Blockchain-based Big Data Transaction Pricing Mechanism

Huimin Liu

***Abstract* — With the rapid development of information technology, the amount of data traffic on the network is increasing day by day, and big data transactions have emerged as a new business model. However, at present, there is little historical data on big data transactions, and the transaction mechanism is not perfect, and the traditional big data transaction pricing scheme shows many problems, such as: unfairness of the pricing mechanism, collusion and third-party hidden dangers, which cause poor circulation of data resources and the inability to maximize the value of data. In order to solve the above problems, this paper studies and designs a blockchain-based big data transaction pricing mechanism (BDPM) based on the sealed auction theory, combined with blockchain technology, SGX technology and IPFS.**

***Index Terms*—big data transactions,sealed auctions,big data pricing,blockchain.**

I. Introduction

With the development of the global digital economy, the huge profit value hidden in data is gradually revealed, and new business models for big data transactions have emerged. In recent years, in order to promote economic development, scientific and technological innovation and social progress, countries attach great importance to the construction of big data trading platforms, with the rapid construction and development of big data trading platforms, the value of big data formation has been rapidly improved, which makes big data transactions in the big data industry more and more important. In the process of trading big data as a commodity, the price of data commodities reflects the real value of big data to a certain extent, and reasonable pricing can make the optimal allocation of data resources and maximize the profit value, so it is urgent to formulate a fair and reasonable big data transaction pricing mechanism.

II. Background

In order to avoid the dominance of data prices completely in the hands of sellers, resulting in high prices, this article will take the form of auction to allow both buyers and sellers to participate in the pricing process, so that the pricing is more fair.

Regarding the research on data auction pricing, Wei X et al. mentioned that in the big data trading market, auction is the most effective method to achieve the optimal allocation of data resources due to the relativity of data value and the information asymmetry between data buyers and sellers [1]. Cai et al. propose a real dual auction mechanism for the data trading market, which solves the diversified market preferences faced by big data auction pricing, the complex conflict of interest relationship of data consumers, and the strategic behavior problems of both parties[2]. Chen et al. propose an iterative dual auction mechanism that solves the problem that the pricing rules of buyers and sellers are designed to induce participants to submit quotations and determine the amount of transaction data between buyers and sellers and their prices[3]. An et al. propose a multi-round pseudonymous proof data auction scheme to protect the best auction results from manipulation by pseudonymous bidding attacks[4]. CAO et al. propose an iterative auction mechanism to solve the data transaction problem in a data marketplace with multiple data owners, collectors, and users [5]. Bichler M et al. mainly analyzed the protection of buyers' rights and interests and maximized the benefits of buyers, and proposed that the scheme of auctioning with multiple indicators is more conducive to the successful execution of the auction than the auction scheme that only analyzes the price [6]. S.D. Jap analyzes the impact of dynamic pricing between sellers and buyers when participating in the auction, and analyzes the impact of sealing the buyer's bid price during the auction process [7]. In 2021, Shih D H used auctions for data price pricing, implemented an auction platform with price recommendations through SGX and blockchain combination, used SGX to execute price prediction algorithms off-chain, announced reasonable prices to buyers, and reduced losses caused by the winner's curse, but the auction did not mention how to solve the problem of collusion attacks and denial [8]. In general, in the existing big data transaction auction pricing research scheme, the research focus is mainly divided into two aspects, on the one hand, to study the rationality and effectiveness of auction pricing methods, on the other hand, it is mainly based on practical experience and existing strategies to solve technical problems in the auction pricing process.

However, the current big data auction pricing mechanism generally has two problems, one is the problem of collusion in the transaction process; Second, from the perspective of the privacy of bidding data, the privacy of failed data buyers cannot be guaranteed; The third is the data retention problem generated by the data auction platform; The fourth is the problem of repudiation by sellers or buyers. Therefore, this paper introduces IPFS to store ciphertext data commodities to solve the data retention problem of auction platform, and introduces SGX [9], [10], [11] to realize sealed auction to solve the privacy problem of buyers' bidding data. A new blockchain-based Dig Data Transaction Pricing Mechanism (BDPM) is proposed, which solves the problems of collusion,

Manuscript received December 07, 2022

Huimin Liu School of Computer Science and Technology,Tiangong University,Tianjin,China

repudiation and third-party hidden dangers in the current data transaction pricing.

III. RELATED WORK

A. *BDPM General Frame*

The overall framework of BDPM for blockchain-based big data transaction pricing mechanism is shown in Figure 1: There are 6 roles, 4 entity roles: data seller, data buyer, trusted node, arbitration agency, and 2 non-entity roles: IPFS and smart contract.

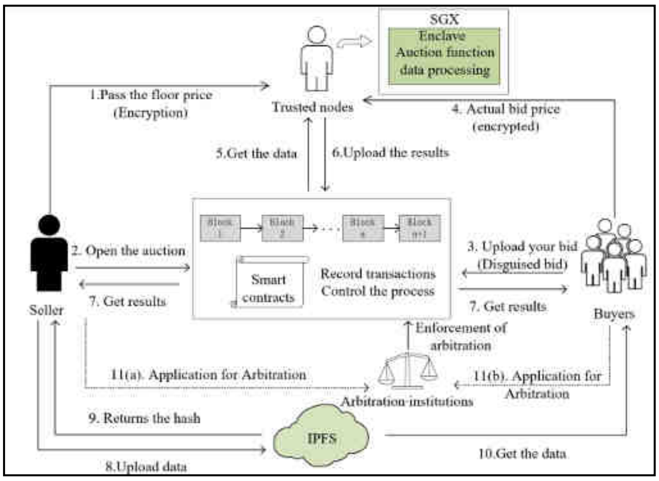

Figure 1. Overall frame diagram of BDPM

The detailed definition and description of the participating roles in the big data transaction pricing mechanism are as follows:

1) Sellers: the data owner, and hopes to sell his own data through auction.

2) Buyers: Data buyers, and want to buy data by participating in the auction.

3) Trusted nodes: Nodes in the blockchain that have SGX and want to get paid for performing tasks.

4) Arbitration institutions: professional institutions that only provide arbitration services as trusted by the buyers and sellers of data.

5) IPFS: To store big data to reduce the burden on the chain.

6) Smart contract: it is used to connect buyers and sellers with trusted nodes, control the auction process and record transaction information, manage the seller's margin and the bid price of the buyers, and provide fund settlement services.

Because the big data transaction pricing mechanism proposed in this paper BDPM is based on block chain, the parties involved in the transaction pricing need to have block chain encryption currency account, and the entity role involved in BDPM transaction encryption and signature operation, so in the subsequent research assume the entity role in the model have cryptocurrency account and AES symmetric encryption algorithm key, and have RSA public and private key pair for signature and transaction.

B. *Process Design*

The BDPM blockchain-based big data transaction pricing mechanism proposed in this paper adopts a single round of sealed bidding, which is characterized by a sealed base price and bidding data in the bidding process. According to the design idea and the overall framework diagram, the general process of BDPM can be divided into four stages, described as follows:

1) Transaction preparation stage:

The data seller obtains the response of the trusted node by publishing the auction task. The seller selects the trusted node and initiates the remote authentication, which safely transmits the key that can be used to decrypt the base price to the secure area of the trusted node. In addition, the encrypted base price is also transmitted to the trusted node. Later, the data owner calls the smart contract to release the information related to the commodity auction, including the information of the selected trusted node, the hash value of the explicit data, the deposit, the remuneration, the auction time, etc.

2) Auction pricing stage:

Data buyers decide whether to participate in the auction by looking at the big data auction information on the blockchain. Interested buyers request to participate in the big data auction by paying the transaction margin to the smart contract. The margin is larger than the real bid to hide the real bid, and when the buyer wants to deny it, they can directly deduct the corresponding amount from the deposit to prevent denial. In addition, the data buyer establishes a safe channel through remote authentication and the trusted node designated by the seller, and securely transmits the key that can decrypt the bid price to the secure area of the trusted node, and transmits the sealed real bid price to the trusted node.

Auction time deadline, when the credible node SGX has the seller's floor price and buyer bid price and the corresponding decryption key, trusted node inside the security zone price, bidding, and query the buyer on the smart contract deposit, executed within the security zone against the auction pricing algorithm, will be greater than the base price and lower than the disguised bid real bid price sorting, the maximum for the commodity price, the corresponding buyers will get the auction victory, trusted node will auction results using security zone internal private key signature, uploaded to the smart contract.

3) Settlement and delivery:

Considering that the volume of data to be traded may be large, the AES symmetric encryption method is adopted to encrypt the data. After the previous stage, the pricing of the big data goods and the winner have been determined, and then the data seller encrypts the data with their own AES key, uploads the encrypted data to the IPFS and obtains the storage address. The data seller then views the winning buyer's information through the smart contract, encrypts his AES key with the public key of the winning buyer, and then passes the encrypted key and the data storage address to the winning buyer. The buyer obtains the ciphertext data from the IPFS, decrypts the explicit data, calculates the hash value of the explicit data, and determines whether it is the hash value that is the same as the seller. If the verification is successful, the correct data will be confirmed to the smart contract, and the smart contract will conduct the fund settlement between the buyer and the seller and the trusted node.

4) Arbitration:

Before the fund settlement, if the ash value verification failure, the buyer judge data problems, can submit arbitration, the arbitration institution accept arbitration application and according to the evidence submitted for arbitration, after the arbitration, the arbitration institution call contract arbitration award, if the verification is no problem, will charge arbitration fee to the arbitration failure party, transaction according to the original plan for settlement, if the verification has a problem, will deduct the deposit of dishonest participants, paid to other participating entities.

IV. FEASIBILITY ANALYSIS

A. *Performance Analysis*

Compared with the traditional big data transaction pricing mechanism, the changes in the big data transaction pricing mechanism proposed in this paper mainly focus on the introduction of SGX technology to solve the pricing problem of sealed auction pricing, and the addition of SGX remote authentication and SGX sealing function for the transmission and storage of key data. Therefore, the data seller or the data buyer and the trusted node to establish a safe channel through SGX remote authentication, to transmit and store the key time is the focus of the performance test of this mechanism.

Since there are multiple data buyers based on the auction theory, the number of buyers n is set to 20,40,60,80 and 100 respectively to test the time of remote authentication between the buyers and sellers and the trusted node. The test results are shown in Figure 2:

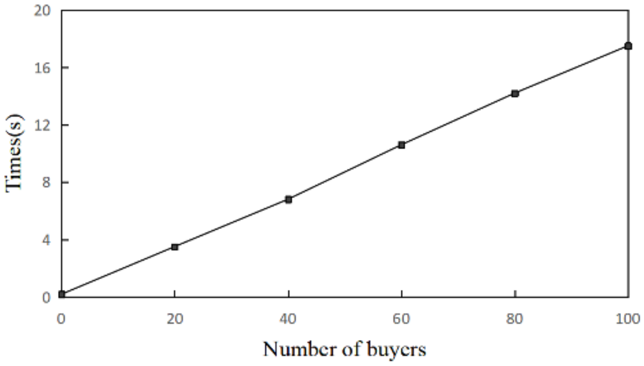

Figure 2. Key transfer time

It can be concluded from Figure 2 that with the increase of the number of data buyers participating in the auction, the time of key transmission and storage also increases, because each data buyer needs to remotely authenticate with the trusted node. In addition, when the number of buyers is 1, only the data sellers need to carry out remote authentication with the trusted node.

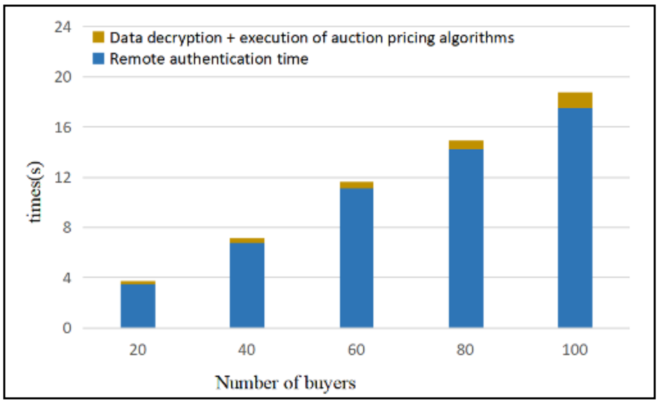

Figure 3. Time for the critical operations in the SGX

Figure 3 shows the time needed for the key operation in SGX, where the time required for data decryption and auction algorithm execution is similar to the traditional scheme, so the extra time of the scheme is the time, when the buyer number of 50, the total time is less than 9 seconds, for the trading role to ensure the data transaction fairness and data privacy, the extra time spent by the remote authentication in this mechanism is within the acceptable range.

B. *Security Analysis*

In this section, the security analysis of the BDPM based on the threat model and security assumptions.

1) Malicious data seller: After receiving the payment from the data buyer, the malicious data seller may not provide the data or provide incomplete data. In this mechanism, the data seller needs to take their data to the hash value and publish the hash value to the chain. Therefore, if the seller changes the data, the buyer will find the malicious behavior of the seller by verifying the hash value. In addition, the data seller needs to upload the margin at the beginning of the transaction, and if he / she is dishonest, the margin will be deducted and compensated to other entities involved in the transaction.

2) Malicious data buyer: do not pay fees to the data seller after receiving the data. In this mechanism, the data buyers need to be at the beginning of the transaction will hide bid storage to the smart contract, and requires the buyer cover bid higher than the real bid, so when the buyer wins, the cover bid must be higher than the data pricing, when payment, smart contract will deduct from the buyer cover bid data price amount, in order to prevent data buyers against goods costs.

3) Malicious trusted nodes: maliciously trusted nodes tamper with the auction results or obtain and leak the explicit data. The explicit data, including the disclosure of the buyer or seller's key, the seller's base price and the buyer's bid price. In this mechanism, the keys of both the buyer and the seller are transmitted to the Enclave security zone of the trusted node through the secure channel established after remote authentication, and the trusted node cannot be stolen. The seller's base price and the buyer's bid price are transmitted in the form of ciphertext, and are only decrypted in the Enclave security zone of the trusted node, so the trusted node cannot steal the relevant plaintext data. In addition in the mechanism of the seller at the beginning of the transaction the trusted node SGX internal public key storage in the smart contract, trusted nodes need to use private key will perform the result of the auction pricing algorithm for signature, because of SGX

unique security features the whole signature process is invisible for trusted nodes, after the signature, smart contract will use the data seller of the auction results information signature verification, verification through, the result is regarded as effective, so the trusted node cannot successfully tamper with the auction result.

4) The collusion of two malicious characters: data sellers and data buyers have conspired to provide false transactions and conduct DOS attacks. In the design of this mechanism, sellers and buyers need to pay the trusted nodes for the task, and the transaction, so it is costly and difficult to cause DoS attacks; the data seller and the trusted nodes conspired to improve the final pricing of big data auction. As mentioned in the security analysis of a single malicious trusted node, the trusted node cannot obtain the explicit data of the buyer bidding or tamper with the result. If the seller sets a high base price, no buyer will successfully auction and the data cannot be sold, so the data seller and the trusted node cannot form collusion; the data buyer and the trusted node conspired to reduce the lowest level of data auction.

V. Conclusion

This paper first analyzes the traditional big data auction and pricing scheme, and summarizes the problems of mechanism fairness, collusion, denial and third-party hidden dangers existing in the traditional data transaction and data auction and pricing process. Based on the above problems, this paper proposes a big data transaction pricing mechanism BDPM. Which introduced the SGX technology, Combining the characteristics of the SGX's remote authentication and data sealing, To aled the bidding data and base price, A new sealed bidding scheme was realized; Designed the anti-communist auction pricing algorithm, Combined with the new sealed bid scheme, Prevents collusion attacks in the auction pricing process, Implement the auction pricing algorithm in the SGX secure zone, Save the untrusted auctioneer, Solve the problem of big data commodity security, fair pricing; In addition, through the features of the blockchain and smart contracts, Introducing a margin mechanism, Ask trading participants to advance deposits in smart contracts, Increase the cost of denial, Prevent the buyer refused to pay, the seller back and other denial problems, Ensure the smooth implementation of data transactions; Then, by introducing a distributed file system, IPFS, Store encrypted big data goods into the IPFS, It avoids the hidden danger of retaining data on unreliable third-party platforms. Finally, through the characteristics of cryptographic signature authentication and hash function, the correctness of data transaction pricing results and the integrity of data goods are guaranteed, and the rights and interests of both buyers and sellers are guaranteed.

References

[1] Wei X , Li X . Anti-collusion data auction mechanism based on smart contract[J]. Information Sciences, 2020, 555(1).

[2] Cai H , Zhu Y , Jie L I , et al. Double Auction for a Data Trading Market with Preferences and Conflicts of Interest.

[3] Chen C , Wu J , Lin H , et al. A Secure and Efficient Blockchain-based Data Trading Approach for Internet of Vehicles[J]. IEEE Transactions on Vehicular Technology, 2019, PP(99):1-1.

[4] An D , Yang Q , Yu W , et al. Towards truthful auction for big data trading[C]// IEEE International Performance Computing & Communications Conference. IEEE Computer Society, 2017:1-7.

[5] Cao X , Chen Y , Liu K . Data Trading With Multiple Owners, Collectors, and Users: An Iterative Auction Mechanism[J]. IEEE Transactions on Signal and Information Processing over Networks, 2017.

[6] Bichler M. An experimental analysis of multi-attribute auctions[J]. Decision support systems, 2000, 29(3): 249-268.

[7] Jap S D. The impact of online reverse auction design on buyer – supplier relationships[J]. Journal of Marketing, 2007, 71(1): 146-159.

[8] Shih D H, Wu T W, Shih M H, et al. A Novel Auction Blockchain System with Price Recommendation and Trusted Execution Environment[J]. Mathematics, 2021, 9(24): 3214.

[9] McKeen F, Alexandrovich I, Berenzon A, et al. Innovative instructions and software model for isolated execution[J]. Hasp@ isca, 2013, 10(1).

[10] Hoekstra M, Lal R, Pappachan P, et al. Using innovative instructions to create trustworthy software solutions[J]. HASP@ ISCA, 2013, 11(10.1145): 2487726-2488370.

[11] Anati I, Gueron S, Johnson S, et al. Innovative technology for CPU based attestation and sealing[C]//Proceedings of the 2nd international workshop on hardware and architectural support for security and privacy. New York, NY, USA: ACM, 2013, 13(7).