A Contribution Evaluation and Reward Framework Based on Shapley Value and Smart Contract

Shengdong Chen[1], Zhanwei Guo[2], Tianyou Liu[3],Tianwen Zhang[4*]

[1] Guangzhou Institute of Software, 510000, China

[2] Zhongke Software Testing (Guangzhou) Co., Ltd., 510000, China

[3] Guangzhou Zhiying Weilai Technology Co., Ltd.,Guangzhou,510000,China

[4] Hebei Zhuowei 'er information technology Co., LTD., 073000, China

Corresponding Email:zhangtianwen1988@gmail.com

https://doi.org/10.70695/0k3pep85

Abstract

Currently, Shapley value-based contribution assessment in incentive mechanisms has become a major research focus due to its advantages in explaining model decisions. This approach ensures the effectiveness and fairness of contribution assessment algorithms. However, there is a significant issue regarding the reliability of the computational results. Additionally, traditional Shapley value-based contribution processes face an issue where computational complexity increases exponentially with the number of participants. Addressing these issues, this paper proposes a contribution assessment and reward framework based on Shapley value and smart contract technology using alliance blockchain. This framework overcomes the reliance on third-party institutions that is characterized by traditional incentive mechanisms, leveraging the decentralized nature of blockchain to eliminate trust issues associated with a single centralized entity. Furthermore, the traceability of blockchain ensures the transparency and accountability of the contribution assessment and reward distribution processes.

Keywords Federated Learning; Incentive mechanism Shapley value; Smart contract Contribution Evaluation

1 Introduction

Federated Learning (FL) is a machine learning technique that enables multiple participants to collaboratively train a shared model while preserving privacy. Under this mechanism, each client retains a local copy of the global initial model and trains it using their own local data [1]. After training, the updated model parameters are uploaded to a central aggregation server. In recent years, federated learning has been widely applied across various scenarios, making incentive mechanisms more significant. This research aims to address critical issues in traditional incentive models and introduces a novel framework combining Shapley value and smart contracts to improve trust, efficiency, and fairness in federated learning. In the medical field, incentive mechanisms based on contribution assessment are used to encourage hospitals to share their computational resources and high-quality data, collectively developing more accurate disease prediction models, thereby enhancing the accuracy and reliability of the models [2].

However, the implementation of incentive mechanisms faces several challenges in practice. Participants in federated learning vary significantly in terms of data quality, data volume, and computational resources [3]. Different participants may come from various geographic locations, industries, or organizations, resulting in diverse types, scales, and qualities of collected data. Federated learning scenarios often involve processing large amounts of data. The high number of participants leads to substantial computational demands for contribution assessment. Incentive mechanisms in practical federated learning applications typically need to simultaneously achieve multiple objectives. For instance, an incentive mechanism may aim to encourage participants to provide high-quality data while ensuring privacy protection during data processing and model training, and maintaining fairness in reward distribution to prevent manipulation [4].

Blockchain technology provides a solution in which smart contracts can automatically distribute rewards to clients based on their contributions, using cryptocurrencies or tokens as forms of compensation [5]. Furthermore, blockchain's decentralized nature reduces dependence on central servers, effectively preventing single point of failure issues. Blockchain's immutability ensures both traceability and auditability of each client's contributions during the federated learning process. Therefore, applying blockchain technology to incentive mechanisms in federated learning has become a prominent research focus in both academic and industrial fields [6].

Traditional incentive mechanisms often rely on centralized third-party institutions for contribution assessment and reward distribution, leading to trust issues and potential data privacy breaches, which hinder active participant engagement.. Therefore, to effectively encourage entities with abundant data and resources to participate actively in federated learning, a framework based on Shapley value and smart contracts has been proposed. This framework leverages consortium blockchain and smart contract technologies to address trust dependencies and the credibility of computation results inherent in traditional incentive mechanisms. Additionally, this study introduces a multiple bootstrap truncated gradient Shapley algorithm, which optimizes the computation process of Shapley value, thereby further enhancing the efficiency and accuracy of the incentive mechanism.

2 Methodology and Data

Within the contribution assessment and reward framework, participants who provide personal data serve as the actual data contributors in the federated learning system, while computing nodes are responsible for calculating the Shapley value. In a blockchain network, the model gradient update information from participants along with the contribution calculation results from the computing nodes will be recorded in transactions on the blockchain. The model gradient update information trained using participants' own data will be uploaded to the smart contract by the FL server, allowing it to synchronize with the blockchain network. The Shapley value of the participants will also be stored in the blocks and permanently preserved on the blockchain network, ensuring that the values are tamper-proof.

In federated learning systems, model updates must be recorded and stored on the blockchain, generating large amounts of data during training. This imposes high demands on the consensus efficiency of the blockchain. Furthermore, the participant model gradient information stored on the blockchain involves the privacy of participants' data and cannot be accessed freely by external parties. Therefore, traditional public chain architectures based on consensus algorithms such as Proof of Work are no longer suitable for current needs. A consortium blockchain, which restricts data visibility to consortium nodes and offers rapid consensus mechanisms, is evidently more appropriate. The information on nodes in a consortium blockchain is genuine, backed by real entities, and is trustworthy. Additionally, consortium blockchains have robust access control strategies, and the aforementioned characteristics of consortium chains meet the needs for both the privacy of participant model data and the rapid consensus of blockchain. Currently, there are many consortium blockchain platforms supporting smart contracts, with well-known examples including Hyperledger Fabric and FISCO BCOS, which each support the deployment and operation of smart contracts. In this section's experiment, FISCO BCOS will be used as the underlying blockchain platform, and smart contracts will be implemented using Solidity.

To further evaluate the performance of FISCO BCOS, the following experiments will be conducted using Hyperledger Caliper to assess the performance of the FISCO BCOS blockchain and smart contracts. Hyperledger Caliper is a versatile blockchain performance testing framework that allows users to test the performance of various blockchains using custom use cases. For FISCO BCOS, the first step is to deploy the FISCO BCOS platform. Next, a new network configuration file needs to be added, along with the creation of a test script that includes initialization, execution, and completion phases. Finally, this new test script will be incorporated into the test configuration file, ensuring that the correct callback is specified for Caliper. To assess the performance of the FISCO BCOS platform, two fundamental interfaces, "set" and "get," have been prepared. The "set" interface is responsible for generating and deploying the "hello world" smart contract, while the "get" interface is tasked with calling the smart contract and outputting "hello world." The number of transactions will be set to 5000, and the tests will then be conducted.

3 Results

The results indicate that as the sending rate increases, the TPS of the "get" interface increases linearly, while the TPS of the "set" interface reaches a bottleneck at around 1500 (see Figure 1a). In terms of delay, the latency of both the "set" and "get" interfaces remains largely unaffected by changes in the sending rate. The latency of the "set" interface is slightly higher than that of the "get" interface, which exhibits nearly negligible delay (Figure 1b).

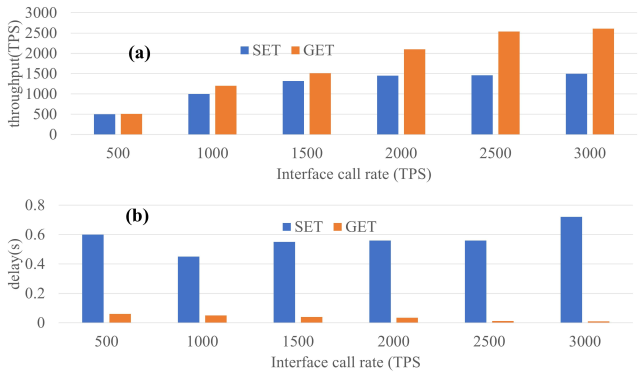
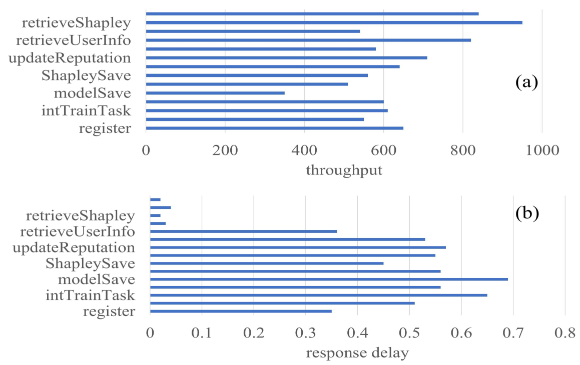


Fig. 1. FISCO BCOS performance test

After deploying smart contracts for applications and production environments, the performance declined due to increased business complexity, but still maintained a high level.

Figure 2 (a) shows the throughput of the interfaces provided by each major smart contract when the number of requests reaches 1000. It can be seen that the throughput of a single interface in the system is between 300 times/second and 900 times/second. Figure 2 (b) shows the latency of each interface when the caller is 100 at the same time. Experiments have shown that smart contracts built on FISCO BCOS have lower latency. The lower latency and reasonable throughput indicate that the current smart contract platform is sufficient to meet the performance requirements of computing tasks.



(a) The average throughput of the interface, (b) the average response time of the interface

Fig. 2. The interface performance of smart contracts when the number of client requests is 1000

4 Conclusion

We conducted performance testing on the alliance blockchain FISCO BCOS and Solidity based smart contracts for this framework. The experimental results indicate that the performance of the FISCO BCOS consortium blockchain is sufficient to meet the contribution computing performance requirements under multitasking. Smart contracts based on Solidity may experience a slight performance degradation due to the complexity of the business, but they are still sufficient to meet the required performance. This proves the performance superiority of FISCO BCOS consortium chain and smart contracts.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Fadlullah Z M, Kato N. HCP: Heterogeneous computing platform for federated learning based collaborative content caching towards 6G networks. IEEE Transactions on Emerging Topics in Computing, 2020, 10(1): 112-123.
2. Rodrigues T K, Suto K, Nishiyama H, et al. Machine learning meets computation and communication control in evolving edge and cloud: Challenges and future perspective. IEEE Communications Surveys & Tutorials, 2019, 22(1): 38-67.
3. Powell K. Nvidia clara federated learning to deliver ai to hospitals while protecting patient data. Nvidia Blog, 2019: 512.
4. Liu Z, Chen Y, Zhao Y, et al. CAreFL: Enhancing smart healthcare with Contribution-Aware Federated Learning. AI Magazine, 2023, 44(1): 4-15.
5. Zhao Y, Zhao J, Jiang L, et al. Privacy-preserving blockchain-based federated learning for IoT devices. IEEE Internet of Things Journal, 2020, 8(3): 1817-1829.
6. Zhao B, Liu X, Chen W. When crowdsensing meets federated learning: Privacy-preserving mobile crowdsensing system. arXiv preprint arXiv:2102.10109, 2021.