and proposed decentralized systems. Multiple domain authority settings impose some overhead for DA registration during the system initialization phase, which is the same as a centralized approach. The suggested system simply requires the administrator to deploy credential details and upload the public parameter to block chain, resulting in a 3|G1| communication overhead. The IPFS creates DKuid for the user during the key distribution phase. The size of the element DKuid is 3|G2|. Because the suggested system does not need to connect with each DA to gather each user’s credentials, the communication overhead on the user side is constant and clearly lower than centralized.

5.2.2. Computation overhead

The calculation overhead of decryption on the user side, is shown in Fig. 5. On the user side, the suggested solution provides superior decryption efficiency because blockchain handles the majority of the computing duties. The domain authority acquires the ciphertext submitted by the data owner from the block chain

![Table 2](#)

<table>
<thead>
<tr>
<th>Method</th>
<th>System Initialization</th>
<th>Key Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchal Centralized System (Li et al., 2016)</td>
<td>(Nw * NDA)</td>
<td>G 1</td>
</tr>
<tr>
<td>Proposed Semi-Decentralized System</td>
<td>3</td>
<td>G1</td>
</tr>
</tbody>
</table>

![Fig. 5](#)
during the decryption phase and performs decryption using the decryption engine associated with the user. Subsequently attaining the decryption keys, the user only requires to do an arithmetic operation and basic multiplication operations to complete the final decryption; thus, the completing time is unaffected by the number of users in the domain. As a result, the user’s decryption time is essentially constant.

6. Conclusion & future work

In this paper, a secure cloud access control is offered utilizing a semi-decentralized hierarchical architecture based on blockchains. Two primary difficulties have been recognized in centralized systems: the first is data manipulation by internal personnel, and the second is single point failure. A decentralized system based on blockchain is presented to solve the difficulties of a single point of failure, data modification by internal employees, excessive processing, and overhead during communication at the data consumer side. The introduction of block chain-based distributed IPFS technology transforms the centralized system. The distribution key is no longer reliant on the center authority to prevent an attack on it. IPFS is used to build a distributed access control system with indirect data owner and data user interaction. Experiments have shown that the cost of accessing files at the user level is quite low. The original data on a decentralized network cannot be modified by any other member or user because blockchain creates a new hash code for manipulated data every time it is changed. As a result, decentralized data access is more secure than centralized data access. Because blockchain functions like a distributed system, the decentralized system also avoids the problem of single point failure. Every node in a blockchain network stores a copy of the data of the other nodes. Because there is no single authority, data can be recovered quickly in the event of a failure. According to the security analysis, the proposed system could efficiently resist individual and collaborative harmful users, as well as not entirely trustworthy cloud servers. Furthermore, the blockchain-based method can record trustable and immutable access logs, allowing data owners to easily monitor users’ access activity in future. Other decentralized storage platforms, such as Storj, and other blockchain technologies, such as Ethereum and Hyperledger Fabric etc. may eventually replace cloud storage platforms for further work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The author would like to thank the anonymous reviewers for their constructive suggestions and comments, which have been incorporated into the article.

References


Steichen, M., Fitz, B., Norvill, R., Shbair, W., State, R., 2018. Blockchain-based, decentralized access control for IPFS. 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2018. IEEE, 1499-1506.


