

LCHAIN: A Secure Log Storage Mechanism using IPFS and Blockchain Technology

Parin Patel¹, Hiren Patel²

¹Research Scholar

Kadi Sarva Vishwavidyalaya

Gandhinagar-382015, Gujarat, India

patelparinv@gmail.com

²Sarva Vidyalaya Kelavani Mandal

Gandhinagar-382015, Gujarat, India

hbpatel1976@gmail.com

Abstract— Data security is a very important and crucial part of Cloud storage. Day by day, thousands of operations function over the Cloud, verify logs generated from all transactions is very difficult. The attacker may temper or remove targeted logs or attack traces on the system with stealthy techniques. It is required to maintain the security of the log to trace back all the transactions to identify such tempering and loss of logs. Temper-proof log storage is a challenging issue on the Cloud. To overcome the issue, we propose strong and secure log storage using Blockchain technology and IPFS. Detecting log tempering and tracing is challenging due to large log volumes. We recommend the usage of Blockchain technology due to its inherent feature of immutability to address the issue of log tempering. We present LChain which provides immutable storage of logs with tracing. Blockchain technology helps to create immutable logs but also offers non-repudiation and scalability. Smart contracts are used for efficient searching and enhancing computational power.

Keywords- Log Storage, Cloud, Blockchain, Smart Contract, Integrity, Access Control

I. INTRODUCTION

Logs are used as pieces of evidence when issues arise in the system. For instance, forensic investigation logs help identify the pattern or strategy of the attacker. To keep an organization safe and secure, it is vital to monitor and verify logs periodically. An effective log monitoring and analysis helps recognize the loopholes in the system which can be utilized by the malicious user(s).

Due to the real-time scenario, most monitoring mechanisms have a minimum window time frame with a small time duration to deal with unexpected situations, which helps the attackers deceptively attack for a long period. Therefore, it is significant to store and maintain the huge volume of logs for a long period and use them for threat detection. It requires a huge investment to collect, store, index, or sort the logs. Many systems are available which can be used to identify various attacks and types of threats [1][2].

Through this article, we have studied various situations for the Cloud domain which show the importance of secure log storage management. For instance, in infrastructure-as-a-service (IaaS), the data resides at a service provider's site which makes it vulnerable, especially in the situation when the service provider behaves maliciously and provides incomplete or tempered log reports. This results in a dependency on the provider and leaves the data owner with a choice of nothing but

believing the service provider as it has sole and full authority over the data. The issue becomes more complex in the scenario of a community cloud where infrastructure is shared between several organizations. This leads us to identify some mutually trusted solutions that can assure the client's data security.

To conclude the above discussion, it is required to have a trusted, traceable, immutable and verifiable log system to maintain trust among Cloud service providers and their users. There are multiple solutions available for log tampering detection over the Cloud [3], [4] but merely tampering detection may not be enough hence we focus on immutable log storage using Blockchain technology.

In this paper, we have studied various solutions provided to address security issues of Cloud log storage using Blockchain technology. The objective of our research is to propose an architecture LChain that provides immutable log storage using Blockchain technology. We use an Interplanetary system to overcome the big file-handling issue of Blockchain. We also focus on efficient searching which results in less gas consumption by using mapping instead of the array in the smart contract. The usage of smart contracts improves computational efficiency. The rest of the paper is arranged as follows. In section 2, we discuss contemporary work which addresses security issues in log storage over the Cloud. Section 3 discusses some preliminary studies related to our proposed work. We propose our system in Section 4. Section 5 demonstrates the

implementation result with a detailed analysis. We conclude our study in section 6 followed by a list of references at the end of the article.

II. RELATED WORK

An efficient way to manage log storage has attracted many researchers' attention. They have focused on many prospective solutions that could detect and prevent issues in log storage. Authors of [5] have proposed a technique Logcrypt to store a log cryptographically in such a way that it cannot be altered. Authenticity is achieved through the usage of public key cryptography. Multiple logs were maintained concurrently and verified logs with a single initial value. Authors of [6] have proposed algorithms to transfer logs over the Cloud storage. Log monitors are stored in centralized servers that generate queries if they find any unusual activities by focusing on integrity and security concerns. Authors of [7] have proposed a healthcare management system that uses Blockchain technology to store patient data. Every transaction is stored on Blockchain to provide security to patients' data. Data of patients is encrypted using cryptographic techniques, which also guarantee pseudonymity. They examine the data processing practices as well as the affordability of the smart contracts employed in their system. Authors of [8] have proposed solutions that provide immutable log storage using Blockchain technology. In this system, all transactions are stored on a block of Blockchain that is immutable and tamper-proof. In this system, every log is stored with the signing key using a logging Cloud server for long-term analysis and maintenance. All logs are stored on a local log server which is monitored using a log monitor client. Authors of [9] have proposed a Blockchain-based query-response system that works efficiently. They have used a permissioned Blockchain network with a multichain module called streams which uses the key value to store data. Only indexes of data are stored on the Blockchain network. Authors of [10] proposed a system LCaaS that uses a hierarchical ledger and a repository that can be accessed by all Cloud participants. It stores all logs on the hierarchical ledger, hence it provides immutability to the storage and also avoids tempering or stealing logs. Authors of [11] have used smart contracts to create an index of every log file and store all transactions over the Blockchain. The proposed system uses Anchor and Query modules where Anchor upload logs on the storage and Query handles the requests by obtaining index value through a smart contract. The query identifies the hash index of the file and replays the log file content to the client. In this proposed scheme they have used private Ethereum and IPFS clusters. Authors of [12] have designed a log storage system using HyperLedger permissioned Blockchain network. This system uses two operations: data appending and querying. Data appending uses a trigger function and appends data to the Blockchain. For the

query process, it uses the function that retrieves the log data from the storage using key elements. This proposed system only works for elastic search for the data storage system. Authors of [26] proposed a solution where the encrypted file is stored on the cloud and IPFS generates the hash value of the file stored on the blockchain network. Here they remove single-point failures like a centralized system. Domain authority will manage all the access controls from the clients or data users. Administrator manage the system by setting up the system and parameter initialization.

In this research, we aim to propose a scheme wherein we store the logs on a Cloud server using Interplanetary File System (IPFS). The hash value of the log is stored on the Blockchain network. Further, we use the smart contract to perform cryptographic operations like retrieval of requested data and appending of data as per the request of the client, which results in trust among the non-trusted entities and offers a reduction in communication and computational overhead. We proposed a novel solution using Blockchain, smart contracts, and IPFS to secure log storage without using a third-party monitoring system.

The major contribution of this paper is as follows.

1. We achieve integrity of data through the inherent feature of the immutability of Blockchain as data once stored can't be altered. Further, the exact location can't be predicted due to the fact that the Blockchain is distributed in nature resulting in hard to predict the exact file location.
2. We offer high availability and the system resists a single point of failure.
3. Performance in the form of search time is also improved by using mapping technology in a smart contract.
4. We compare our work with the existing system on criteria such as computation, size, and validation time.

III. PRELIMINARIES

In this section, we analyze our scheme Lchain proposed through figure 1 and figure 2 with the usage of Blockchain, Smart Contract, and IPFS.

A. Blockchain

Blockchain technology consists of a block structure that provides an immutable distributed public ledger. Blockchain prevents any unethical activities like tempering or removing data from storage. Blockchain is a chain of blocks and each block contains the header and set of a transaction where the block header stores hash value, Merkle Root, Timestamp, version, nonce value, etc. A new hash of the block is calculated using the previous block hash value, hence it is very difficult to temper the block as each block is linked with the previous and next blocks. Blockchain provides trust, traceability, security, and

transparency. Due to these advanced features, Blockchain technology can be used in log storage management as the log once stored on Blockchain cannot be altered.

Ethereum (a decentralized, open-source Blockchain with smart contract functionality) enables the development and execution of a smart contract and distributed application without any central authority. It is an open-source platform and uses its currency (Ether). It uses the Turing programming language to build and run applications.

B. Smart Contract

We use smart contracts because it manages all data access mechanisms, and use the ownership certificate. It also manages all transaction details on the Blockchain network. All the transactions are saved on the Blockchain network that can't be modified. Smart contracts are easy to build and deploy and use less computational power. Smart contract performs all operations between Cloud and Data Owner / User without the interference of third-party. A smart contract helps the user to store metadata on the Blockchain when any type of update occurs. A smart contract is a decentralized platform that validates and enforces agreements through transactions.

C. Smart Interplanetary File System (IPFS)

An IPFS is a system that is used to store and access files, data, websites, or applications. It uses content addressing where all files or data are identified by the use of the content, not by the location. It generates a cryptographic hash value that is unique for every content. It searches the data using that hash value not by the location so it gives faster results than other searching devices. As the decentralized version of HTTP, IPFS (Interplanetary File System) is a peer-to-peer network of nodes that collaborate for the purpose of sharing and storing information. IPFS retrieves the data based on what data is, not where data is stored like cloud storage servers. You won't have to search through cloud folders as cloud servers try to remember where you put them.

Owner: The Owner makes requests to access the logs if any suspicious activities are found. The owner uploads the data on a server. The owner will get an integrity report so he can identify the loss of data or any fraud.

Smart Contract: Smart contracts are the functions that automatically manage all functions like storing hash values to the Blockchain, retrieving the data from the Cloud log storage server, and identifying the integrity of data. All the operations are handled by smart contract without interference from any trusted third party. We use mapping to store log file metadata using the hash value returned by IPFS so it will reduce timing in searching for the requested log file data.

Interplanetary File System (IPFS): Interplanetary File System stores the log files and generates the hash of that file. It stores the metadata, and public parameters and maintains the integrity, and immutability of data. Domain authority can manage user credentials and help businesses to create a trusted environment. It is a distributed hash table that maps unique hash to data which is generated by the log generator devices and uploaded on servers. It is a decentralized storage system where the file is distributed in such a way that it's not easy to identify the location. Storing all data on a Blockchain is very expensive so we use a solution like IPFS. It is distributed file system that is available free to anyone. You can store any type of file on servers. It provides security, reliability, and privacy to our system.

Blockchain: Generated logs have a unique hash value which was generated by Interplanetary File System(IPFS) stores on blocks. It provides fast searching of data in the Blockchain by using mapping of a hash value to the file metadata. Blocks store the Blockchain ID, IPFS Hash value, timestamp, and file ID as transactions over the Blockchain. We save only metadata on Blockchain hence it became cheaper for our system.

B. Architecture Overview

An overview of the proposed architecture is illustrated in Fig 1.

IV. SYSTEM MODEL

This section provides a detailed overview of the Lchain system model. We describe all participants, the flow of our proposed system model, and all phases with detailed algorithms.

A. Entities of system model

The components of the system model are explained as follows.

Log Generators: These are the devices that generate logs that are stored on the Interplanetary File System(IPFS). It collects the logs from various log generator devices which are connected through a network. It uploads the logs on the server and sets the threshold values for several logs that can be stored in one batch.

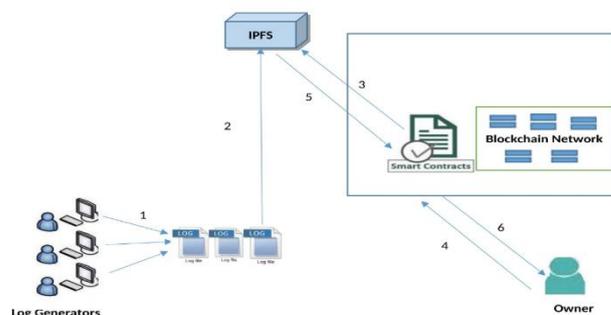


Figure 1. Lchain system model using Blockchain and IPFS.

An overview of the LChain System Model using Blockchain is illustrated in Fig. 1. All model steps are outlined below.

1. Log generator devices will generate log files.
2. Encrypted log files uploaded to Interplanetary File System(IPFS).
3. Interplanetary File system generates a unique hash code for the encrypted log file. Generated unique hash code will be saved on the Blockchain using the function UploadHash.
4. When the Owner finds any suspicious activity he sends a query request using the RetriveLog function of a smart contract.
5. Requested log file will be searched using the file index which is saved on the Blockchain. The Search function will search file hash from the Blockchain using file index and replay to the IPFS. The GetFile function searches log files on the Cloud servers using hash value and replays to the Owner. GetHash function gets IpfsHash using the GetHash function. The IpfsHash value generated from IPFS and FHash value received from the Blockchain will be compared.
6. An integrity report will be provided to the client. Hence Client can easily identify tempering or loss of the logs. ‘

C. Smart contract Design

Below section, we describe a smart contract algorithm used in our proposed work. A smart contract is written in solidity language and deployed on the Ethereum Blockchain platform using the Truffle compiler.

1. UploadHash: The SendHash function gets IpfsHash from IPFS and saves that hash value on the Blockchain network. Log file data is stored using mapping instead of the array so searching for data becomes faster in the Blockchain network.

Algorithm 1 Upload hash value on the Blockchain

```

Procedure UploadHash
//SendHash procedure will return the hash
value of the file stored on IPFS.
IpfsHash = SendHash(memory);

//Upload data on Blockchain for fileid
mapping(IpfsHash=>FileId) lfile.txt
Add_mappingfile (Sha256(Blockchain_Id,
File_Id, IpfsHash, timestamp) {
append (Sha256(Blockchain_Id, File_Id,
IpfsHash, timestamp)-> lfile.txt
    
```

```

Upload (SHA256(Blockchain_Id, File_Id,
IpfsHash, timestamp)
    
```

```

//GetHash function will return the stored
hash value on the Blockchain for the
requested file.
IpfsHash=GetHash(Fileindex);
end if
end procedure
    
```

2. RetriveLog: This function is used when a client is requesting for the log file when any suspicious activity is identified. It will compare new and old hash values and notify the client about the integrity of the log.

Algorithm 2 Log retrieved from Cloud storage and get the new hash value of the log file from IPFS. Compare the hash value with the hash stored on the Blockchain which will be retrieved from the Blockchain.

```

Procedure RetriveLog
//request received from the Owner
Request(Fileindex);

//Find the stored hash value of the requested
file from the Blockchain
Fhash=search(Fileindex);

// Send Fhash to IPFS and receive the requested
log file. The GetHash function will return a new
hash value for the requested file.
IpfsHash=GetHash(GetFile(Fhash));

//compare the new hash with the hash value
stored on Blockchain
If(IpfsHash, Fhash) =True then
Message (CID," file is correct");
else
Message (CID," file is Incorrect");
end if
end procedure
    
```

V. IMPLEMENTATION RESULTS AND ANALYSIS

All experiments were carried out under the system windows 11, 16 GB RAM and 2.7 GHz intel i7 11 generation processor. Implementation has been done using Node.js and the npm version used 16.15.0. An interplanetary file System(IPFS) is used to upload the file to the Cloud storage server. Ganache v2.5.4 is the Ethereum Blockchain that is used to run and test the contract files. Truffle v5.6.4 is used to build, debug and deploy smart contracts over the Ethereum Blockchain network. Visual studio code is used as an editor to write a smart contract and other files.

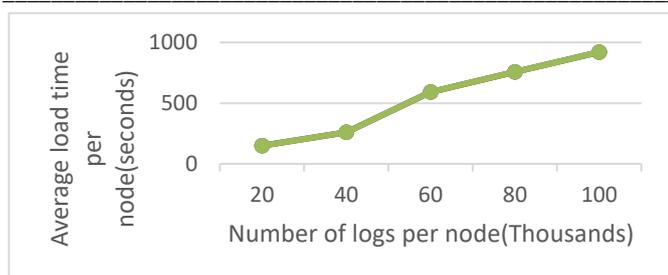


Figure 2. Average time to load number of logs on the log server

The Lchain system model is executed and obtained results using above mention configuration and technology. Fig 2. shows the average load time per node with respect to the number of logs loaded at a time. The graph shows a linear variation number of log entries. If the buffer size is large then retrieval of logs is also faster. We set the buffer size to 104 as per the transaction limit reached. Fig 3. shows the validation overhead when the client requests the log files. In such cases, we have compared 10 times to identify the result of validation time. The proposed system model also shows linear variation in validation time for 10 log files. Compare to the existing system our proposed system shows minimum delay in the verification of the logs.

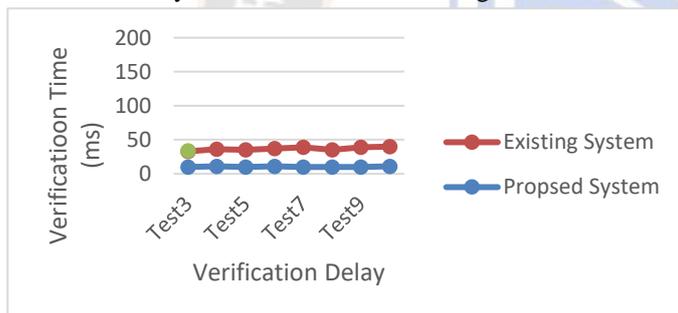


Figure 3. Verification Delay comparison with the existing system

This paper proposes a system model which provides immutable log storage with integrity verification using Blockchain technology. We have also discussed existing technology which has identified some problems like integrity, security, and performance of the system. Hence we have built a solution using Blockchain technology and an IPFS-based Cloud log storage system that overcomes existing problems in current Cloud storage platforms. IPFS stores files on the Cloud in a distributed manner so it is very hard to locate files per user. The hash value of the corresponding file is stored on the Blockchain which is immutable, transparent, and fully secure. Trust between the Cloud log storage server and the client is established with the system.

Through this research, our major contributions are as under.

1. Privacy: Real data stored on the storage server can't be directly identified due to IPFS, where IPFS stores and search file using the hash value. Hash value stores on the

Blockchain without file location so again there is no direct access available publically. Hence Privacy for a client on Blockchain and contents are not accessible using a user id.

2. Distributed Ledger: It is shared between the client and the server. Access management of data is possible with the system so anyone can't access the data stored on the server.
3. Immutability: File hash is generated using IPFS and files are distributed over the servers so it shows immutability in storage. The hash value of the file stored on the Blockchain. In the Blockchain, it is impossible to alter transactions in a block due to hashing calculation of Blockchain technology. Blockchain also guarantees immutable property to our system.
4. Scalability: Due to the parallel processing feature of Blockchain multiple blocks are created at the same time, added to super blocks at the same time, and upload in the Blockchain network. Because of these characteristics, it provides scalability to our system. Blockchain technology consists of so many servers that any block can have added to any chain at the same time so there is no loss of data.
5. Efficient Searching: We used mapping instead of an array in a smart contract to store values on the Blockchain. Mapping is used to fetch data on a key-based or value-based. So it gives faster results. Hence our proposed system has less time in searching the values compared to the existing system.

VI. CONCLUSION

In this paper, we have presented a design and implementation of a model Lchain that makes use of Blockchain technology and IPFS to offer security to data. Blockchain technology offers immutable and tamper-proof data storage. We use smart contracts to overcome the problem of trusted third parties or any centralized authority which may create a single point of failure. As the servers only store the logs of data or transactions, users' anonymity is protected. Mapping permits smart contracts to fetch data faster from the Blockchain. Our system has less validation and storage overhead. The proposed system creates trust between Cloud users and Cloud storage servers using IPFS and Blockchain technology.

REFERENCES

- [1] J. Buric and D. Delija, "Challenges in network forensics," 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, 2015, pp.1382-1386.
- [2] N. Virvilis, B. Vanautgaerden and O. S. Serrano, "Changing the game: The art of deceiving sophisticated attackers," 2014 6th International Conference On Cyber Conflict (CyCon 2014), Tallinn, 2014, pp. 87-97.

- [3] M. Sato and T. Yamauchi, "Vmm-based log-tampering and loss detection scheme," *Journal of Internet Technology*, vol. 13, no. 4, pp. 655–666, 2012.
- [4] R. T. Snodgrass, S. S. Yao, and C. Collberg, "Tamper detection in audit logs," in *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30. VLDB Endowment*, 2004, pp. 504–515.
- [5] Holt JE. Logcrypt: forward security and public verification for secure audit logs. *Proceedings of the 4th Australasian workshops on grid computing and e-research (ACSW '06)*, Tasmania, Australia, 2006; 203–211.
- [6] I. Ray, K. Belyaev, M. Strizhov, D. Mulamba and M. Rajaram, "SecureLogging as a Service—Delegating Log Management to the Cloud," in *IEEE Systems Journal*, vol. 7, no. 2, pp. 323-334, June 2013.
- [7] Omar, Abdullah & Bhuiyan, Md & Basu, Anirban & Kiyomoto, Shinsaku & Rahman, Shahriar (2019). A privacy-friendly platform for healthcare data in Cloud based on Blockchain environment. *Future Generation Computer Systems*. 95C. 511-521. [10.1016/j.future.2018.12.044](https://doi.org/10.1016/j.future.2018.12.044).
- [8] Dr. Manish Kumar, Ashish Kumar Singh, Dr. T V Suresh Kumar, (2018) Secure Log Storage Using Blockchain and Cloud Infrastructure, 9th ICCCNT 2018, IISC, Bengaluru, India, IEEE'.
- [9] Ozdayi, M.S., Kantarcioglu, M. & Malin, B. Leveraging Blockchain for immutable logging and querying across multiple sites. *BMC Med Genomics* 13 (Suppl 7), 82 (2020). <https://doi.org/10.1186/s12920-020-0721-2>
- [10] W. Pourmajidi and A. Miransky, "Logchain: Blockchain-Assisted Log Storage," 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), 2018, pp. 978-982, doi: [10.1109/CLOUD.2018.00150](https://doi.org/10.1109/CLOUD.2018.00150).
- [11] W. Huang, "A Blockchain-Based Framework for Secure Log Storage," 2019 IEEE 2nd International Conference on Computer and Communication Engineering Technology (CCET), 2019, pp. 96-100, doi: [10.1109/CCET48361.2019.8989093](https://doi.org/10.1109/CCET48361.2019.8989093).
- [12] H. Wang, D. Yang, N. Duan, Y. Guo and L. Zhang, "Medusa: Blockchain Powered Log Storage System," 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 2018, pp. 518-521, doi: [10.1109/ICSESS.2018.8663935](https://doi.org/10.1109/ICSESS.2018.8663935).
- [13] S. Ali, G. Wang, B. White and R. L. Cottrell, "A Blockchain-Based Decentralized Data Storage and Access Framework for PingER," 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 2018, pp. 1303-1308, doi: [10.1109/TrustCom/BigDataSE.2018.00179](https://doi.org/10.1109/TrustCom/BigDataSE.2018.00179).
- [14] M. Castro, B. Liskov, Practical Byzantine fault tolerance and proactive recovery, *ACM Transactions on Computer Systems (TOCS)* 20 (4) (2002) 398–461.
- [15] Z. Zheng, S. Xie, H.-N. Dai, H. Wang, Blockchain challenges and opportunities: A survey, *Work Pap.*–2016 .
- [16] Jiaying Li, Jigang Wu, Long Chen, Block-Secure: Blockchain Based Scheme for Secure P2P Cloud Storage, *Information Sciences* (2018), doi: [10.1016/j.ins.2018.06.071](https://doi.org/10.1016/j.ins.2018.06.071)
- [17] Ilya Sukhodolskiy, Sergey Zapechnikov, A Blockchain-Based Access Control System for Cloud Storage, *IEEE* 2018, 978-1-5386-4340-2/18
- [18] Xue, Jingting & Xu, Chunxiang & Zhang, Yuan & Bai, Lanhua. (2018). DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain: 18th International Conference, ICA3PP 2018, Guangzhou, China, November 15-17, 2018, Proceedings, Part III. [10.1007/978-3-030-05057-3_30](https://doi.org/10.1007/978-3-030-05057-3_30).
- [19] Xueping Liang, Sachin Shetty, Deepak Tosh, Charles Kamhoua, Kevin Kwiat, and Laurent Njilla, ProvChain: A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability, 2018 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing
- [20] S. Wang, X. Wang, and Y. Zhang, "A Secure Cloud Storage Framework With Access Control Based on Blockchain," in *IEEE Access*, vol. 7, pp. 112713-112725, 2019, doi: [10.1109/ACCESS.2019.2929205](https://doi.org/10.1109/ACCESS.2019.2929205).
- [21] Pengcheng Wei, Dahu Wang, Yu Zhao, Sumarga Kumar Sah Tyagi, Neeraj Kumar, Blockchain data-based Cloud data integrity protection mechanism, *Future Generation Computer Systems*, Volume 102, 2020, Pages 902-911, ISSN 0167-739X, <https://doi.org/10.1016/j.future.2019.09.028>.
- [22] W. Pourmajidi and A. Miransky, "Logchain: Blockchain-Assisted Log Storage," 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), 2018.
- [23] D. Yue, R. Li, Y. Zhang, W. Tian and C. Peng, "Blockchain Based Data Integrity Verification in P2P Cloud Storage," 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), Singapore, Singapore, 2018, pp. 561-568. doi: [10.1109/PADSW.2018.8644863](https://doi.org/10.1109/PADSW.2018.8644863)
- [24] Sutton A., Samavi R. (2017) Blockchain Enabled Privacy Audit Logs. In: d'Amato C. et al. (eds) *The Semantic Web – ISWC 2017*. ISWC 2017. Lecture Notes in Computer Science, vol 10587. Springer, Cham.
- [25] Xuanmei Qin, Yongfeng Huang, Zhen Yang, Xing Li, A Blockchain-based access control scheme with multiple attribute authorities for secure Cloud data sharing, *Journal of Systems Architecture*, Volume 112, 2021, 101854, ISSN 1383-7621, <https://doi.org/10.1016/j.sysarc.2020.101854>.
- [26] Smita Athanere, Ramesh Thakur, Blockchain based hierarchical semi-decentralized approach using IPFS for secure and efficient data sharing, *Journal of King Saud University - Computer and Information Sciences*, Volume 34, Issue 4, 2022, Pages 1523-1534, <https://doi.org/10.1016/j.jksuci.2022.01.019>.