Research Article

A Blockchain Consensus Optimization-Based Algorithm for Food Traceability

Peng Liu [ID],[1] **Shuang Ren** [ID],[2] **Jun Wang** [ID],[3] **Shanshan Yuan** [ID],[1] **Yiying Nian** [ID],[1] **and Yifan Li** [ID][2]

[1]*China National Institute of Standardization, Beijing, China*
[2]*Beijing Jiaotong University, Beijing, China*
[3]*China Agricultural University, Beijing, China*

Correspondence should be addressed to Peng Liu; liupeng@cnis.ac.cn

At present, blockchain technology is more and more widely used in the field of food traceability, and good results have been achieved. However, many of the current blockchain technologies and algorithms are not developed for the specific situation of food traceability, resulting in resource waste and low computational efficiency. In view of these problems, this paper analyzes and summarizes the classic distributed consensus mechanism in blockchain technology, focusing on the PBFT (practical Byzantine fault tolerance) consensus mechanism and the existing problems related to the improvement scheme. In order to solve the problem of low efficiency of a consensus algorithm in a food traceability scenario, this paper proposes a blockchain consensus algorithm suitable for the food traceability scenario based on clustering and food credit. In addition, the differences between the improved algorithm and the classical Byzantine consensus algorithm in consensus algorithm time and communication times are analyzed through experiments and simulations. The consensus efficiency of the improved algorithm in this paper is significantly improved, which can greatly reduce the application difficulty of blockchain in food traceability.

1. Introduction

With the continuous development of blockchain technology, the applications based on the blockchain technology have gradually been recognized and utilized by different sectors and provided the adequate technical platform for the application and research of food quality safety traceability [1]. The food circulation scale gradually expands, ensuring that the food quality safety plays a significant role in promoting the efficient, healthy, and green ecological development of foods. The food quality safety involves different links such as the production, processing, transportation, and sales. Among them, the safety hazards may exist at every link. Therefore, the accurate traceability of real information at every link of food is the significant method to guarantee the food quality safety and also conforms to the consumer's expectations on food quality safety. Most of traditional food supply chain traceability systems combine the identification technologies such as bar code technology and radio frequency identification (RFID) and Internet of Things (IoT) technology for traceability and adopt the centralized database to store the traceability information [2]. The traditional traceability system adopts the centralized management, and the food terminal users such as breeding manufacturer, processing factory, logistics company, sales company, and government regulator are separated from each other, which results in disadvantages such as the low transparency of traceability information, low regulatory force of the government, and easy tampering of data. Therefore, the truthfulness and reliability of the data are low, and it is hard to position the problem links at the first time, in case of the occurrence of any food safety accident.

The blockchain is a sort of decentralized distributed ledger system, which is used to register and issue the digital

assets, equity security, and integral and conduct the transfer, payment, and transaction via the point-to-point method. Compared with the traditional centralized ledger system, the blockchain system has advantages such as full disclosure, tamper proof, and multipayment prevention and is independent of any trusted third party. Higher network delay exists under the point-to-point network [3]. Therefore, it is impossible that the sequential order of affairs observed on every link is completely consistent. As a result, a mechanism shall be designed for the blockchain system, identifying the sequential order of affairs occurring at approximately the same time. This algorithm through which a consensus will be reached for the sequential order of affairs within a time window, is called "consensus algorithm" [4]. At present, there are several major categories of consensus algorithms, namely, PoW, Pos, DPos, Pool, and PBFT, and the use of the consensus algorithm is of great significance to guarantee the data consistency [5].

The consensus algorithm of existing blockchain is often aimed at the application scenario of financial property, but it is not very applicable to the features required under the application scenario of food traceability such as high crash fault tolerance, frequent search, and having no unknown roles in traceability [6]. For the safety, scalability, and maturity of every blockchain protocol, the clearly defined and implemented consensus algorithm may be the most significant function, and the selection of the appropriate consensus algorithm is crucial to support the mutual trust in the distributed business network. It adopts the computational complexity theory and statistics principle to analyze, test, and compute the different consensus algorithms such as classic Byzantine fault tolerance and SOLO consensus model, adopts Occam's razor principle to optimize the relevant consensus algorithms, and researches and obtains the optimized consensus algorithm that is suitable for the traceability chain from farm to table, which lays a solid foundation for improving the efficiency of a blockchain-based food traceability system.

2. Present Status for Construction of the Blockchain-Based Food Traceability System

With the popularization and development of digital currency, the blockchain technology enters the public's vision and has been hailed as the fourth milestone in credit history, the cornerstone of future credit [7]. At present, many scholars have proposed the solution that realizes the blockchain-based food supply chain traceability.

Chen and Li [8] used the K-medoids clustering algorithm to cluster and hierarchically divide the large-scale network nodes participating in the blockchain consensus based on features, and then, the improved multicentered PBFT consensus algorithm is applied to the clustered model. Moreover, this paper improved the K-medoids algorithm. The experimental results show that K-PBFT optimizes the consensus process involving large-scale nodes, so that the blockchain technology can be applied to a wider range of application scenarios. Chen and Di [9] used the credit model to improve the PoW algorithm and then proposed the
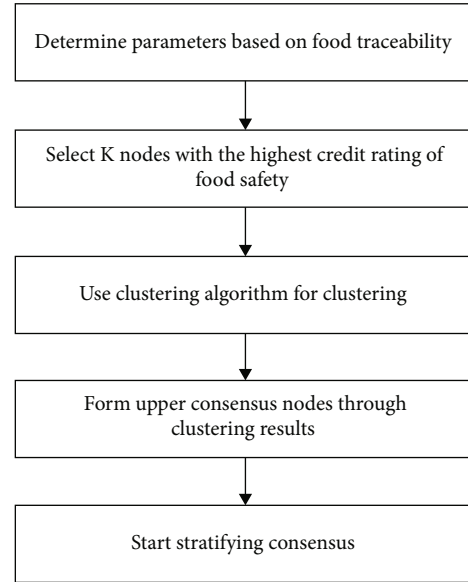

Figure 1: Process of the optimized blockchain consensus algorithm.

TCCM consensus algorithm based on a threshold cryptography scheme and finally added it to the blockchain to construct a new data sharing and traceability scheme. The results show that the randomness of the CPoW algorithm is improved, the calculation amount of TCCM is more than 1 W, and the resource consumption is low, which improves the security of data sharing and realizes data traceability. It shows that the traceability scheme based on blockchain data sharing technology is feasible and can provide new technical ideas for big data sharing and data flow detection. Bo et al. [10] used blockchain, asymmetric encryption, digital signature, privacy computing, and other technologies to build a trusted food safety traceability platform. The blockchain food safety traceability system that meets the needs of government supervision and meets the needs of enterprise privacy protection solves the problem of enterprise data privacy protection. Li and Jiang [11] adopted the blockchain technology of agricultural product circulation in the mode of alliance chain, constructed the framework structure and technical system of Wen's agricultural product traceability platform, and expounded the mechanism of the traceability system to solve the food safety of agricultural products. George [12] presented a prototype of restaurant food quality traceability using blockchain and food quality data indexing. The prototype obtains data from various stakeholders in the food supply chain, analyzes and integrates the data, and applies the Food Quality Index (FQI) algorithm to generate FQI values. Based on certain parameters, FQI values help determine whether a food is fit for consumption. In addition to enhancing the traceability of food, the prototype also helps in grading the quality of food for human consumption. Mondal et al. [13] proposed a blockchain-inspired IoT architecture that uses a proof-of-object-based authentication protocol and a proof-of-work protocol similar to digital currency, by integrating RFID-based technology at the physical layer (RFID) sensors and integrating blockchain at the
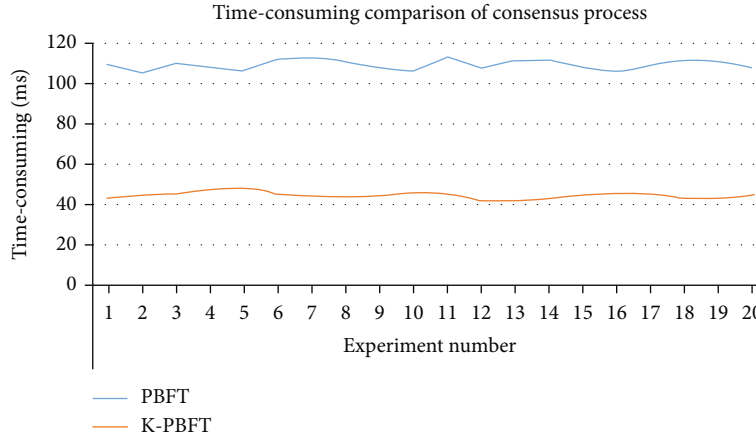
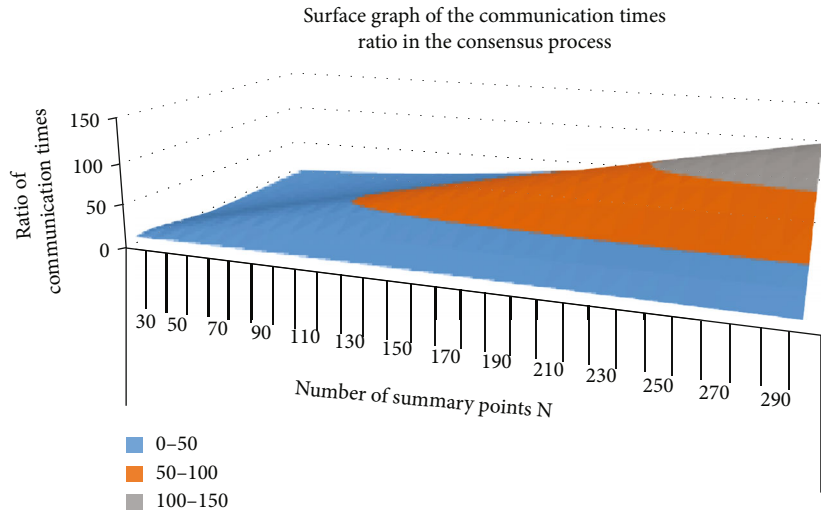FIGURE 2: Time-consuming comparison of the consensus process.

FIGURE 3: Surface graph of the communication time ratio in the consensus process.

network layer to achieve a complete architecture. Liu [14] used blockchain technology and radio frequency identification technology to construct the "blockchain+RFID" two-in-one collaborative sharing model. This model uses the alliance chain to build the food traceability system, and the Go language is used to develop the chain in Fabric Code. Node. It is used to write the client program. The mobile terminal realizes the food traceability query by scanning the code, and the client uses certificate authentication account to query the web page. Through the interconnection of all aspects of food circulation, we can quickly locate responsible parties for food safety incidents. Wang et al. [15] constructed a food safety credit system model based on a consortium chain, which consists of food manufacturers, processors, distributors, operators, consumers, governments, media, banks, and credit agencies to maintain daily traceability data, transaction data, credit data, and other credit information data.

In the blockchain system, no fixed centralized entity exists, the supervision, validation, and decision-making of the system and the generation and maintenance of data are participated by every node in the system, and it is not required to establish the foundation of trust between nodes. Therefore, in light of consensus reaching, the consensus mechanism is always indispensable and urgently needed. In fact, the process in which a consensus is reached via the consensus mechanism is such a process that ensures the validation and update of data by node in the system. Afterwards, a unified consensus result will be provided to the outside.

The consensus mechanism mainly includes two parts, namely, consistency and consensus. Among them, consistency mainly refers to whether the node data in the blockchain are consistent, and the consensus refers to a method of reaching consensus by the node under the premise of consistent data. The consensus algorithm falls into two major categories, namely, Byzantine fault tolerance type and crash fault tolerance type.

Both Paxos [16] and Raft [17] are classified as the crash fault tolerance type consensus mechanism, which are realized by assuming that all nodes are honest nodes, without any malicious nodes. Therefore, these two consensus

mechanisms are not applicable to network environment in which the malicious nodes exist.

PBFT is a classical Byzantine fault tolerance-type consensus mechanism, which solves the Byzantine Generals Problem [18]. In the network, some malicious nodes inevitably exist, and PBFT adopts a method of neglecting the malice of a few nodes by the honesty of most nodes; that is, it may accommodate about 33% malicious nodes. However, in the system where the PBFT is adopted, the number of nodes involved in the consensus is fixed, instead of any dynamic increase or decrease. In case of any change to the number of nodes, the PBFT needs to be restarted. Therefore, PBFT is not applicable to the dynamically changing network environment.

With regard to the problem that PBFT lacks the dynamism, the algorithms such as GBC, dynamic PBFT, and S-PBFT appear [19]. GBC uses three communication methods and unified data structure of the gossip protocol, enabling it to have the dynamism; dynamic PBFT adds the node information record form and NDC protocol to deal with the dynamic change of nodes; S-PBFT adopts the short-term signature and key distribution mechanism to regularly update the keys and remove the dishonest nodes, which reduces the time delay and improves the scalability [20, 21].

In light of the network congestion and low consensus efficiency problems of PBFT resulting from the increase in nodes, the K-PBFT algorithm also appears [22]. K-PBFT fully utilizes the improved K-medoids clustering algorithm to classify and layer the nodes in the system. At the same time, it also completes the consensus in combination with the PBFT's "three-stage" consensus process and further improves the consensus efficiency.

3. Research Method

When selecting the sample feature space coordinate system of the clustering algorithm to compute the Euclidean distance, it is possible to consider selecting the feature spaces such as network delay among nodes and geospatial location, enabling the clustering algorithm to better classify the nodes with higher similarity into a cluster; alternatively, combine many features to form the high-dimensional feature space coordinate system, enabling the cluster center nodes selected by the clustering algorithm to be better nodes evaluated under many feature spaces such as intracluster hardware environment and network environment, ensuring that various feature differences among slave nodes of the same cluster and among center nodes of various clusters are minor and further obtaining higher consensus efficiency. Finally, determine the following node dimensions:

D1: food safety credit grade: it shall be determined by food safety regulatory authorities (it falls into four grades, namely, A, B, C, and D).

D2: product category: according to food and food additive class label in the catalogue of food production licenses, multiple choices are allowed.

D4: food safety risk grade: it shall be determined by food safety regulatory authorities (it falls into four grades, namely, A, B, C, and D).

D5: type of industrial chain for the node: plantation, cultivation, production of raw materials, production of finished products, storage, transportation, and sales.

D6: geospatial position of the node (according to three levels, namely, province, city, and county).

During the process of iteration, the K-medoids algorithm needs to repetitively compute the distance from each node to other nodes within the cluster, which increases the computing complexity of the algorithm. In combination with the application scenario of blockchain, the number of node clusters involved in the blockchain consensus is unstable, the hardware conditions vary, and the operating environmental conditions are uncertain. Therefore, under certain scenarios, after the top supervisors select the center nodes of initial cluster, it is not expected to waste the computing resources of the consensus cluster to frequently replace the intracluster center. According to the application features of blockchain, the Kmedoids clustering algorithm will be improved hereunder.

Assume that n nos. of p-dimension feature object nodes are used as the clustering dataset. $v_i = \{v_{i1}, v_{i2}, v_{i3}, \cdots, v_{ip}\}$ ($i = \{1, 2, \cdots, n\}$) is used to indicate the coordinate value when the number i node is mapped in the p-dimension feature space R^P. All the above node datasets $V = \{v_1, v_2, \cdots, v_n\}$ may be expressed as follows:

$$V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1p} \\ v_{21} & v_{22} & \vdots & v_{2p} \\ \vdots & \vdots & & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{np} \end{bmatrix}. \tag{1}$$

The K-medoids algorithm is the function that uses the Euclidean distance to express the similarity of two nodes and uses equation (2) to express the distance between two nodes, namely, v_i and v_k.

$$d(v_i, v_k) = \|v_i - v_k\| = \sqrt{\sum_{j=1}^{p} (v_{ij} - v_{kj})^2}. \tag{2}$$

The K-medoids algorithm is applied in the blockchain environment, and the initial K nos. of center nodes are not randomly selected from all samples. Instead, K nos. of nodes with the highest food safety credit grade are directly selected as the initial clustering center nodes. In this case, we do not expect to adjust the center nodes again. The improved K-medoids algorithm may better control the replacement probability of cluster center nodes, and the detailed procedure is shown below.

Step 1. Directly select K nos. of nodes with the highest food safety credit grade in the dataset V as the initial cluster center nodes.

Step 2. Compute the Euclidean distance from each node to K nos. of cluster centers according to equation (2).

Step 3. Complete the cluster classification.

Different from the traditional K-medoids algorithm, the improved K-medoids algorithm directly adopts the initial node as the clustering center, which will not further carry out the clustering iteration. Thus, it is possible to better control the computing power resources consumed when the node runs the clustering algorithm, facilitate selecting the node set of the backbone consensus cluster in the blockchain environment, and improve the controllability of the blockchain platform.

The K-medoids algorithm is combined with PBFT, which utilizes the clustering algorithm to indiscriminately classify the clustering and conduct the layering for all consensus nodes in the original PBFT, adopts the cluster center nodes after the clustering as the master node of all nodes in this cluster, and adopts all nonclustering center nodes in every cluster as the slave nodes in this cluster. This paper calls every cluster as a consensus subcluster, the nonclustering center nodes form the K nos. of consensus subclusters led by K nos. of master nodes, and then the collection of all K nos. of master nodes forms a backbone consensus cluster.

The main process stages of the K-PBFT consensus algorithm are shown below.

(1) Consensus stage of the consensus subcluster: every node will send the request to the master node of its cluster (cluster center node). After the master node receives the request within a period, it packages several requests into a block and then broadcasts this block to its consensus subcluster for a PBFT consensus

(2) Backbone cluster consensus stage: after the block passes the consensus validation process of the consensus subcluster, the consensus confirmation of the secondary PBFT will be conducted in the backbone consensus cluster. K nodes in the backbone consensus cluster take turns to broadcast through the polling method the block that has passed the consensus validation of its consensus subcluster for consensus

(3) Submission stage: after the block passes the consensus of the backbone consensus cluster, all master nodes will carry out the digital signature for this block and receive the digital signature from other master nodes, which represents the recognition of the truth and effectiveness of this block; then, the digital signature collection of this backbone consensus cluster together with the block itself will be packaged into all slave nodes in its consensus subcluster to which the message is submitted and broadcasted, which means that the on-chain operation may be made for this block. Until now, all nodes of any consensus subcluster will receive the submitted message of this block

(4) Execution stage: after the slave node receives the submitted message from the master node, it will validate the digital signature collection accompanying the block and may judge whether this block has passed the consensus validation of backbone consensus cluster. If the validation fails, it may be considered that the master node of this slave node has the malicious behavior, and this illegal operation may be reported to the blockchain supervisor, realizing the supervision upwards from the node; if the validation succeeds, it may execute the request contents of this block and realize the on-chain of block records

So the process of the optimized blockchain consensus algorithm is shown in Figure 1.

The optimized blockchain consensus algorithm better classifies the nodes with higher similarity into a cluster; alternatively, combine many features to form the high-dimensional feature space coordinate system, enabling the cluster center nodes selected by the clustering algorithm to be better nodes evaluated under many feature spaces such as product category and food safety risk grade, ensuring that various feature differences among slave nodes of the same cluster and among center nodes of various clusters are minor and further obtaining higher consensus efficiency.

4. Experiment and Results

This paper will conduct the contrast experiment for the optimized blockchain consensus algorithm after the improvement and traditional PBFT consensus algorithm, respectively, in terms of two aspects, namely, single consensus time-consuming and communication times in the single consensus process.

4.1. Consensus Time-Consuming Experiment. In the experiment, without considering factors such as the processing time-consuming after the single node receives the information, the resource utilization on the single computer, CPU processing speed, disk read/write speed, and network congestion and assuming that the receiving and sending of all messages on the single computer at the same time will occur in parallel, it only computes the time-consuming from the start of the consensus process to the end of the consensus process. In order to compare the consensus time-consuming of two models, 20 independent contrast experiments are made (neglect the polling process of the backbone consensus cluster). In every experiment, randomly generate the delays between nodes according to the above simulation item and then record the time-consuming for ms consensus with the optimized PBFT consensus algorithm and ordinary PBFT consensus algorithm. The contrast results are shown in Figure 2.

The average time-consuming computed for the PBFT single consensus through 20 times of repeated contrast experiment is 109.25 ms, and the time-consuming for the optimized PBFT single consensus is 44.25 ms. As shown by the experiment results, the time-consuming for the single consensus with the optimized PBFT algorithm is shortened by 60%, compared with that with the PBFT algorithm.

As known from the above experiment, every consensus of the traditional PBFT algorithm needs a three-stage

consensus process conducted by all consensus nodes within the global scope. As a result, a great number of network nodes will act as the consensus nodes and participate in the PBFT consensus process, which will easily cause the network congestion due to the frequent broadcast of a consensus message among nodes. However, the optimized PBFT algorithm that is improved with the K-medoids clustering algorithm will adopt food safety credit grade, product category, food safety risk grade.etc as the evaluation features with the node clustering method, carry out the screening, layering and clustering on the consensus nodes, and then conduct the multicentralized consensus on a small scale in the better intercluster and intracluster network environment. As shown from the experimental results, if there are a great number of consensus nodes, the optimized PBFT algorithm that adopts the network delay between consensus nodes as the feature space effectively shortens the time-consuming of the consensus process and improves the consensus efficiency, compared with the PBFT.

Meanwhile, due to the layered and multicentralized structure of optimized PBFT, the partial consensus process may be executed in parallel between single consensus subclusters and between the consensus subcluster and backbone consensus cluster. Therefore, in the real network environment, the optimized PBFT algorithm may process the requests raised by multiple consensus subclusters in parallel.

4.2. Contrast of Communication Times. In order to validate whether the improved consensus algorithm model is better than the traditional PBFT consensus algorithm in light of communication times, this paper sets up the contrast experiment for communication times of the consensus process as shown in Figure 3.

Through the above experimental analysis, the optimized blockchain consensus algorithm may effectively reduce the communication times of the single consensus process. The optimized blockchain consensus algorithm conducts the consensus after the node clustering and layering at a small scale, which effectively reduces the communication times required for the single consensus. Meanwhile, even when the number of consensus nodes on the food tracing platform is huge, it is not required to maintain network connection between every consensus node and all other nodes. Instead, the network connection is only required for every consensus node with the nodes in the same consensus subcluster, which reduces the resources consumed for establishing and maintaining the network connection on a single computer.

5. Conclusion

This paper puts forward an improved PBFT blockchain model. The optimization algorithm reselects the feature vector according to the particularity of the food traceability scene and redefines the K-medoids selection basis for selecting K-medoids clustering, which greatly improves the calculation speed and efficiency of the consensus algorithm. The experimental results of consensus time and consensus communication times show that the improved k-PBFT algorithm is superior to the traditional PBFT algorithm in terms of

consensus efficiency and consensus process communication times in the food traceability scenario. The improved algorithm will greatly reduce the technical difficulty of the food traceability system using blockchain and effectively promote the application of blockchain technology in the food traceability system to ensure food safety. The improved algorithm only can be used in the food traceability system.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J. Kim, "Blockchain technology and its applications: case studies," *Journal of System and Management Sciences*, vol. 10, no. 1, pp. 83–93, 2020.

[2] Y. Chai and Q. Li, "Research on influencing factors of knowledge sharing in supply chain enterprises under blockchain environment," *Tehnički vjesnik*, vol. 28, no. 5, pp. 1553–1559, 2021.

[3] Y. Wang, "Analysis on the application of blockchain technology in the field of logistics and supply chain," *Logistics Engineering and Management*, vol. 42, no. 10, 2020.

[4] S. Y. Lee, "A fast healthcare interoperability resources based blockchain framework for medical data management," *Journal of System and Management Sciences*, vol. 10, no. 4, pp. 125–136, 2020.

[5] F. U. Xiang, W. A. N. G. Huaimin, and S. H. I. Peichang, "A survey of blockchain consensus algorithms: mechanism, design and applications," *Science China (Information Sciences)*, vol. 64, no. 2, pp. 100–114, 2021.

[6] F. M. Tseng, E. I. Gil, and L. Y. Lu, "Developmental trajectories of blockchain research and its major subfields," *Technology in Society*, vol. 66, article 101606, 2021.

[7] M. Swan, *Blockchain: blueprint for a new economy*, O'Reilly Media, Inc., 2015.

[8] Z. Chen and Q. Li, "Improved PBFT consensus mechanism based on K-medoids," *Computer Science*, vol. 46, no. 12, pp. 101–107, 2019.

[9] C. Liu and D. Zhao, "Traceability consensus mechanism based on blockchain data sharing technology," *Information Technology*, vol. 12, pp. 112–117, 2021.

[10] B. Yang, Y. Wang, and Y. Liu, "Construction and implementation of asymmetric privacy protection food safety traceability system based on blockchain," *China Auto-ID*, vol. 6, pp. 41–47, 2021.

[11] Y. Li and J. Jiang, "Construction of agricultural product traceability platform based on blockchain technology," *Grain*

Science and Technology and Economy, vol. 46, no. 6, pp. 49–53, 2021.

[12] R. V. George, H. O. Harsh, P. Ray, and A. K. Babu, "Food quality traceability prototype for restaurants using blockchain and food quality data index," *Journal of Cleaner Production*, vol. 240, article 118021, 2019.

[13] S. Mondal, K. P. Wijewardena, S. Karuppuswami, N. Kriti, D. Kumar, and P. Chahal, "Blockchain inspired RFID-based information architecture for food supply chain," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5803–5813, 2019.

[14] Z. Liu, "A food traceability framework based on permissioned blockchain," *Food & Machinery*, vol. 2, no. 2, pp. 107–113, 2020.

[15] J. Wang, S. Zeng, J. Guo, and G. Liu, "The application of blockchain technology in the construction of food safety credit system," *Wireless Internet Technology*, vol. 16, no. 21, pp. 137–140, 2019.

[16] L. L. Paxos, "Made simple," *ACM SIGACT News*, vol. 32, no. 4, pp. 18–25, 2001.

[17] D. Ongaro and J. Ousterhout, "In search of an understandable consensus algorithm," in *Proceedings of the ATC14,USENIX Annual Technical Conference*, pp. 305–319, Philadelphia, PA, USA, 2014.

[18] M. Castro and B. Liskov, "Practical Byzantine fault tolerance," in *Proceeding of the 1999 Third Symposium on Operating System Design and Implementation*, pp. 173–186, Berkeley, CA, USA, 1999.

[19] S.-j. Zhang, J. Chai, Z.-h. Chen, and H.-w. He, "Byzantine consensus algorithm based on gossip protocol," *Computer Science*, vol. 45, no. 2, pp. 20–24, 2018.

[20] X. Hao, L. Yu, L. Zhiqiang, L. Zhen, and G. Dawu, "Dynamic practical Byzantine fault tolerance," in *2018 IEEE Conference on Communications and Network Security (CNS)*, pp. 1–8, Beijing, China, 2018.

[21] X. Fan, "Scalable practical Byzantine fault tolerance with short-lived signature schemes," in *Proceedings of 28th Annual International Conference on Computer Science and Software Engineering (CASCON)*, p. 12, New York, NY, USA, 2018.

[22] Y.-h. Zhou, J.-b. Fang, Y.-x. Jia, L.-y. Jia, and W.-m Shi, "Consortium Blockchain Consensus Algorithm Based on PBFT," *Computer Science*, vol. 48, no. 11, pp. 1331–2141, 2021.