BLOCKCHAIN ASSISTED COLLABORATIVE SERVICE RECOMMENDATION SCHEME WITH DATA SHARING

N.SRINIVASA RAO [1],**PEDAPUDI SATISH**[2].

[1] Assistant Professor, **DEPT OF MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh**

Email:- naagaasrinu@gmail.com

[2]**PG Student of MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh**

Email:- satishpedapudi13@gmail.com.

ABSTRACT

With the rapid development of cloud computing, a large number of web services have been emerging quickly, which brings a heavy burden for users to choose the services they preferred. In order to suggest web services for users, recommendation algorithms are needed and many of them have been investigated recently. However, most of the existing recommendation schemes are based on centralized historical data, which may lead to single point of failure. Generally, the data contains a lot of sensitive information that cloud may expose the privacy of users, which makes most cloud platforms reluctant to share their own data. In order to solve the above issues, the secure data sharing among cloud platforms is necessary for better recommendation, which can maximize the profits. In this paper, we propose a blockchain-assisted collaborative service recommendation scheme (*BC-SRDS*). Specifically, we adopt the ciphertext-policy attribute-based encryption (CP-ABE) algorithm to encrypt the data, which ensures the data confidentiality and realizes secure data sharing. Then, we utilize the blockchain to share data, such that the DoS attack, DDoS attack and single point of failure can be avoided. Meanwhile, the data integrity, tampering-proof of data are guaranteed through the blockchain. And we use locality-sensitive hashing algorithm to recommend the services for users. Finally, it is proved through the security analysis that *BC-SRDS* is capable of achieving data confidentiality, data integrity and tampering-proof. A series of experiments show that *BC-SRDS* achieves better recommendation accuracy compared with the existing schemes.

I. INTRODUCTION

For the rapid development of the Internet and computer technology, a large number of network information services have entered people's daily life, providing users with many conveniences. Meanwhile, due to the rapid growth of Internet users, the information generated by users also presents an explosive growth, leading to the problem of ``information overload''. It is an important challenge for producers and consumers in today's society: for Internet service providers, it is very difficult for Internet service providers to select the services through user information quickly and make their services popular with the public. For Internet users, it is not easy to select the services they are interested in from the vast amount of information, which requires a lot of time and energy.

In order to solve above problems, a lot of recommendation algorithms have been proposed, among which collaborative filtering recommendation is a common method. It provides personalized recommendation for target user according to the scoring records of resources. Moreover, the algorithm has a good recommendation effect and is widely used in personalized sites, e-commerce, and other fields. Although the collaborative filtering recommendation algorithm has high recommendation accuracy, it still faces a series of challenges: the data used in the collaborative filtering recommendation algorithm is often stored on the centralized server, so there is no historical data for new users or new items to refer to, and it may encounter cold start problem and cannot complete the recommendation. However, by collecting the data of users on different platforms, the problem of cold start can be solved effectively. For example, user *A* has called the services of Amazon and user *B* has called the services of IBM. If *A* and *B* are similar users, they can recommend services to *A* by analyzing the services of *B*, or recommend services to *B* by analyzing the services of *A*. Nonetheless, a malicious recommendation may lead to data abuse or privacy leakage issues. Therefore, in order to protect users' privacy information, Amazon and IBM are reluctant to share user data with each other. In this case, we cannot obtain similar users of new users, which greatly reduces the quality of recommendation. Besides, assume that two platforms agree to share data, and obtain the data from different platforms to improve the recommendation accuracy. But there still exists the problem that the data stored in different distributed platforms greatly increases the communication overhead between cloud platforms, which cannot response quickly when online.

In order to solve the above problems, in this paper, we propose a novel block chain-assisted collaborative service recommendation scheme with data sharing (*BC-SRDS*), with which every platform can share their data securely based on block chain. And locality-sensitive hash (LSH) is an efficient data search algorithm that can quickly find similar data, therefore, the LSH is adopted to perform quick recommendation. The main contribution can be summarized as follows V

1) To the best of our knowledge, many recommendation algorithms are based on centralized data, but little of them adopt block chain to realize the recommendation. Moreover, it is difficult to combine distributed cloud platforms to share their data with each other for recommendation because of users' privacy. In order to realize better recommendation, in this paper, we introduce block chain to provide a secure sharing environment for accurate recommendation.

2) Different from most existing recommendation works, we adopt the CP-ABE (cipher text-policy attribute based encryption) to promote the data sharing among the cloud platforms and combine block chain to ensure the security of data provenance, avoid the risk on the failure of single point, improve data integrity, and defend against DOS or DDOS attacks.

3) Based on the real distributed QOS (quality of service) dataset WS-DREAM,1 the experimental results show the effectiveness of *BC-SRDS*. The analysis of the metrics such as CPU consumption, memory consumption, throughput, latency, MAE (Mean Absolute Error), RMSE (Root Mean Square Error), the usage of gas and the number of similar neighbors proves that *BC-SRDS* can significantly improve the accuracy and gain mor pro_ts.

II. SYSTEM ANALYSIS

EXISTING SYSTEM

- Sarwar *et al.* [6] proposed a collaborative filtering algorithm based on items, which mainly calculates the similarity between items by analyzing the user's historical data. Because the number of users are much larger than the number of items, the computing of similarity between items is less sparse than that between users. Meanwhile, because the properties of items are fixed, we can calculate the similarity between items offline. Moreover, the calculation performance between items is high. Compared with collaborative filtering algorithm based on users, the recommendation quality is greatly improved. However, the item-based collaborative filtering algorithm also brings new challenges: Firstly, the algorithm does not consider the difference between users, thus, the recommendation quality is poor; secondly, when a new item is added to the system, it is difficult to recommend the items that are similar to the new item to users, because it is not scored or there are few scores about the new item.

- In order to solve these problems, Jiang *et al.* [4] proposed a hybrid recommendation algorithm, which takes the advantages of the two algorithms that are item-based collaborative filtering algorithm and user-based collaborative filtering algorithm. They use the characteristics of attributes of users or items to obtain a better recommendation effect. However, the algorithm is based on specific applications and it is difficult to transplant and expand.

- Cheng *et al.* [7] proposed a method to improve the recommendation quality by introducing metric learning into collaborative filtering algorithm. By measuring the distance between target user and candidate set, it separates th candidate set. And the items that have high similarity to user preference are close to the users while the items that have low similarity to user preference are far away from users, which reduces the influence of sparse data on the recommendation. Although the algorithm can alleviate the influence of sparsity on the recommendation, but it does not propose an effective method to deal with sparse data.

- In [8], Mnih proposed the matrix decomposition algorithm, which is one of the most popular collaborative filtering algorithms based on model. Compared with user-based collaborative filtering algorithm, it shows better performance in dealing with sparse problems. In [9], Yu *et al.* added the dimension of geographic location. They pointed out that users in similar geographical location usually invoked the same service, and they often have the same service requirements. Besides, the quality of web service is highly related to time.

Disadvantages

- The system is less secured since the system is not implemented Collaborative service recommendation.
- The system is not implemented CP-ABE (ciphertext-policy attribute based encryption)which is to promote the data sharing among the cloud platforms and combine blockchain to ensure the security of data provenance

Proposed System

- The system proposes a blockchain-assisted collaborative service recommendation scheme with data sharing. Different from the scheme4 in [25], the system adopts CP-ABE to realize secure data sharing so that we can share data securely while guaranteeing that the cloud platforms can control their own data. CP-ABE embeds the access strategy into the ciphertext, which means that the data owners can determine that who can access these user data. Further, it prevents user historical service data from being illegally accessed. Therefore, the scheme can meet the needs for security of user historical service data.

- ❖ Moreover, our scheme also introduces blockchain to realize data sharing. Its features of tampering proof, decentralization can further ensure the security of ciphertexts. Our scheme can easily detect whether the data has been tampered with. Therefore, our scheme is secure for recommendation. Moreover, the system model of our scheme is shown in this system. In this scheme, it mainly includes three layers: User layer, Data sharing layer, and Data layer.

1) User layer: This layer mainly includes various devices belonged to users, through which users can access the web services. They will produce large-scale data on the cloud platforms. Note that these users are distributed in different platforms.

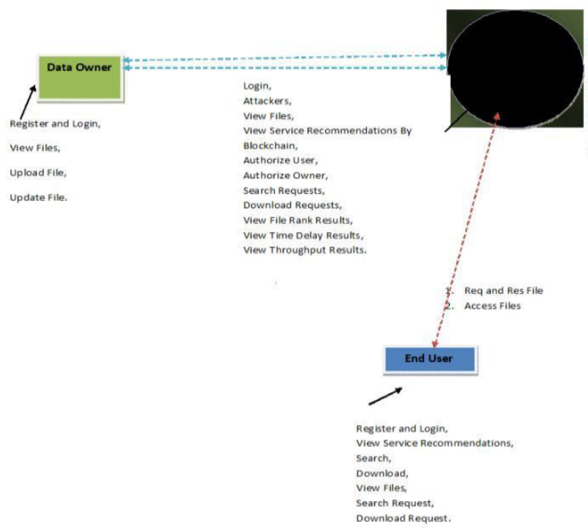
2) Data sharing layer: The cloud platforms that agree to share the data construct a consortium blockchain, on which they can share and utilize message. They will record the shared data on the consortium blockchain. Using the data on consortium blockchain, the platforms can merge it with their own data to perform accurate recommendation, obtaining more profits. Every platform maintains a public ledger together.

3) Data layer: In this layer, assisted with the cloud servers, we store the large les (e.g. videos, images) on it, because blockchain is not suitable for storing large amounts of data.

Advantages

- ➢ The system proposes for **Defending against DoS or DDoS attacks.**
- ➢ **Tamper-proofing.** All nodes should hardly tamper with the shared data. Once the shared data is tampered with by a node, the data can be detected.
- ➢ **Avoiding single point of failure.** The scheme should meet the requirement that when a cloud platform was attacked, other cloud platforms can also run normally.
- ➢ **Data integrity.** When the shared data is tampered with,the scheme can quickly determine whether the data is complete.

Architecture Diagram



III. IMPLEMENTATION

- **Data Owner Module**

In this module, the data owner uploads their data in the cloud server. For the security purpose the data owner encrypts the data file's blocks and then store in the cloud. The data owner can check the replication of the file's blocks over Corresponding cloud server. The Data owner can have capable of manipulating View Files, Upload File, and Update File.

- **Cloud Server Module**

The cloud service provider manages a cloud to provide data storage service. Data owners encrypt their data file's blocks and store them in the cloud for sharing with Remote User. To access the shared data file's blocks, data consumers download encrypted data file's blocks of their interest from the cloud and then decrypt them and perform the following operations such as View Attackers, View Files, View Service Recommendations By Blockchain,Authorize User, Authorize Owner, Search Requests, Download Requests, View File Rank Results, View Time Delay Results, View Throughput Results.

- **End User**

In this module, remote user logs in by using his user name and password. After he will request for secrete key of required file's blocks from cloud servers, and get the secrete key. After getting secrete key he is trying to download file's blocks by entering file's blocks name and secrete key from cloud server and performs the following operations View Service Recommendations, Search, Download, View Files, Search Request, Download Request.

- **Data Encryption and Decryption**

All the legal users in the system can freely query any interested encrypted and decrypted data. Upon receiving the data from the server, the user runs the decryption algorithm Decrypt to decrypt the cipher text by using its secret keys from different Users. Only the attributes the user possesses satisfy the access structure defined in the cipher text CT, the user can get the content key.

IV. CONCLUSION

In this project, we propose a service recommendation scheme named *BC SRDS* which not only supports the data sharing among different platforms based on the consortium block chain but also provides an accurate service recommendation for users. Moreover, to guarantee data security, we encrypt the data by CP-ABE algorithm before sharing them to other cloud platforms. Through block chain, the cloud platforms can easily get the shared data and use it to maximize the pro_ts, meanwhile, the DOS attack, DDOS attack and single point of failure are avoided. The security analysis show that *BC-SRDS* is capable of achieving data confidentiality, data integrity and tampering-proof. Finally, we evaluate our scheme based on WS-DREAM and carry out a series of experiments. The experimental results show that *BC-SRDS* can achieve a higher accuracy than other three schemes. Besides, its gas cost is entirely acceptable for the cloud platforms. Moreover, from metrics of the resource consumption, throughput and latency on consortium block chain, we can conclude that our scheme based on the consortium block chain is feasible.

REFERENCES

[1] A. Popescul, L. H. Ungar, D. M. Pennock, and S. Lawrence, ``Probabilistic models for uni_ed collaborative and content-based recommendation in sparse-data environments," 2013, *arXiv:1301.2303*. [Online]. Available: http://arxiv.org/abs/1301.2303

[2] G. Arora, A. Kumar, G. S. Devre, and A. Ghumare, ``Movie recommendation system based on users similarity," *Int. J. Comput. Sci. Mobile Comput.*, vol. 3, no. 4, pp. 765_770, 2014.

[3] J. Bobadilla, F. Ortega, and A. Hernando, ``A collaborative _ltering similarity measure based on singularities," *Inf. Process. Manage.*, vol. 48, no. 2, pp. 204_217, Mar. 2012.

[4] C. Jiang, R. Duan, H. K. Jain, S. Liu, and K. Liang, ``Hybrid collaborativ _ltering for high-involvement products: A solution to opinion sparsity and dynamics," *Decis. Support Syst.*, vol. 79, pp. 195_208, Nov. 2015.

[5] H. G. Rong, S. X. Huo, C. H. Hu, and J. X. Mo, ``User similarity-based collaborative _ltering recommendation algorithm,'' *J. Commun.*, vol. 35, no. 2, pp. 16_24, 2014.

[6] K.-Y. Chung, D. Lee, and K. J. Kim, ``Categorization for grouping associative items using data mining in item-based collaborative _ltering,'' *Multimedia Tools Appl.*, vol. 71, no. 2, pp. 889_904, Jul. 2014.

[7] C.-K. Hsieh, L. Yang, Y. Cui, T.-Y. Lin, S. Belongie, and D. Estrin, ``Collaborative metric learning,'' in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 193_201.

[8] A. Mnih and R. R. Salakhutdinov, ``Probabilistic matrix factorization,'' in *Proc. Adv. Neural Inf. Process. Syst.*, 2008, pp. 1257_1264.

[9] C. Yu and L. Huang, ``A Web service QoS prediction approach based on time- and location-aware collaborative _ltering,'' *Service Oriented Comput. Appl.*, vol. 10, no. 2, pp. 135_149, Jun. 2016.

[10] X. Zhang, Z.Wang,W. Zhang, and F. Yang, ``A time-aware QoS prediction approach toWeb service recommendation,'' in *Proc. 4th Int. Conf. Comput. Eng. Netw.* Cham, Switzerland: Springer, 2015, pp. 739_748.

[11] H. Li, ``Learning to rank for information retrieval and natural language processing,'' *Synth. Lect. Hum. Lang. Technol.*, vol. 4, no. 1, pp. 1_113, Apr. 2011.

[12] S. Badsha, X. Yi, I. Khalil, D. Liu, S. Nepal, E. Bertino, and K.-Y. Lam, ``Privacy preserving location-aware personalized Web service recommendations,'' *IEEE Trans. Services Comput.*, early access, May 22, 2018, doi:10.1109/TSC.2018.2839587.

[13] J. Zhu, P. He, Z. Zheng, and M. R. Lyu, ``A privacy-preserving QoS prediction framework for Web service recommendation,'' in *Proc. IEEE Int. Conf. Web Services*, Jun. 2015, pp. 241_248.

[14] T. R. Hoens, M. Blanton, A. Steele, and N. V. Chawla, ``Reliable medical recommendation systems with patient privacy,'' *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 4, pp. 1_31, Sep. 2013.

[15] H. Polat and W. Du, ``Privacy-preserving collaborative _ltering using randomized perturbation techniques,'' in *Proc. 3rd IEEE Int. Conf. Data Mining*, Nov. 2003, pp. 625_628.