

## **BSSPD: A Blockchain-Based Security Sharing Scheme for PersonalData with Fine-Grained Access Control.**

**Mr.MV. Sathya Narayana<sup>1</sup>, C.Chandana<sup>2</sup>,Ch.Chaitanya Jyothi<sup>3</sup>, D.Varshitha<sup>4</sup>, B.Himani<sup>5</sup>**  
<sup>2,3,4,5</sup> UG Scholars, Department of CSE, *MALLA REDDY ENGINEERING COLLEGE FOR WOMEN*, Hyderabad, Telangana, India.

<sup>1</sup>Assistant Professor, Department of CSE, *MALLA REDDY ENGINEERING COLLEGE FOR WOMEN*, Hyderabad, Telangana, India.

### **ABSTRACT:**

Privacy protection and open sharing are the core of data governance in the AI-driven era. A common data-sharing management platform is indispensable in the existing data-sharing solutions, and users upload their data to the cloud server for storage and dissemination. However, from the moment users upload the data to the server, they will lose absolute ownership of their data, and security and privacy will become a critical issue. Although data encryption and access control are considered up-and-coming technologies in protecting personal data security on the cloud server, they alleviate this problem to a certain extent. However, it still depends too much on a third-party organization's credibility, the Cloud Service Provider (CSP). In this paper, we combined blockchain, ciphertext-policy attribute-based encryption (CP-ABE), and InterPlanetary File System (IPFS) to address this problem to propose a blockchain-based security sharing scheme for personal data named BSSPD. In this user-centric scheme, the data owner encrypts the sharing data and stores it on IPFS, which maximizes the scheme's decentralization. The address and the decryption key of the shared data will be encrypted with CP-ABE according to the specific access policy, and the data owner uses blockchain to publish his data-related information and distribute keys for data users. Only the data user whose attributes meet the access policy can download and decrypt the data. The data owner has fine-grained access control over his data, and BSSPD supports an attribute-level revocation of a specific data user without affecting others. To further protect the data user's privacy, the ciphertext keyword search is used when retrieving data. We analyzed the security of the BSSPD and simulated our scheme on the EOS blockchain, which proved that our scheme is feasible. Meanwhile, we provided a thorough analysis of the storage and computing overhead, which proved that BSSPD has a good performance

### **INTRODUCTION:**

The development of 5G and Internet of Things technology provides a large amount of training data for the rapid implementation of artificial intelligence (AI). At the same time, data security and privacy protection have become the most interesting topics in data governance and sharing. Powerful data mining and analysis have brought potential threats to personal privacy protection. Traditionally, most people choose to

outsource their data to cloud servers for sharing and dissemination. However, most of the data stored in the cloud is very sensitive, especially those data generated by IoT devices that are closely related to human life. These data have their particularities and may contain personal-related information such as life, work, and healthcare; once personal data is stolen or leaked illegally and linked to the data owner's real identity, it may bring great trouble to an individual. Therefore,

integrating data and generating value while ensuring data security and privacy have become a significant challenge for all contemporary companies that use big data and AI.

At present, researchers have proposed many secure sharing schemes in the cloud environment [1–9]. These schemes seem to solve the security and privacy issues during data sharing. Nevertheless, these schemes all have a standard feature: they are overly dependent on the Cloud Service Provider (CSP). They believe that the CSP is a trusted third-party organization, and their security models assume that the CSP is semitrustable, which means that the CSP will be curious about the data but will not destroy it. It means that the following situations are always inevitable.

(1)The CSP itself may make profits from the user's private data, or its insiders may do evil and cause the user's privacy disclosure. Although some methods, such as attribute-based encryption algorithms, can achieve user-defined access policies that seem user-centric, these methods still require a trusted third party to generate and manage user keys. It is impossible to exclude the possibility of collusion between these trusted centers. All these will lead to the fact that once the data owners upload their data to the cloud server, they will no longer have their data's absolute possession

(2)The data is centrally stored on cloud servers and managed by the CSP. An inevitable single point of failure may lead that users cannot obtain their data generally by using the cloud service. The CSP can improve data security and service stability by utilizing disaster recovery backup. However, some

irresistible factors will prevent users from using cloud services to obtain their data, such as political factors

(3)To provide better service, the CSP needs to spend more money to buy servers, hire better employees, rent the data center venues, and so on. These costs are increasing gradually, and the CSP cost is also increasing and the construction of the management platform. Users ultimately pay the operating costs of the CSP

#### LITERATURE SURVEY:

**g our logging mechanism together with users' data and policies. We leverage the JAR** Data sharing becomes an exceptionally attractive service supplied by cloud computing platforms because of its convenience and economy. As a potential technique for realizing finegrained data sharing, attribute-based encryption (ABE) has drawn wide attentions. However, most of the existing ABE solutions suffer from the disadvantages of high computation overhead and weak data security, which has severely impeded resource-constrained mobile devices to customize the service. The problem of simultaneously achieving fine-grainedness, highefficiency on the data owner's side, and standard data confidentiality of cloud data sharing actually still remains unresolved. This paper addresses this challenging issue by proposing a new attribute-based data sharing scheme suitable for resource-limited mobile users in cloud computing. The proposed scheme eliminates a majority of the computation task by adding system public parameters besides moving partial encryption computation offline. In addition, a public ciphertext test phase is performed before the decryption phase, which eliminates most of

computation overhead due to illegitimate ciphertexts. For the sake of data security, a Chameleon hash function is used to generate an immediate ciphertext, which will be blinded by the offline ciphertexts to obtain the final online ciphertexts. The proposed scheme is proven secure against adaptively chosen-ciphertext attacks, which is widely recognized as a standard security notion. Extensive performance analysis indicates that the proposed scheme is secure and efficient.

#### **Ensuring distributed accountability for data sharing in the cloud:**

Cloud computing enables highly scalable services to be easily consumed over the Internet on an as-needed basis. A major feature of the cloud services is that users' data are usually processed remotely in unknown machines that users do not own or operate. While enjoying the convenience brought by this new emerging technology, users' fears of losing control of their own data (particularly, financial and health data) can become a significant barrier to the wide adoption of cloud services. To address this problem, in this paper, we propose a novel highly decentralized information accountability framework to keep track of the actual usage of the users' data in the cloud. In particular, we propose an object-centered approach that enables enclosing programmable capabilities to both create a dynamic and traveling object, and to ensure that any access to users' data will trigger authentication and automated logging local to the JARs. To strengthen user's control, we also provide distributed auditing mechanisms. We provide extensive experimental studies that demonstrate the efficiency and effectiveness of the proposed approaches.

#### **Key-aggregate cryptosystem for scalable data sharing in cloud storage:**

Data sharing is an important functionality in cloud storage. In this paper, we show how to securely, efficiently, and flexibly share data with others in cloud storage. We describe new public-key cryptosystems that produce constant-size ciphertexts such that efficient delegation of decryption rights for any set of ciphertexts are possible. The novelty is that one can aggregate any set of secret keys and make them as compact as a single key, but encompassing the power of all the keys being aggregated. In other words, the secret key holder can release a constant-size aggregate key for flexible choices of ciphertext set in cloud storage, but the other encrypted files outside the set remain confidential. This compact aggregate key can be conveniently sent to others or be stored in a smart card with very limited secure storage. We provide formal security analysis of our schemes in the standard model. We also describe other application of our schemes. In particular, our schemes give the first public-key patient-controlled encryption for flexible hierarchy, which was yet to be known.

#### **Collective data-sanitization for preventing sensitive information inference attacks in social networks:**

**AUTHORS: Zhipeng Cai, Zaobo He, Xin Guan, and Yingshu Li,**

—Releasing social network data could seriously breach user privacy. User profile and friendship relations are inherently private. Unfortunately, sensitive information may be predicted out of released data through data mining techniques. Therefore, sanitizing network data prior to release is necessary. In

this paper, we explore how to launch an inference attack exploiting social networks with a mixture of non-sensitive attributes and social relationships. We map this issue to a collective classification problem and propose a collective inference model. In our model, an attacker utilizes user profile and social relationships in a collective manner to predict sensitive information of related victims in a released social network dataset. To protect against such attacks, we propose a data sanitization method collectively manipulating user profile and friendship relations. Besides sanitizing friendship relations, the proposed method can take advantages of various data-manipulating methods. We show that we can easily reduce adversary's prediction accuracy on sensitive information, while resulting in less accuracy decrease on non-sensitive information towards three social network datasets. This is the first work to employ collective methods involving various data-manipulating methods and social relationships to protect against inference attacks in social network

**A private and efficient mechanism for data uploading in smart cyber-physical systems:**

**AUTHORS:** Zhipeng Cai<sup>1</sup>, Xu Zheng<sup>1,2</sup>, Student Member

—To provide fine-grained access to different dimensions of the physical world, data uploading in smart cyber-physical systems suffers novel challenges on both energy conservation and privacy preservation. It is always critical for participants to consume as little energy as possible for data uploading. However, simply pursuing energy efficiency may lead to extreme disclosure of private information, especially when the uploaded contents from participants are more

informative than ever. In this paper, we propose a novel mechanism for data uploading in smart cyber-physical systems, which considers both energy conservation and privacy preservation. The mechanism preserves privacy by concealing abnormal behaviors of participants, while still achieves an energy-efficient scheme for data uploading by introducing an acceptable number of extra contents. To derive an optimal uploading scheme is proved to be NP-hard. Accordingly, we propose a heuristic algorithm and analyze its effectiveness. The evaluation results towards a real-world dataset demonstrate that the results obtained through our proposed algorithm is comparable with the optimal ones

**Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly big data:**

**AUTHORS:** [Xiaokang Zhou](#), [Wei Liang](#)

Scholarly big data, which is a large-scale collection of academic information, technical data, and collaboration relationships, has attracted increasing attentions, ranging from industries to academic communities. The widespread adoption of social computing paradigm has made it easier for researchers to join collaborative research activities and share academic data more extensively than ever before across the highly interlaced academic networks. In this study, we focus on the academic influence aware and multidimensional network analysis based on the integration of multi-source scholarly big data. Following three basic relations: Researcher-Researcher, Researcher-Article, and Article-Article, a set of measures is introduced and defined to quantify correlations in terms of activity-based

collaboration relationship, specialty-aware connection, and topic-aware citation fitness among a series of academic entities (e.g., researchers and articles) within a constructed multidimensional network model. An improved Random Walk with Restart (RWR) based algorithm is developed, in which the time-varying academic influence is newly defined and measured in a certain social context, to provide researchers with research collaboration navigation for their future works. Experiments and evaluations are conducted to demonstrate the practicability and usefulness of our proposed method in scholarly big data analysis using DBLP and ResearchGate data

**A differential-private framework for urban traffic flows estimation via taxi companies:**

**AUTHORS: Zhipeng Cai<sup>1</sup>, Xu Zheng<sup>2</sup>, Member, Jiguo Yu<sup>3,4,5</sup>,**

—Due to the prominent development of public transportation systems, the taxi flows could nowadays work as a reasonable reference to the trend of urban population. Being aware of this knowledge will significantly benefit regular individuals, city planners, and the taxi companies themselves. However, to mindlessly publish such contents will severely threaten the private information of taxi companies. Both their own market ratios and the sensitive information of passengers and drivers will be revealed. Consequently, we propose in this work a novel framework for privacy-preserved traffic sharing among taxi companies, which jointly considers the privacy, profits, and fairness for participants. The framework allows companies to share scales of their taxi flows, and common knowledge will be derived from these statistics. Two algorithms are proposed

for the derivation of sharing schemes in different scenarios, depending on whether the common knowledge can be accessed by third parties like individuals and governments. The differential privacy is utilized in both cases to preserve the sensitive information for taxi companies. Finally, both algorithms are validated on realworld data traces under multiple market distributions.

**Bitcoin: a peer-to-peer electronic cash system:**

A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a best effort basis, and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone.

**EXISTING SYSTEM:**

As early as 2015, Swan pointed out that there was not yet an acceptable “health data



common” model [15] with appropriate privacy and reward systems for public sharing of personal health data and quantified self-tracking data. Simultaneously, the author believes that blockchain can precisely provide such a structure for creating a secure, remunerated, and owner-controlled health data sharing. Zyskind et al. described a distributed personal data management system [16] that ensures users own and control their data. The system encrypts the data collected from the user’s mobile phone and stores it off-chain and only stores the data’s hash value on the blockchain. Meanwhile, two acceptable transaction types named Taccess and Tdata are defined, in which Taccess is used to implement access control management, and Tdata is used for data storage and retrieval. Azaria et al. proposed MedRec system [17], a blockchain-based decentralized record management system for electronic medical records (EMRs). MedRec provides patients with a comprehensive and immutable log, and the patients can access their medical information at any time across providers and locations. However, the system implements permissionless blockchain with PoW consensus, lacking data security, data privacy, and throughput. Xia et al. proposed MeDShare [18], a system that solves the problem of sharing medical data in a trustless environment by custodians of medical big data. Dubovitskaya et al. have proposed a framework for managing and sharing EMR data for cancer patient care [19]. It uses a permission chain to maintain metadata and access control policies and uses cloud services to store the encrypted data. Patients can define their access control policies to ensure data security and availability. The above-

mentioned data-sharing schemes based on blockchain give an ideal blueprint, but most of them only describe the scheme’s outline and do not provide the implementation details of the required protocol.

### **PROPOSED SYSTEM:**

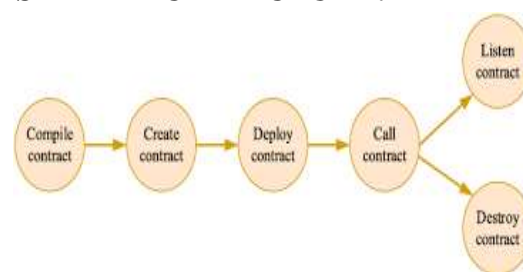
In the following years, many researchers have designed and implemented more robust access control protocols on blockchain to protect data privacy and security during sharing. Liang et al. used the consortium chain Hyperledger Fabric to realize a user-centric health data-sharing model [20] in which the cloud storage is used as a data warehouse and the blockchain ledger is constructed to store operations such as query and update. At the same time, it uses the member management service provided by Hyperledger Fabric to strengthen the users’ identity authentication and the channel model to protect users’ privacy. Fan et al. focused their attention on mobile network data sharing and privacy protection in the 5G era and proposed an efficient sharing scheme based on blockchain [21]. The main idea is to define a transaction format on blockchain to represent an access strategy. The strategy includes access requestor, content provider, visitor, and the beginning and ending time of access allowed, which is a role-based access control model. Zhang et al. proposed a blockchain-based data-sharing scheme for AI-powered network operations [22]. The scheme sets up two different types of chain, in which DataChain is used as access control tools for data, and BehaviorChain is used to store access records and ensure they cannot be tampered with. They divide access permissions into four levels. Zhou et al. proposed a blockchain-

based file-sharing system [23] to address inefficient file sharing during the review of academic papers. The scheme uses Access Control Language (ALC) to exercise access control over the information stored on-chain. It needs to define an access policy on the blockchain for each pair of users and resource. Patel proposed a crossdomain image-sharing framework based on blockchain [24], which uses blockchain as data storage and allows patients to define an access policy. They pointed out that this approach can protect the data from unrelated parties, but no research has been conducted on privacy and security. Tan et al. have proposed a blockchain-based access control scheme for Cyber-Physical Social System (CPSS) big data [25], called BacCPSS. BacCPSS uses an address of blockchain as the user's identity and maintains a user access matrix on the Smart Contract, ensuring that only operations authorized in the access matrix can be performed. The access control methods implemented in the above data-sharing schemes either need to maintain large numbers of access rules on the chain or cannot achieve fine-grained access control. Neither the access control matrix nor the RBAC is suitable for distributed environments like blockchain.

ABE is considered the most appropriate technology to solve data security and privacy protection problems in a distributed environment. Therefore, recently, researchers have used ABE to achieve fine-grained access control over data on the blockchain. Jemel and Serhrouchni proposed a decentralized access control mechanism [26]. For the first time, researchers used blockchain nodes to execute a CP-ABE algorithm to verify user

access rights' legitimacy. The scheme designs two types of transactions: SetPolicy and GetAccess. But it does not use Smart Contracts, and it is obvious that the scheme is unable to achieve more complex requirements. Sun et al. constructed a model of secure storage and effective sharing for electronic medical data based on ABE and blockchain [27], which provides better access control. Doctors use ABE to encrypt patients' medical data and store it on IPFS. However, it also does not use Smart Contracts. It only broadcasts some ABE parameters stored in transactions, which cannot achieve more complex business functions. Wang et al. proposed a sharing scheme [28] in which users distribute secret keys. It realizes that the data owner has a fine-grained access control on his data. At the same time, the Ethereum Smart Contract is used to realize the retrieval of ciphertext keywords. However, it requires multiple off-chain communication between users, and more importantly, it does not implement the permit revocation. Pournaghi et al. proposed a secure and efficient sharing scheme based on blockchain and ABE entitled MedSBA to record and store medical data [29]. It implements the update and revocation of permissions by broadcasting a new strategy to cover the previous transaction, but this will lead to users who do not want to be revoked to update their keys.

**SYSTEM ARCHITECTURE:**



**IMPLEMENTATION:**

**MODULES:**

upload MRI images dataset : use this button to get upload images.

Generate images train & test model : use this button to get generate images train & test model.

Generate deep learning CNN model : use this button to get deep learning CNN model.

Get drive HQ images: using this button to get open drive HQ

Predict tumor :use this button to get predict tumor.

**(a) Learn Linear Algebra and Multivariate Calculus**

Both Linear Algebra and Multivariate Calculus are important in Machine Learning. However, the extent to which you need them depends on your role as a data scientist. If you are more focused on application heavy machine learning, then you will not be that heavily focused on maths as there are many common libraries available. But if you want to focus on R&D in Machine Learning, then mastery of Linear Algebra and Multivariate Calculus is very important as you will have to implement many ML algorithms from scratch.

**(b) Learn Statistics**

Data plays a huge role in Machine Learning. In fact, around 80% of your time

as an ML expert will be spent collecting and cleaning data. And statistics is a field that handles the collection, analysis, and presentation of data. So it is no surprise that you need to learn it!!!

Some of the key concepts in statistics that are important are Statistical Significance, Probability Distributions, Hypothesis Testing, Regression, etc. Also, Bayesian Thinking is also a very important part of ML which deals with various concepts like Conditional Probability, Priors, and Posteriors, Maximum Likelihood, etc.

**Types of Machine Learning**

- **Supervised Learning** – This involves learning from a training dataset with labeled data using classification and regression models. This learning process continues until the required level of performance is achieved.
- **Unsupervised Learning** – This involves using unlabelled data and then finding the underlying structure in the data in order to learn more and more about the data itself using factor and cluster analysis models.
- **Semi-supervised Learning** – This involves using unlabelled data like Unsupervised Learning with a small amount of labeled data. Using labeled data



vastly increases the learning accuracy and is also more cost-effective than Supervised Learning.

- **Reinforcement Learning** – This involves learning optimal actions through trial and error. So the next action is decided by learning behaviors that are based on the current state and that will maximize the reward in the future.

**Advantages of Machine learning :-**

**1. Easily identifies trends and patterns -**

Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans. For instance, for an e-commerce website like Amazon, it serves to understand the browsing behaviors and purchase histories of its users to help cater to the right products, deals, and reminders relevant to them. It uses the results to reveal relevant advertisements to them.

**2. No human intervention needed (automation)**

With ML, you don't need to babysit your project every step of the way. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own. A common example of this is anti-virus softwares; they learn to filter new threats as they are

recognized. ML is also good at recognizing spam.

**3. Continuous Improvement**

As **ML algorithms** gain experience, they keep improving in accuracy and efficiency. This lets them make better decisions. Say you need to make a weather forecast model. As the amount of data you have keeps growing, your algorithms learn to make more accurate predictions faster.

**4. Handling multi-dimensional and multi-variety data**

Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

**5. Wide Applications**

You could be an e-tailer or a healthcare provider and make ML work for you. Where it does apply, it holds the capability to help deliver a much more personal experience to customers while also targeting the right customers.

**Disadvantages of Machine Learning :-**

**1. Data Acquisition**

Machine Learning requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality. There can also be times where they must wait for new data to be generated.

**2. Time and Resources**

ML needs enough time to let the algorithms learn and develop enough to fulfill their purpose with a considerable amount of accuracy and relevancy. It also needs massive resources to function. This can mean additional requirements of computer power for you.

**3. Interpretation of Results**

Another major challenge is the ability to accurately interpret results generated by the algorithms. You must also carefully choose the algorithms for your purpose.

**4. High error-susceptibility**

Machine Learning is autonomous but highly susceptible to errors. Suppose you train an algorithm with data sets small enough to not be inclusive. You end up with biased predictions coming from a biased training set. This leads to irrelevant advertisements being displayed to customers. In the case of ML, such blunders can set off a chain of errors that can go undetected for long periods of time. And when they do get noticed, it takes quite some time to recognize the source of the issue, and even longer to correct it.

**Results**

To run project first double click on ‘Start\_IPFS.bat’ file to start IPFS server and get below screen



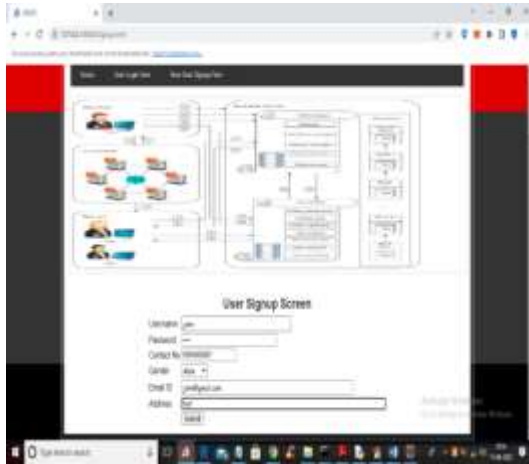
In above screen IPFS server started and now double click on ‘runServer.bat’ file to start python DJANGO server and get below screen



In above screen python DJANGO server started and now open browser and enter URL as ‘http://127.0.0.1:8000/index.html’ and press enter key to get below screen



In above screen click on ‘New User Signup Here’ link to add new user to Blockchain



In above screen user is signup and press button to get below output



In above screen user signup process completed and similarly you can add any number of users and now click on ‘User Login Here’ link to get below login screen



In above screen user is login and press button to get below output



In above screen user logged in successfully and now click on ‘Share Secured Data’ link to share data with other users



In above screen user can enter some message and then upload image and by holding CTRL KEY you can select names of users with

whom you want to share this data and press button to get below output



In above screen 'John' is sharing data with user 'aaa' and 'bbb' and both users can decrypt and view data but user 'ccc' cannot view it.

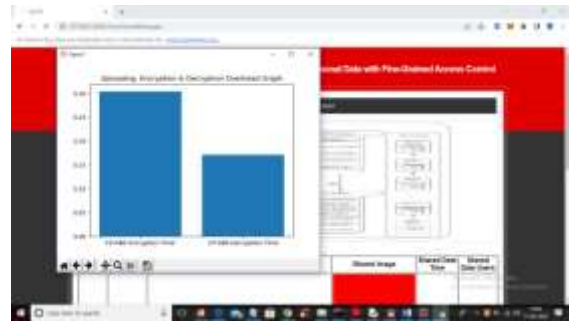


In above screen we can see sharing attributes stored at Blockchain and images and decryption keys stored at IPFS and now click on 'View Shared Messages' link to view own messages and other users shared messages so 'John' is the data owner so he can view his own upload and others shared data.



In above screen we can see data owner name, shared messages with IPFS address and we

can see names of shared users list and now we can check weather aaa or bbb can view this data or not and now click on 'Storage Overhead Graph' link to view encryption and decryption time overhead



In above screen x-axis represents encryption and decryption and y-axis represents time overhead and now logout and login as 'bbb' user to view shared data.



In above screen shared user 'bbb' is login and after login will get below output



Now in above screen 'bbb' can click on 'View Shared Messages' link to view all users shared data



In above screen ‘bbb’ can view shared data from aaa and john and now logout and login as ‘ccc’ and nobody shared data with ‘ccc’ so he cannot access any data

In above screen user ‘ccc’ is login and after login will get below screen

Now in above screen ‘ccc’ can click on ‘View Shared Messages’ link to get below output



In above screen ‘ccc’ can get empty table as nobody shared data with him. Similarly any number of user can signup and share data.

## Conclusion

In the AI-driven era, a user-centered sharing model is proposed to open data while ensuring data privacy. We combined blockchain, CP-ABE, and IPFS to propose a blockchain-based security data-sharing scheme with fine-grained access control and permission revocation. In our proposed scheme, the *DO* encrypts his data and uploads it to IPFS, then encrypts the returned address and decryption key by CP-ABE. Only *DUs* whose attributes satisfy the access policy can decrypt and obtain the data. There is no

centralized node in the scheme, and the *DO* has complete control over his shared data, which promises privacy and security. To achieve the goal, we have implemented our scheme on the EOS blockchain. The security and performance analysis proves that our scheme is feasible and practical and has a good performance. We can also add a cryptocurrency to introduce an economic system for data sharing and further enrich our scheme’s functions. At the same time, there are many shortcomings in our scheme. For example, the CP-ABE we designed with permission revocable does not have the best performance. There are also many types of research on CP-ABE . We can use a CP-ABE with better performance to improve our scheme. Besides, for the searchable encryption algorithm used in our scheme, the *DO* needs to distribute a secret key for each *DU* and store it on-chain. It also needs to maintain large amounts of indices for each shared data, which can be further optimized. At present, some researchers have proposed using blockchain to solve the fairness problem in searchable encryption algorithm . In the future, we will study and discuss the endowment of a better ciphertext searchable algorithm to further optimize our scheme. Simultaneously, to make our scheme more practical, we can combine some studies with ours and put forward a data governance scheme that is more in line with the practical application.

## REFERENCES

- [1] J. Li, Y. Zhang, X. Chen, and Y. Xiang, “Secure attribute-based data sharing for resource-limited users in cloud computing,” *Computers & Security*, vol. 72, pp. 1–12, 2018.



- [2] S. Sundareswaran, A. Squicciarini, and D. Lin, "Ensuring distributed accountability for data sharing in the cloud," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 4, pp. 556–568, 2012.
- [3] Cheng-Kang Chu, S. S. M. Chow, Wen-Guey Tzeng, Jianying Zhou, and R. H. Deng, "Key aggregate cryptosystem for scalable data sharing in cloud storage," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 468–477, 2014.
- [4] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," 2008, <https://bitcoin.org/bitcoin.pdf>.
- [5] Y. Xu, C. Zhang, G. Wang, Z. Qin, and Q. Zeng, "A blockchain-enabled duplicatable data auditing mechanism for network storage services," *IEEE Transactions on Emerging Topics in Computing*, 2020.
- [6] M. Swan, "Blockchain thinking: the brain as a decentralized autonomous corporation [commentary]," *IEEE Technology and Society Magazine*, vol. 34, no. 4, pp. 41–52, 2015.
- [7] G. Zyskind, O. Nathan, and A. Pentland, "Decentralizing privacy: using blockchain to protect personal data," in *2015 IEEE Security and Privacy Workshops*, pp. 180–184, San Jose, CA, 2015.
- [8] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "Medrec: using blockchain for medical data access and permission management," in *2016 2nd International Conference on Open and Big Data (OBD)*, pp. 25–30, Vienna, 2016.
- [9] Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani, "Medshare: trust-fewer medical data sharing among cloud service providers via blockchain," *IEEE Access*, vol. 5, pp. 14757–14767, 2017.
- [10] Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, and F. Wang, "Secure and trustable electronic medical records sharing using blockchain," *AMIA Annual Symposium Proceedings*, vol. 2017, pp. 650–659, 2017.
- [11] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, "Integrating blockchain for data sharing and collaboration in mobile healthcare applications," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5, Montreal, QC, 2017.
- [12] K. Fan, Y. Ren, Y. Wang, H. Li, and Y. Yang, "Blockchain based efficient privacy preserving and data sharing scheme of content-centric network in 5g," *IET Communications*, vol. 12, no. 5, pp. 527–532, 2017.
- [13] G. Zhang, T. Li, Y. Li, P. Hui, and D. Jin, "Blockchain-based data sharing system for AI powered network operations," *Journal of Communications and Information Networks*, vol. 3, no. 3, pp. 1–8, 2018.

- [14] 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), pp. 1–10, Gold Coast, Australia, 2019.
- [15] V. Patel, "A framework for secure and decentralized sharing of medical imaging data via blockchain consensus," *Health informatics journal*, vol. 25, no. 4, pp. 1398–1411, 2018.