Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

DAGShare: A Secure and Efficient Distributed Personal File Sharing System

Hira Arshad
Information Technology University
Sumbal Fatima
Information Technology University
Adnan Noor Mian
adnan.noor@itu.edu.pk

Information Technology University

Research Article

Keywords: Inter-Planetry-File-System (IPFS), IOTA, File Sharing, Blockchain, Access Control

Posted Date: October 21st, 2024

DOI: https://doi.org/10.21203/rs.3.rs-5283218/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

DAGShare: A Secure and Efficient Distributed Personal File Sharing System

Hira Arshad[a,1], Sumbal Fatima[b,2] and Adnan Noor Mian[a,*,3]

[a]Department of Computer Science, Information Technology University, , Lahore, 54000, Pakistan

ARTICLE INFO

Keywords:
Inter-Planetry-File-System (IPFS)
IOTA
File Sharing
Blockchain
Access Control

ABSTRACT

Many approaches have been proposed to share data in a decentralized manner using blockchain, which solves many security and data management issues but has some limitations, such as transaction fees, mining time, power consumption, etc. The alternative to a traditional blockchain is using IOTA, a directed acyclic graph (DAG) based system, but this also has the limitation of on-chain storage. To address these challenges we propose DAGShare, a DAG-based system with off-chain storage on an interplanetary file system (IPFS), a content-based addressable system. The DAGShare combines the IOTA network and IPFS to provide a secure and scalable personal file-sharing system in which members can create new groups and join other groups. The system inherits blockchain features and resolves its limitations. The system also provides an access control mechanism to grant system access rights by using IPFS-cluster-service that is not present in the IPFS server and IOTA. We provide a security evaluation of our system using the AVISPA tool and show that it is not only secure but also scalable as compared with IOTA using file sharing and retrieval experiments.

1. Introduction

The internet has made it possible to distribute and share huge amounts of data, including personal information. Data exchange can potentially add value to its consumers and their owners. However, unrestricted sharing of personal data can lead to security breaches. The personal data assets can be managed in either a centralized or decentralized fashion. In the former instance, service providers frequently regulate access to data, and data is at high risk. While sharing personal data from one entity to another, parties want their data to be secured from unwanted access and corruption of data. Sharing data in centralized systems is inefficient because it leads to the single point of failure and has limited trust and authority. There are other security concerns while sharing personal data i.e., data mutability, Sybil attack, and man in the middle attack, etc.

Compared to conventional centralized file-sharing systems, decentralized file-sharing offers improved security, privacy, and resilience by dispersing data over a network, removing single points of failure, and giving users more control over their files. To make data transparent, immutable, interoperable, and cryptographically secure blockchain-based decentralized systems have grabbed many researchers' attention after Satoshi Nakamoto's paper was published in 2008 [1]. It uses a mechanism for verification of transactions by going through a mathematical process done by miners for the generation of tokens along with decentralized nodes to maintain the decentralization of the network. Due to its immutable nature in which data cannot be altered, reduces the possibility of tampering in blockchain. It also offers several other benefits i.e. decentralization, trackability, and incorruptibility. The term *decentralization* refers to the fact that the system is not controlled by a single body. The terms *traceability* and *incorruptibility* refer to the fact that transaction outcomes are linked together via cryptographic mechanisms, making the chain's results distinctive.

File-sharing can utilize blockchain for immutable file history, transparency of data, secure access control mechanism, decentralized storage solutions, and enhanced security for a file-sharing system. In the literature, numerous systems have previously been suggested that leverage blockchain technology for file storage, file-sharing [2, 3], file security [4], access control, and file authenticity [5]. These systems utilize blockchain to uphold the security and reliability of their data.

Several applications have leveraged blockchain technology to enhance file storage, file-sharing [2, 3, 6], file security [7], access control, and file authenticity [5]. However, These applications encounter certain limitations such as scalability, transaction costs, and latency. For instance, educational systems struggle with scalability and transaction costs when the number of users and transactions increases [8]. Vehicular networks face latency concerns. Blockchain transactions cannot meet the real-time requirements of vehicle communication, and managing a large volume of real-time data can be challenging [9]. Edge computing encounters scalability issues when dealing with a massive number of transactions from edge devices. Additionally, the consensus processes introduce high latency, impacting the efficiency of the system [10]. Blockchain-based crowdsourcing platforms face transaction costs and scalability problems in managing large volumes of data [9]. Energy markets have concerns about energy consumption and scalability when handling real-time data [11]. System designs suffer from excessive energy consumption and scalability issues while processing a lot of operations [12]. In personal data-sharing systems individuals should also be able to cope with high quantities, velocity (speed of data generation and distribution), and

*Corresponding author
hiraarzo@gmail.com (H. Arshad); sumbalsohaib@gmail.com (S. Fatima); adnan.noor@itu.edu.pk (A.N. Mian)
ORCID(s):

veracity (data accuracy and truth-fullness) of data utilizing distributed cloud storage systems [12].

However, blockchain is limited in achieving excellent quality of service for distributed operations in regards to velocity and veracity [13]. Unfortunately, both the storage space and the speed with which information is processed are limited in blockchains [14]. Furthermore, proof of work (PoW) maintains the security of the system but miners are required to verify transactions, which takes a certain amount of time to verify each transaction, which makes blockchain inefficient in terms of scalability. To speed up the transactional throughput of blockchain, there have been proposed substitutes in recent research. A lightning network [15] is used to speed up the transactions. It is a layer-2 technology to increase the speed of transactions, which consists of many payment channels between users. However, these networks are highly vulnerable to cyber-attacks because they need to stay online all the time and failure of one node will crash the whole network. With the establishment of a multi-chain network, in which different blockchains can exchange information connected to the network, Polkadot [16] was aiming for a solution for the isolation of blockchain. For the calculation of data and information processing of transactions, Polkadot proposed using parallel chains to achieve this goal.

To improve limitations of security risks and computational overheads in the techniques described above, our study proposes a distributed file-sharing and storing system for personal data based on IOTA [17] and IPFS [18] to reduce computing overheads and security threats. IOTA is a decentralized DAG-based system and distributed ledger technology [19] that inherits all blockchain features but is more scalable than other blockchain environments [20]. It is fee-less, partition tolerant, and provides quantum-resistant cryptography [21]. IPFS is a decentralized file-sharing network built on a P2P network having a DAG-based structure using content-addressing to identify the files [22, 23]. It provides large storage to other blockchain-based systems and saves data off-chain. For distributed access control policies and group key management, we use an IPFS-cluster-service. Using the decentralized file distribution mechanism of IPFS nodes, the files are encrypted by the IPFS-cluster-service and kept on the IPFS node's server. Furthermore, we deploy IOTA as a trackable server to record file sharing/retrieving information in the network, preventing data tampering and solving the inherent issues faced by IPFS such as data integrity, node-level data duplication, and time taken in retrieving data. The IPFS node and IOTA network, combined with the use of the IPFS-cluster-service, form a secure file-sharing system in which system users can create new groups or join various groups at their leisure, and only group members have access to the approved files. The main contributions of the paper are

- We design and implement a DAG-based system with the integration of IPFS and IOTA and compare the proposed system with an IOTA-based system.
- We provide access control mechanism and key management by IPFS-cluster-service for personal file sharing system and then perform the security analysis of our proposed system on AVISPA to prove its safety.
- We perform experiments to show that the proposed system is efficient in file sharing and retrieval as compared with the IOTA-based system.

We conducted a security evaluation of the system using the AVISPA tool and determined that DAGShare is secure. Additionally, we conducted experiments on a network of 20 nodes, sharing files of various sizes. Our results indicate that DAGShare, with off-chain storage, outperforms IOTA with on-chain storage in terms of file sharing and retrieval. However, this improved performance comes at the expense of higher CPU utilization. To the best of our knowledge, there is no DAG-based implemented file-sharing system to improve scalability.

The paper is organized in the following manner. The background of IOTA and IPFS technologies are discussed in section II. Related work is described in section III. The proposed, DAGShare, personal data-sharing system is presented and describes how to combine IOTA and IPFS to get their benefits in section IV. The security analysis of our system on AVISPA, a widely used tool for security verification, is provided in section V. Results of the proposed are presented in section VI. Finally, the conclusion and future work are presented in section VII.

2. Preliminaries

In this section, we discuss the working of the two important entities, IOTA and IPFS used in the proposed system.

A. IOTA

IOTA was proposed by Dr. Serguei Popov in 2015 for immutable data transfer in a network, specially designed for IoT devices for machine-to-machine (M2M) communication [21]. IOTA uses DAG-based architecture, called Tangle, rather than a simple block-based architecture in blockchain [24]. The structure of the tangle is shown in Fig. 1. In tangle, each transaction uses a Random Markov Chain Monte Carlo (MCMC) algorithm to select two random unconfirmed transactions (called Tips) and attach proof of work to the transaction. The transaction is then broadcast to the peer-to-peer (P2P) network using the gossip protocol. Some basic properties of IOTA tangle are [17].

1. It is used without minors and thus no mining fee is charged.
2. In contrast to other blockchain technologies, the more transactions, the better it is for the tangle network.
3. It uses the Ed25519 signature scheme, which reduces the transaction size so that network bandwidth and processing time are also reduced.

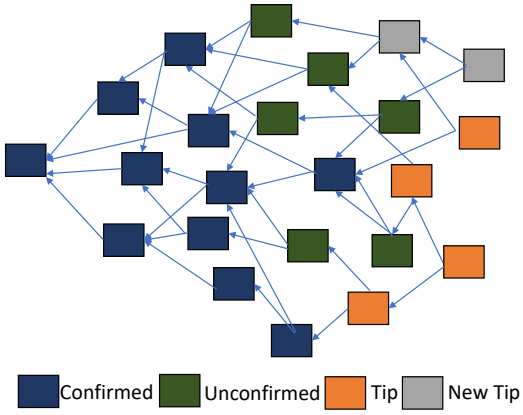

Figure 1: Structure of tangle

4. It allows offline transactions, which can become part of the main tangle when the connection is restored.

Consensus is obtained without minors, so no mining fee is charged [21]. Expanded transactions in the network do not hinder the IOTA and execution improves as the number of transactions increases because of the attribute of synchronous validation [25], [26]. It uses the Ed25519 signature scheme, which reduces the transaction size so that network bandwidth and processing time are also reduced [27]. IOTA can run on small and heterogeneous devices due to less computational power and energy consumption [25], [28], [29]. A white flag approach is used to improve the efficiency of tip selection and reduce the need for reattachment as well as eliminate several network attacks [30]. To resist various types of attacks and to reduce the negative impact of lazy nodes, three kinds of tip selection algorithms are used: uniform random, unweighted random walk, and weighted random walk algorithms. The most advanced tip selection method is the weighted random walk algorithm and Markov Chain Monte Carlo (MCMC) techniques are used in the Weighted Random Walk. Authors in [31] proposed a new technique for selecting 3 tips, two previous transactions and a third one from the transactions that are left behind, to increase the confidence of all transactions in the tangle. If there is no left-behind transaction it will simply select only two previous transactions. The process of updating data involves the creation of a new transaction and its validation through the Tangle. When a user in the network wants to update data, they create a new transaction. To submit the new transaction, the user needs to choose two existing transactions (tips) to validate. This process helps in achieving consensus and confirming the integrity of the new transaction.

The validation mechanism while retrieving a transaction in the IOTA network involves ensuring the consistency and integrity of the transaction data. This mechanism typically operates by confirming the coherence and integrity of the transaction data as follows.

- A user needs $TRAN_{ID}$ to retrieve specific data that uniquely points to that specific transaction in the tangle.
- When a query is made, a node in the network responds by providing the information of the requested transaction.
- Retrieved transaction is validated by the user by performing a consistency check. This verified that the transaction adheres to the rules of the tangle, the referenced tips are valid and PoW is correct. They may also check the cumulative weight of the transaction.

As transactions are public, an authentication method known as Masked Authenticated Messages (MAM) is employed to prevent spam data and monitor user fake data which is a cryptographic protocol for data transmission. MAM creates a linked data flow from messages by forwarding each message to a new address using Tangle transactions. Each transaction n contains a pointer to the transaction n+1, and each transaction is linked to its subsequent transaction but they do not know where the transaction n-1 took place [32]. MAM implements the Merkle signature scheme (MSS), which is based on Merkle Hash Tree (MHT) combined with One-Time Signatures (OTS) and resistant to quantum computer attacks [28], [33]. The Tangle differs in that the messages are linked to one another. These messages are necessary for the ledger, and when they are uploaded to the distributed network, they increase the total hashing power while also benefiting from the network's information integrity features, while different transactions refer to them implicitly. At the point when Tangle was launched, it utilized an organizer, called coordinator, to forestall vindictive transactions since it needed more nodes to lead approval of transactions. This coordinator has been out of date after the IOTA network turns out to be more settled. The coordicide had replaced the coordinator [34].

B. IPFS

According to the content of the file in IPFS, every stored file is identified by a unique hash. When a user requests a file, the selected hash is used as the matching file address [22]. As long as there are duplicated copies of the files in the system, the files are permanent. IPFS combines several key characteristics. First, in IPFS, Distributed Hash Tables (DHT) is a data structure that efficiently coordinates data to search between nodes quickly. Due to the advantages of decentralization, data can be transferred without the need for central coordination with DHT. Second, IPFS incorporates BitSwap, a peer-to-peer file-sharing technology [35]. BitSwap functions as a marketplace where nodes can purchase blocks. The blocks that make up a file can come from any file on the file-system, even if they are unrelated.

Another essential feature of IPFS is the Merkle DAG. It connects data blocks that represent the target's cryptographic hashes. The Merkle DAG ensures data identification as well as tamper resistance. Version Control Systems (VCS), like Merkle DAG, is another component of IPFS that helps users

maintain their data more easily. It enables users to edit several versions of a file and efficiently share them. The self-certifying file system (SFS) [36] is the IPFS's final concept. It establishes the Interplanetary Name Space (IPNS), an SFS that enables users to validate things published on the network. The publishers' private keys can be used to sign the objects, which can then be validated using the publisher's public key.

3. Related Work

A solution, RapidChain [37], was proposed for the solution of problems such as storage capacity, scalability, and transaction per second (TPS) that were low in Bitcoin. A reference committee C_R is formed by nodes by using a consensus mechanism in RapidChain and to iterate the members of the reference committee authors introduced the concept of an epoch. The construction of the committee is updated consistently to circumvent Sybil attack. Kad (Kademlia) is the routing protocol used between the committee and the client. A distributed and peer-to-peer (P2P) platform for trading is provided to the users by a software layer, Omnilayer [38], that was built on top of Bitcoin. It supports smart properties and currencies generated by the users on the other hand it is vulnerable to attacks. In smart manufacturing systems to reduce data processing time, reduce the cost of production, and increase productivity rate, a system model is proposed by Shahbazi and Choel with the integration of blockchain, machine learning, and edge computing [39]. Although previous techniques improve blockchain they still require transaction fees and depend on miners which are not suitable for voluntary file-sharing systems such as personal data-sharing networks.

There have been many approaches proposed to take advantage of blockchain with IPFS [13]. Authors in [40] presented a file-sharing system that is both scalable and trustworthy. Their approach employs a different re-encryption technique [41] and a scalable key management method to support many users' access. The encryption key is distributed, making it more resistant to collusion attempts by revoked users. Because the encryption key's one sub key is supplied to the authorized user, the key must be handled by the user, increasing security issues. While authors in [42] proposed an agent-oriented framework but their system inherits the limitations and challenges of blockchain based on distributed technology. Authors in [43] proposed a solution that can audit and trace access events in IPFS. The authors evaluated the scalability by varying the number of nodes and measured the latency of different events. But the scale of their experiments is very small i.e. number of nodes ranging from 3 to 27 which makes the scalability of their system unknown on a large scale. An incentive-based system was proposed in [44] which uses filecoin as an incentive mechanism with IPFS, it can be used when you want to persuade, users to share files. A secure file-sharing system was proposed in [45] which uses blockchain and IPFS along with IPFS proxy. For exchanging IoT data, the authors in [46] designed a blockchain-based storage solution. In the proposed scheme when an authorized user is revoked a re-encryption token is used to manage the keys, which ensures that the key encrypting the data is updated preventing future access to the system. The owner of data, on the other hand, is responsible for encryption of key and updating operations, which considerably increases the data owner's computing overheads.

Fortunately, IOTA improves the verification mechanism of transactions in the blockchain [17]. In the absence of miners, the tangle system does not take much time for the validation of a transaction hence considered to be a scalable system, and having no fee for transaction makes it cost-effective. The inherent features of blockchain make it secure for data-sharing and being a decentralized system makes it tamper-proof from a single point of failure. IOTA is considered to have the potential to solve the problems of high transaction fees and low throughput of blockchain. Along with IOTA use of IPFS solves the problem of storage for larger files.

To the best of our knowledge, there is no study available that combines IOTA and IPFS. The integration of blockchain and IPFS is being utilized in various systems. Initially, it is important to understand the necessity of security in these systems. To address this need, blockchain systems are employed due to their inherent security characteristics. However, blockchain systems have certain limitations such as transaction fees and mining time leading to scalability issues. To overcome these limitations, the DAGShare protocol is introduced. IOTA offers a different approach to achieving security and scalability compared to traditional blockchain systems. IOTA also has its limitations in terms of on-chain storage capacity. To overcome this limitation, an off-chain solution using IPFS is implemented. IPFS allows for decentralized and distributed storage of data, ensuring that the system can handle files of larger size. Thus, by combining IOTA, and IPFS, a comprehensive solution is created that addresses the security requirements and storage limitations of the system.

4. Proposed System Model

We propose a secure personal file sharing system *DAGShare* based on, DAG-based architectures, IOTA, and IPFS. The file-sharing methods in the different systems may have permission issues because there is no trusted central entity to validate users. The IPFS-cluster-service is used in the proposed system to provide a distributed access control mechanism and group key management. IOTA is a distributed ledger that allows users to record transactions in a decentralized manner. However, because nodes in the IOTA network have limited memory, only transaction records, which include all can be saved on the IOTA network rather than an entire file. As a result, IPFS is used to satisfy the needs of storing larger files of data. On the other hand, in IPFS and blockchain-based integrated systems, computational overheads and security threats are not taken into

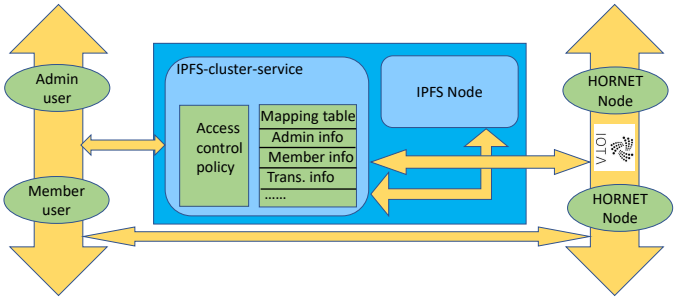

Figure 2: Proposed System Model

account. In this paper, we propose a secure personal file sharing and retrieval mechanism that uses the IPFS-cluster-service to manage access control settings and IOTA for fee-less and fast data retrieval.

Fig. 2 shows the proposed system model and entities in DAGShare. It consists of Admin, Member, IPFS-cluster-service, IPFS node, and IOTA tangle. The system admin is the one who creates a new group. It is the group's first member and can share files within the group. It is also in charge of access policies. When a user becomes a member of the group, he/she gets the right to retrieve the files shared by other members of the group. When a new group is created, the IPFS-cluster-service generates a group key pair. The group key then encrypts the file, ensuring that the file uploaded by the group members remains private. The IPFS-cluster-service manages the mapping table, which contains key lists of the group's members. When a user needs a file he first registers to the group and goes through the registration process for which he sends his member-ID along with its public key and group ID and in return he will get a member ID. After becoming a member, he can access the files by requesting them. The encrypted files are kept on the IPFS node after being encrypted by the IPFS-cluster-service and then the IPFS node sends the encrypted file to the members on request. The file-sharing information can be recorded on the tangle due to the incorruptible property of IOTA. The IOTA is regarded as a trackable server in addition to the identity and hash value, ensuring that the ownership and integrity of files are immutable.

Accessing data from IPFS relies on network bandwidth and in our scenario, when bandwidth is limited speed and efficiency of file retrieval are impacted. The other challenge faced is that IPFS depends on nodes to host and distribute data so for a non-popular file data availability sometimes becomes an issue. The better solution for this is to keep the nodes always on or have many nodes so if one node is down other is available to provide data.

A. The DAGShare Protocol

The DAGShare consists of registration, file sharing, and file retrieval phases. The details are presented below.

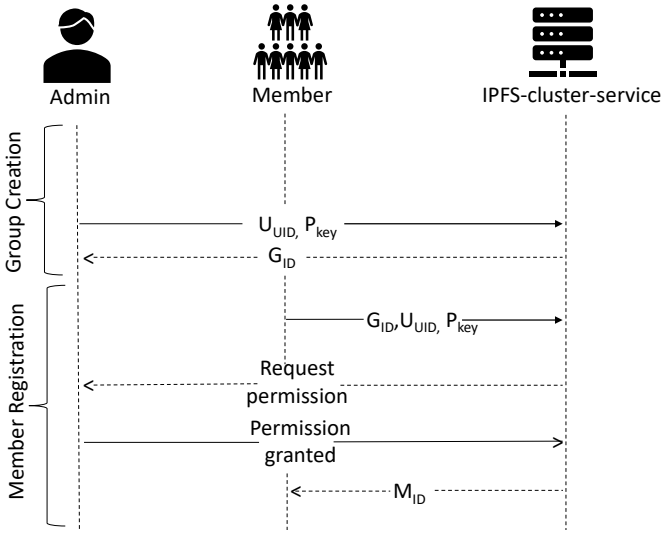

Figure 3: Sequence diagram of the registration process of admin and member

Algorithm 1: Registration

Data: P_{key}, RequestType, U_{UID}
Result: M_{ID},G_{ID}

```
if (RequestType == "Generate G_ID ") then
    call IPFS-cluster-service (P_key, U_UID);
    return G_ID;
    mapping-table.store(G_ID);
end
if (Request-Type == "Request to join group") then
    if (IsValid(U_UID, G_ID, P_key) then
        return M_ID;
        mapping-table.store(M_ID);
    else
        Message.show (Request Denied);
    end
end
```

A.1. Registration

The sequence diagram of registration is shown in Fig. 3 and corresponding Algorithm 1.

- To create a new group in the system, the admin must first register with the IPFS-cluster-service. The admin uploads the public key he created, as well as its unique user ID U_{UID}. Following registration, the IPFS-cluster-service provides the admin with a group ID number, G_{ID}. The IPFS-cluster-service has a mapping table, which contains key lists generated by those who have registered with the system and G_{ID} (line 1-5, Algorithm 1).
- When a user needs to retrieve files belonging to a specific group, he or she must first register to that group by producing a key pair and then upload the key pair which includes the public key, along with

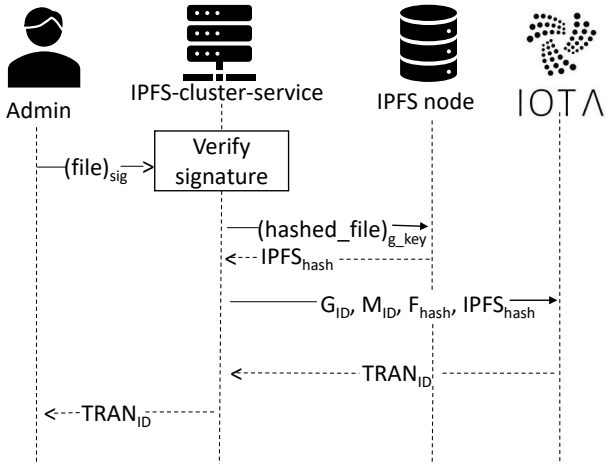

Figure 4: Sequence diagram of file sharing

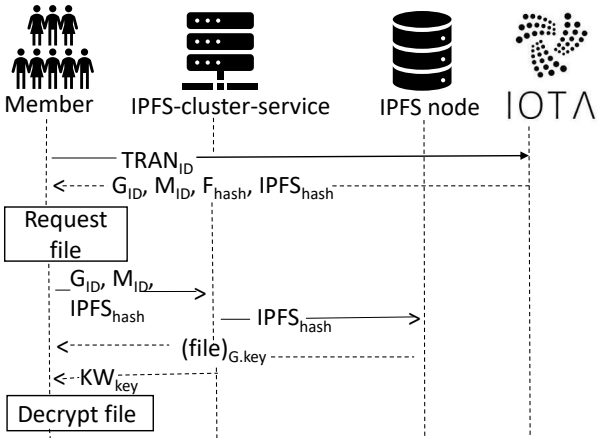

Figure 5: Sequence diagram of file retrieval

Algorithm 2: File Sharing

Data: File
Result: $TRAN_{ID}$

1 Message.add(Signed-file);
2 $IPFS-cluster-service \leftarrow User forward Message$;
3 **if** *(IsValid(Signature(U_{UID}, G_{ID})))* **then**
4 return f_{hash};
5 mapping-table.store(f_{hash});
6 **else**
7 Message.show (Signature not valid);
8 **end**
9 Message.add(Encrypt file with group-key);
10 $IPFS-node \leftarrow IPFS-cluster-service forward Message$;
11 return $IPFS_{hash}$;
12 mapping-table.store($IPFS_{hash}$);
13 Message.add($IPFS_{hash}$, f_{hash}, G_{ID}, U_{UID});
14 $IOTA \leftarrow IPFS$-cluster-$service forward Message$;
15 return $TRAN_{ID}$;
16 mapping-table.store($TRAN_{ID}$);

the U_{UID} and the chosen G_{ID}, to the IPFS-cluster-service. The group admin must grant authorization to a new member. The IPFS-cluster-service uploads the member's public key and its U_{UID} to the mapping table and completes the member's registration procedure after returning a member id M_{ID}, either the member is authorized to view files in the group or not (line 6-13, Algorithm 1).

A.2. File Sharing

The file-sharing mechanism is described by the sequence diagram Fig. 4 and the corresponding Algorithm 2.

- The registered admin/member signs his ready-to-share file. The IPFS-cluster-service checks if the signature is correct (line 1-3, Algorithm 2).
- The IPFS-cluster-service hashes the file and creates a F_{hash} to validate the file's authenticity after properly validating the signature. This hash is primarily used for integrity verification and to ensure that the file has not been tampered with during transmission. This is a crucial step for maintaining the security and trustworthiness of the stored files. It also encrypts the file using group-key. The encrypted file is shared to the IPFS node, to be saved on the server, with group-key saved in the mapping table of IPFS-cluster-service. After saving the encrypted file, the IPFS node returns $IPFS_{hash}$. This $IPFS_{hash}$ is associated with the saved encrypted file on the IPFS node. This hash is utilized for referencing and retrieving the file from the IPFS network (line 3-12, Algorithm 2).
- The IPFS-cluster-service records the combined value of G_{ID}, M_{ID}, $IPFS_{hash}$, and F_{hash} on IOTA tangle network as a transaction after receiving the $IPFS_{hash}$ (line 13-14, Algorithm 2).
- The tangle then publishes transaction and returns IPFS-cluster-service a transaction id $TRAN_{ID}$. The sharing process is completed after the IPFS-cluster-service returns $TRAN_{ID}$ to the group. Every member of the group maintains a ledger that keeps a record of group membership. Whenever ledge receives a $TRAN_{ID}$, it is multicasted to all group members, and thus all group members know the transaction (line 15-16, Algorithm 2).

A.3. File Retrieval

The file retrieval mechanism is described by the sequence diagram Fig. 5 and the corresponding Algorithm 3.

- The authorized member searches the tangle for transaction data using $TRAN_{ID}$ (line 1-5, Algorithm 3).
- The member sends the request to the IPFS-cluster-service with G_{ID}, U_{UID}, $File_{hash}$, and $IPFS_{hash}$ after successfully receiving the required transaction data.(line 6, Algorithm 3)

Algorithm 3: File Retrieval

Data: $TRAN_{ID}$
Result: File
1 **if** *(RequestType == "File Retrieval")* **then**
2 call IOTA ($TRAN_{ID}$);
3 **if** *(IsValid($TRAN_{ID}$))* **then**
4 return Transaction Data;
5 **end**
6 call IPFS-cluster-service(U_{UID}, G_{ID}, $IPFS_{hash}$);
7 **if** *(IsValid(U_{UID}, G_{ID}, $IPFS_{hash}$))* **then**
8 Message.add($IPFS_{hash}$);
9 $IPFS - node \leftarrow IPFS - cluster - service forward Message$;
10 **end**
11 Message.add(Encrypted File);
12 $Member \leftarrow IPFS Node forward Message$;
13 Message.add (KW_{key}) ;
14 $Member \leftarrow IPFS - cluster - service forward Message$;
15 Message.show (File Retrieved);
16 **end**

- IPFS-cluster-service validates the request before querying the IPFS node using the $IPFS_{hash}$. Finally, the IPFS node sends the selected encrypted file to the respective members. (line 7-11, Algorithm 3)
- Because the file is encrypted by the group key generated by the IPFS-cluster-service that is why it is no longer accessible by members. We use a key wrap technique that uses another key to encrypt one key. The IPFS-cluster-service sends the member a keywrap-key KW_key to decrypt the file they received. The member's public key encrypts the group key with KW_key. The member then obtains the group key after decrypting by KW_key. The encrypted file can then be decrypted with the group key (line 12-15, Algorithm 3).

Finally, to validate the file's authenticity, the member hashes the file and compares the result to the file hash acquired from the tangle.

B. Access Control

The designed IPFS-cluster-service is essential to the system. It provides a platform for anyone who wants to connect to the DAGShare system to register and be validated. The member's public key is saved in a mapping table in the IPFS-cluster-service after registration. When a new admin wants registration in the system, the IPFS-cluster-service produces a new group key first. As a symmetric cryptosystem is faster, the group key is symmetric. The mapping table stores the group key with the admin's public key. After the validation is finished, another member's public key is added to the same mapping table to gain access to the same group. The various groups have their own individual mapping tables, which serve as the IPFS-cluster-service's access control policies. The security of the system is managed by the IPFS-cluster-service. To prove the identification of admin and members of the system, their requests are signed before sending to the IPFS-cluster-service which provides proxy services to the system and assures the requests are valid. The same group key is used to encrypt all shared files inside that group. The IPFS-cluster-service issues a new group key that replaces the old one if anyone in a group has been revoked. All the IPFS server's encrypted files that are part of the group need to use the new group key to re-encrypt and $IPFS_{hash}$ are stored on IOTA as a new transaction. The new group key will be shared with all users, except the revoked user, and it will be encrypted using the public keys of the authenticated users. The new files are encrypted by the new group key to prevent the user whose access has been revoked from accessing any future data. The revocation process needs communication overhead on the order of O(n) for groups of n users. It also gets rid of invalid requests to maintain system security.

5. Simulation and verification using AVISPA tool

In this section, we simulate and analyze the security of our proposed protocol with the help of AVISPA (Automated Validation of Internet Security Protocols and Applications) [47, 48], a widely used powerful tool for verification of communication and cryptographic protocols. AVISPA automates the process of analyzing protocol and makes it easier to identify vulnerabilities. The simulation code is written in high-level protocol specification language (HLPSL). It tests protocols using four different back-ends: OFMC, CL-AtSe, SATMC, and TA4SP, which implement different analysis techniques.

In the proposed protocol we have four entities with basic roles defined in the AVISPA implementation for file sharing protocol: Admin, IPFS cluster service, IPFS node, and IOTA. Similarly, file retrieval protocol has basic roles of Member, IPFS cluster service, IPFS node, and IOTA. Both protocols have two composition roles: session and environment, which represent entities and environment parameters with specific roles. We define the fundamental functions of all entities in a module. The details provided in a specific role can only be played by that specific entity. The intruder may perform some roles as a genuine user in AVISPA, which contains one session role and global constants. A statement in the code specifies the intruder's knowledge in which public keys, entities, shared keys, and publicly known activities are all included.

After stating all of the main objectives, we ran the code with the CL-AtSe (Constraint-Logic-based Attack Searcher) back-end and received the result that the proposed protocol is safe. The simulation results of our system are shown in Fig. 6 and Fig. 7, which shows our system DAGShare is secure.

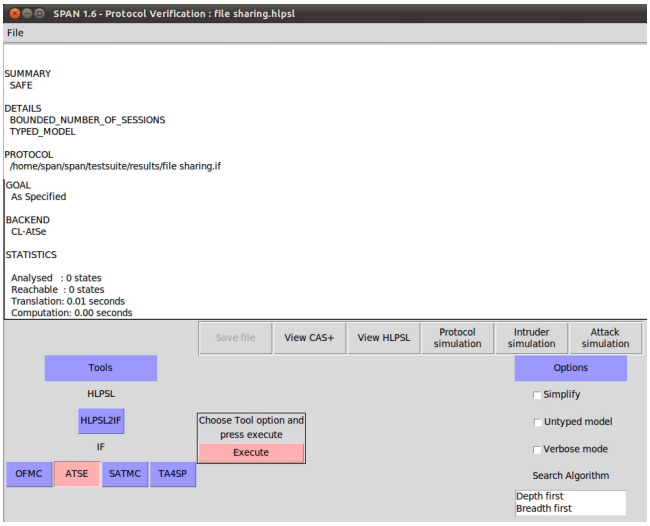

Figure 6: Simulation result while file sharing using CL-AtSe back-end

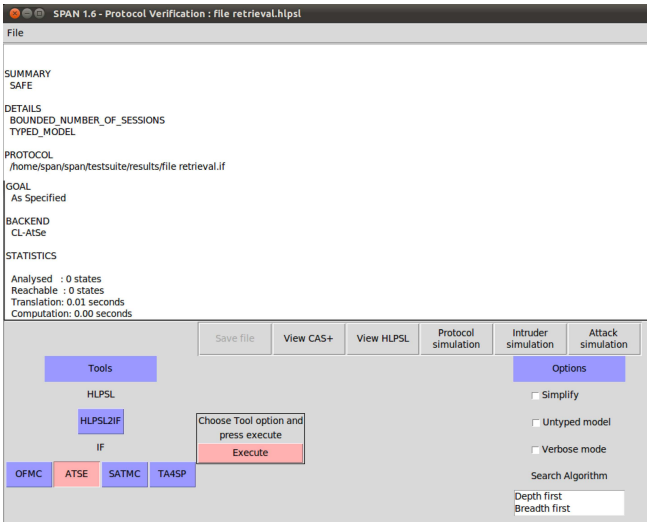

Figure 7: Simulation result while file retrieval using CL-AtSe back-end

6. Experiments and Results

We performed experiments on systems with 8GB RAM with quad-core processors on Ubuntu 20.04 Linux machines. We implemented IOTA 1.5 with HORNET 1.0.6 network node and used its JavaScript and client library. IPFS 0.11.0 was used with CLI. The rest of the work is done using nodejs, Go, and Python programming languages. The results are averaged on 5 runs. We measured the number of transactions processed and their confirmation time with the varying size of the file. Data files used for experiments were of different types i.e. pdf, txt, png that varied in size.

Following are the results of DAGShare measured on a network of 20 nodes in which some are considered as admin and others are members to form a personal file-sharing network. In IOTA, a network of different coordinator groups is used as zero-value transactions to confirm the transactions.

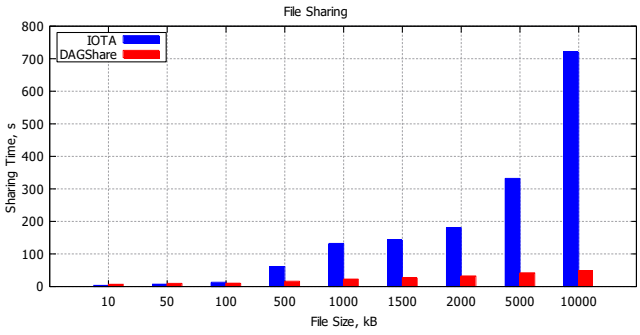

Figure 8: File sharing only through IOTA in comparison with DAGShare

We ran the coordinator every 30 seconds. It can be removed when the network is large enough. We measured the average confirmation rate in DAGShare. It is 0.7 transactions per second for varying numbers of requests. This confirmation rate is low because all transactions cannot be referenced by the coordinator. This transaction confirmation rate increases with the increase in the network size and number of transactions in the network. Furthermore, the increase in the number of transactions leads to congestion. The “snapshot” is a unique property of IOTA which over time reduces the size of transactions to rescue from congestion.

Fig. 8 shows the results of file sharing. We see that as time required to share a file increases in both DAGShare and IOTA as the size of the file increases. However, the time required in the case of DAGShare is much less than that of IOTA. In IOTA the transaction size is limited to 32kB. This makes a large number of transactions to share a large file size and hence much higher time, whereas in the case of DAGShare, the files are stored and shared off-chain on IPFS which does not have such limitations and hence results in much less time. File Retrieval time is shown in Fig. 9. We see that it is less when the file size is small both in IOTA and DAGShare, but increases as the file size increases. However, this time is initially less in the case of IOTA than DAGShare and then from 500kB, it becomes higher in IOTA. Initially, IOTA takes less time than DAGShare because DAGShare involves an access control verification mechanism but as the file size increases, the retrieval time in IOTA increases because a large number of file chunks, previously stored, are now to be retrieved, which takes longer time. Thus, we see that our proposed system improves efficiency of the file sharing and retrieving as compared with IOTA on-chain storage. The performance improvement in file share and retrieval, however, comes at the cost of a slight increase in CPU utilization. Fig. 10 shows the CPU utilization of the system concerning the system running time. We divide this time into four regions. Bootstrapping is the time when neither the DAGShare nor IOTA nodes have yet started, system running idle shows the time when both the DAGShare and IOTA are up and running but not yet started to share or retrieve data, file sharing time means when a file is shared and file retrieval means when a file is being retrieved. We see

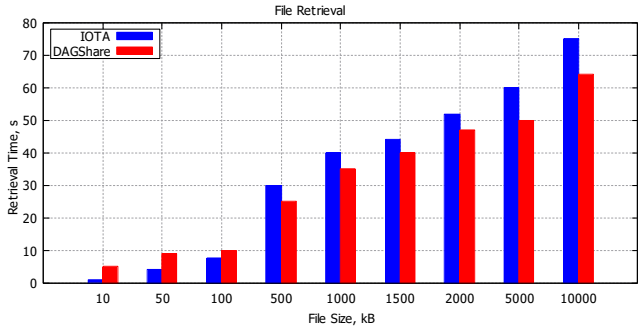

Figure 9: File retrieval on IOTA in comparison with DAGShare

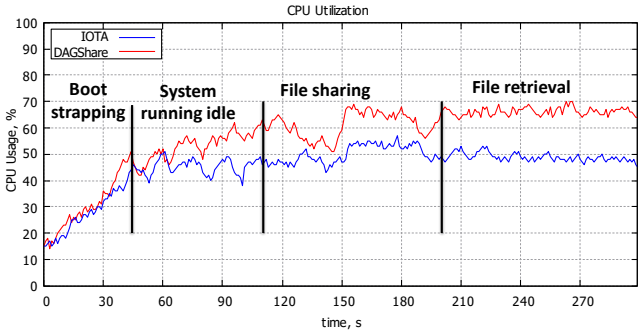

Figure 10: CPU usage in different phases of system

that CPU utilization increases when file-sharing or retrieval takes place. In our system, it remains between 50% and 70%. This can be different for a system with a different processing power. In our experiments, we noticed that the CPU usage is lower when using IOTA compared to DAGShare. This is because in DAGShare, two nodes (IOTA and IPFS), are running on the system simultaneously, whereas in IOTA, there is only one node running on the system. To address this issue, we could utilize a more powerful system, but we were restricted by the processing power limitations. That shows the CPU utilization in DAGShare is always higher than using IOTA only.

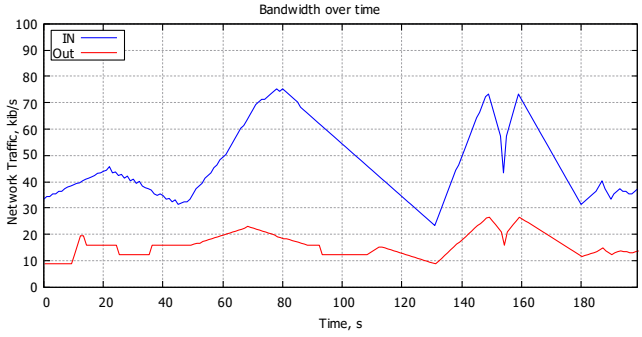

Figure 11: Bandwidth consumption of the system

Fig. 11 shows the bandwidth occupancy/usage of the system over time. Incoming (In) is retrieving data in the system while outgoing (Out) is sending/sharing data from the system. The bandwidth consumption of our system is low it only increases when we interact with the server. our system consumes less bandwidth which makes it suitable for low-power devices and a network with a higher number of users.

Table 1
Comparison of DAGShare, IOTA and Blockchain-based file sharing systems

Features	DAGShare	IOTA	Blockchain
Fees	No	No	Yes
Mining	No	No	Yes
Scalability	High	Medium	Low
Transaction Speed	Fast	Fast	Slow
Quantum Proof	Yes	Yes	No
Smart Contract	No	Yes	Yes
Storage	IPFS	On-chain	On-chain
Structure	DAG	DAG	Blockchain

TABLE 1 shows the comparison summary of the proposed system with IOTA and other conventional blockchain-based file-sharing systems.

7. Conclusion and Future Work

We presented DAGShare, a personal file-sharing system based on IOTA and IPFS. We adopted an IPFS-cluster-service for access control and key management and an IPFS server to store off-chain data rather than on-chain which takes more time for larger files to share and retrieve. The combination of IOTA and IPFS along with IPFS-cluster-service provides a secure and scalable file sharing system. In DAGShare, users can create a new group and join other groups of their choice and also share and retrieve personal files from the authorized group. We provided a security evaluation of the proposed system using the AVISPA tool which shows that DAGShare is safe. We have performed experiments by sharing files of different sizes in a network of 20 nodes. Our experimental results show that our proposed system DAGShare with off-chain storage is performing better than IOTA with on-chain in terms of file sharing and retrieval but at the cost of higher CPU utilization. In the future, we are considering implementing DAGShare on a larger scale and measuring its efficiency for an environment with a large number of users. We also plan to use an attribute-based access control mechanism with IOTA to give the system more access control rights.

8. Declarations

Availability of Data and Material

There is no data or any other material associated with this manuscript.

Competing Interests

The authors proclaim that they have no competing interests.

Code Availability

Not Applicable

Funding
Not Applicable

Acknowledgement
Not Applicable

Author's Contributions
Hira Arshad analyzed the requirement of distributed personal file-sharing System experimented and drafted the manuscript. Also, Sumbal Fatima gave full support in experimenting with using AVISPA to check the security of the algorithm. Adnan Noor Mian contributed by reviewing the work done and revising the content of the manuscript. In addition, all the work is done under Adnan Noor Mian supervision.

References

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized Business Review*, p. 21260, 2008.

[2] I. Zhou, I. Makhdoom, M. Abolhasan, J. Lipman, and N. Shariati, "A blockchain-based file-sharing system for academic paper review," in *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2019, pp. 1–10.

[3] S. Pradhan, S. Tripathy, and S. Nandi, "Blockchain based security framework for p2p filesharing system," in *2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, 2018, pp. 1–6.

[4] S. Park, M. Specter, N. Narula, and R. L. Rivest, "Going from bad to worse: from Internet voting to blockchain voting," *Journal of Cybersecurity*, vol. 7, no. 1, p. tyaa025, 02 2021.

[5] Y. Zhang, C. Xu, N. Cheng, H. Li, H. Yang, and X. Shen, "Chronos+: An accurate blockchain-based time-stamping scheme for cloud storage," *IEEE Transactions on Services Computing*, vol. 13, no. 2, pp. 216–229, 2019.

[6] M. Almasian and A. Shafieinejad, "Secure cloud file sharing scheme using blockchain and attribute-based encryption," *Computer Standards & Interfaces*, vol. 87, p. 103745, 2024.

[7] J. Li, J. Wu, and L. Chen, "Block-secure: Blockchain based scheme for secure p2p cloud storage," *Information Sciences*, vol. 465, pp. 219–231, 2018.

[8] M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić, "EduCTX: A blockchain-based higher education credit platform," *IEEE access*, vol. 6, pp. 5112–5127, 2018.

[9] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1495–1505, 2018.

[10] A. Stanciu, "Blockchain based distributed control system for edge computing," in *2017 21st International Conference on Control Systems and Computer Science (CSCS)*. IEEE, 2017, pp. 667–671.

[11] X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. Bass, C. Pautasso, and P. Rimba, "A taxonomy of blockchain-based systems for architecture design," in *2017 IEEE international conference on software architecture (ICSA)*. IEEE, 2017, pp. 243–252.

[12] V. Stankovski and R. Prodan, "Guest editors' introduction: Special issue on storage for the big data era," *Journal of Grid Computing*, vol. 16, no. 2, pp. 161–163, 2018.

[13] M. I. Khalid, I. Ehsan, A. K. Al-Ani, J. Iqbal, S. Hussain, S. S. Ullah *et al.*, "A comprehensive survey on blockchain-based decentralized storage networks," *IEEE Access*, 2023.

[14] A. I. Sanka and R. C. Cheung, "A systematic review of blockchain scalability: Issues, solutions, analysis and future research," *Journal of Network and Computer Applications*, vol. 195, p. 103232, 2021.

[15] J. Poon and T. Dryja, "The bitcoin lightning network: Scalable off-chain instant payments," 2016.

[16] G. Wood, "Polkadot: Vision for a heterogeneous multi-chain framework," *White Paper*, vol. 21, pp. 2327–4662, 2016.

[17] Dr. Serguei, "IOTA," https://www.iota.org/, Accessed April 7, 2022.

[18] Juan Benet, "IPFS," https://ipfs.io/, Accessed April 7, 2022.

[19] L. Zhao, L. Vigneri, A. Cullen, W. Sanders, P. Ferraro, and R. Shorten, "Secure access control for dag-based distributed ledgers," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[20] P. Ferraro, C. King, and R. Shorten, "Iota-based directed acyclic graphs without orphans," *arXiv preprint arXiv:1901.07302*, 2018.

[21] S. Popov and Q. Lu, "IOTA: feeless and free," *IEEE Blockchain Technical Briefs*, 2019.

[22] J. Benet, "Ipfs-content addressed, versioned, p2p file system," *arXiv preprint arXiv:1407.3561*, 2014.

[23] J. Kaur, R. Rani, and N. Kalra, "Attribute-based access control scheme for secure storage and sharing of ehrs using blockchain and ipfs," *Cluster Computing*, vol. 27, no. 1, pp. 1047–1061, 2024.

[24] S. Popov, "The tangle," *White paper*, vol. 1, no. 3, 2018.

[25] Y. Jiang, C. Wang, Y. Wang, and L. Gao, "A cross-chain solution to integrating multiple blockchains for iot data management," *Sensors*, vol. 19, no. 9, p. 2042, 2019.

[26] D. Assante and R. Crucco, "M2M learning environment for electric applications," in *2019 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 2019, pp. 1518–1522.

[27] U. Sarfraz, M. Alam, S. Zeadally, and A. Khan, "Privacy aware IOTA ledger: Decentralized mixing and unlinkable iota transactions," *Computer Networks*, vol. 148, pp. 361–372, 2019.

[28] J. Brogan, I. Baskaran, and N. Ramachandran, "Authenticating health activity data using distributed ledger technologies," *Computational and structural biotechnology journal*, vol. 16, pp. 257–266, 2018.

[29] X. Zheng, A. Vieira, S. L. Marcos, Y. Aladro, and J. Ordieres-Meré, "Activity-aware essential tremor evaluation using deep learning method based on acceleration data," *Parkinsonism & related disorders*, vol. 58, pp. 17–22, 2019.

[30] M. Conti, G. Kumar, P. Nerurkar, R. Saha, and L. Vigneri, "A survey on security challenges and solutions in the iota," *Journal of Network and Computer Applications*, p. 103383, 2022.

[31] G. Bu, Ö. Gürcan, and M. Potop-Butucaru, "G-IOTA: Fair and confidence aware tangle," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 644–649.

[32] P. Handy, "Introducing masked authenticated messaging," https://medium.com/iotatangle/introducing-masked-authenticated-messaging-e55c1822d50e, Accessed April 7, 2022.

[33] S. Shafeeq, M. Alam, and A. Khan, "Privacy aware decentralized access control system," *Future Generation Computer Systems*, vol. 101, pp. 420–433, 2019.

[34] S. Popov, H. Moog, D. Camargo, A. Capossele, V. Dimitrov, A. Gal, A. Greve, B. Kusmierz, S. Mueller, A. Penzkofer *et al.*, "The coordicide," *Accessed Jan*, pp. 1–30, 2020.

[35] A. De la Rocha, D. Dias, and Y. Psaras, "Accelerating content routing with bitswap: A multi-path file transfer protocol in ipfs and filecoin," 2021.

[36] D. D. F. Mazières, "Self-certifying file system," Ph.D. dissertation, Massachusetts Institute of Technology, 2000.

[37] M. Zamani, M. Movahedi, and M. Raykova, "Rapidchain: Scaling blockchain via full sharding," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018, pp. 931–948.

[38] O. Layer, "Omni layer," http://www.omnilayer.org/, Accessed April 7, 2022.

[39] Z. Shahbazi and Y.-C. Byun, "Improving transactional data system based on an edge computing–blockchain–machine learning integrated framework," *Processes*, vol. 9, no. 1, p. 92, 2021.

[40] S. Cui, M. R. Asghar, and G. Russello, "Towards blockchain-based scalable and trustworthy file sharing," in *2018 27th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2018, pp. 1–2.

[41] C. Dong, G. Russello, and N. Dulay, "Shared and searchable encrypted data for untrusted servers," *Journal of Computer Security*, vol. 19, no. 3, pp. 367–397, 2011.

[42] A. Tenorio-Fornés, S. Hassan, and J. Pavón, "Open peer-to-peer systems over blockchain and ipfs: An agent oriented framework," in *Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems*, 2018, pp. 19–24.

[43] E. Nyaletey, R. M. Parizi, Q. Zhang, and K.-K. R. Choo, "BlockIPFS-blockchain-enabled interplanetary file system for forensic and trusted data traceability," in *2019 IEEE International Conference on Blockchain (Blockchain)*. IEEE, 2019, pp. 18–25.

[44] S. Vimal and S. Srivatsa, "A new cluster P2P file sharing system based on IPFS and blockchain technology," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–7, 2019.

[45] H.-S. Huang, T.-S. Chang, and J.-Y. Wu, "A secure file sharing system based on IPFS and blockchain," in *Proceedings of the 2020 2nd International Electronics Communication Conference*, 2020, pp. 96–100.

[46] H. Shafagh, L. Burkhalter, A. Hithnawi, and S. Duquennoy, "Towards blockchain-based auditable storage and sharing of iot data," in *Proceedings of the 2017 on cloud computing security workshop*, 2017, pp. 45–50.

[47] A. Armando, D. Basin, Y. Boichut, Y. Chevalier, L. Compagna, J. Cuéllar, P. H. Drielsma, P.-C. Héam, O. Kouchnarenko, J. Mantovani *et al.*, "The avispa tool for the automated validation of internet security protocols and applications," in *International conference on computer aided verification*. Springer, 2005, pp. 281–285.

[48] M. A. Akram, H. Ahmad, A. N. Mian, A. D. Jurcut, and S. Kumari, "Blockchain-based privacy-preserving authentication protocol for uav networks," *Computer Networks*, vol. 224, p. 109638, 2023.