International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 8 Issue 6, Nov-Dec 2024 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Research on Agricultural Product Traceability Solutions Based on Blockchain and Machine Learning

Liu Tongjuan[1], Cai Xiaodong[2]

[1]School of Information, Beijing Wuzi University, Beijing, China
[2]School of Postgraduate, Beijing Wuzi University, Beijing, China

ABSTRACT

In an information-driven society, food safety has become a significant concern in people's daily lives. Establishing an agricultural product traceability system plays a crucial role in ensuring food safety. However, traditional agricultural product traceability systems encounter numerous issues in practical application, such as information asymmetry, centralized system structures, and vulnerability to information tampering. Therefore, this research aims to leverage blockchain and Internet of Things (IoT) technologies to address these problems in the existing agricultural product traceability systems, enhancing their efficiency and reliability, while also employing machine learning for supply and demand prediction within the supply chain. The study found that the decentralized, tamper-proof, and transparent nature of blockchain technology, combined with the extensive data collection and processing capabilities of IoT, can effectively improve agricultural product traceability systems. Additionally, machine learning can accurately predict demand, thereby improving the efficiency of supply chain transportation to avoid shortages and overstocking.

KEYWORDS: *Blockchain; Internet of Things; Traceability; Hyperledger structure; Agricultural products; Machine learning*

How to cite this paper*:* Liu Tongjuan | Cai Xiaodong "Research on Agricultural Product Traceability Solutions Based on Blockchain and Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-8 | Issue-6, December 2024, pp.288-295, URL: www.ijtsrd.com/papers/ijtsrd70532.pdf

IJTSRD70532

1. INTRODUCTION

In today's society, food and drug safety are directly related to human health and safety. With the rapid development of the digital economy, the concept of supply chain visibility for food and drug safety has received more attention than ever before.

Automatic identification technologies, such as barcodes, QR codes, radio frequency identification (RFID), and Internet of Things (IoT) data collection and processing technologies, can record and process various types of information regarding product visibility throughout the supply chain, allowing products to be tracked and traced [1]. Traditional product traceability systems are typically based on centralized data storage architectures, where traceability information is usually stored and controlled by third-party agencies. In such systems, it is difficult to ensure data transparency and integrity, and the system may suffer from single points of failure, information tampering, and lack of credibility [2].

Blockchain is a tamper-proof, distributed, decentralized peer-to-peer technology that can be used to track and verify digital transactions. It has many new features, such as distributed data storage, smart contracts, consensus mechanisms, and more [3]. Because blockchain is decentralized, tamper-proof, transparent, and auditable, it can provide a secure environment for capturing data within the supply chain, especially event data created using wired or wireless sensors [4].

Currently, Fabric is the most widely used and well-known permissioned blockchain framework. A recent survey by Rauchs et al. showed that according to data from the Cambridge Centre for Alternative Finance, 48% of permissioned blockchain projects are built on Fabric [5]. IoT, through the use of technologies such as RFID, infrared sensors, and global positioning systems (GPS), and by following specific protocols, connects various items to the internet for information exchange and communication.

The combination of blockchain and IoT can also enhance the transparency and efficiency of the supply chain. Many scholars have proposed various studies on blockchain architectures suitable for IoT applications.

Novo implemented a scalable, easy-to-manage IoT distributed access control system architecture based on blockchain, which stores and distributes access control information via blockchain, effectively addressing the access scalability issues of billions of constrained devices in IoT and improving processing load capacity [6]. Li et al. proposed a multi-layer secure IoT model, which divides IoT into multi-level decentralized networks and adopts blockchain technology at each level, simplifying the actual deployment of blockchain while maintaining its high security [7]. Liu et al. developed a lightweight blockchain architecture suitable for power-constrained industrial IoT environments, effectively balancing resource consumption and device performance limitations [8]. Mondal et al. constructed an IoT architecture for the food supply chain based on a proof-of-object certification protocol, combined with RFID sensors and blockchain technology, to record data in real-time at the physical and network layers, providing an immutable digital history that allows consumers to view the public ledger [9]. The authors of [10] provided a framework for supply chain demand forecasting using machine learning.

Based on the current state of research, the existing traceability systems have the following shortcomings:

Firstly, since blockchain typically stores a complete set of data on each node, uploading all traceability data to the blockchain platform may lead to efficiency issues or even data explosion, as the traceability data in the supply chain can be very large. Secondly, in traditional agricultural product traceability systems, regulatory mechanisms are usually manually executed by regulatory authorities, resulting in a lack of fairness, openness, and transparency in the regulatory process. Some irresponsible individuals may seek personal gain and fail to perform regulatory tasks according to regulations, thus affecting the credibility of traceability data. Thirdly, the high uncertainty of supply and demand is a characteristic of the supply chain, and production and procurement strategies based on inaccurate demand can lead to profit loss and additional costs.

To address these issues, this paper proposes a traceability system based on Hyperledger Fabric blockchain technology and front-end display technology. By combining the analytical capabilities of machine learning and the information collection capabilities of IoT, a more accurate demand forecasting model can be created. By integrating machine learning algorithms into the blockchain framework, enterprises can improve model accuracy and quickly respond to changing market conditions. Utilizing these innovations, enterprises can enhance performance, streamline processes, and respond more effectively to customer needs. Through layered storage and off-chain data management technologies, the load on the blockchain network is reduced, avoiding data explosion issues and ensuring the efficient operation of the system. With smart contracts, automated regulatory processes can be realized, ensuring fairness, openness, and transparency of traceability data, preventing human intervention and data falsification. By using machine learning algorithms to analyze historical and real-time collected data, accurate market demand forecasts can be made, helping enterprises to formulate more effective production and procurement strategies, reduce waste and costs, and increase profits.

Through the above methods, the system proposed in this paper not only improves the efficiency and transparency of agricultural product traceability but also enhances the overall responsiveness of the supply chain, ensuring that enterprises maintain a competitive edge in the market.

2. System Architecture

In this paper, we establish a trustworthy product traceability system based on Fabric and the Internet of Things (IoT).

Currently, blockchain storage mainly uses single-chain structures or "main chain + side chain" architectures. However, both of these architectures store traceability data in the agricultural product supply chain in the form of the main chain. As the usage time increases, the amount of traceability data will grow, occupying more blockchain storage capacity and reducing system performance. Therefore, this system adopts a method of storing data both on-chain and off-chain to enhance system flexibility and alleviate the problem of data explosion.

Additionally, current blockchain-based agricultural product traceability systems lack flexibility and scalability. Traditional blockchains use a sequential execution method, requiring all nodes to execute smart contracts in order, which limits the system's scalability and performance. Therefore, this system uses a modular design method for the blockchain system, isolating various business processes and adhering to the principles of high cohesion and low coupling, thereby improving system flexibility and scalability. By comparing the characteristics of public chains, consortium chains, and private chains, this system uses the consortium chain Hyperledger Fabric to implement the blockchain. The on-chain + off-chain management model is shown in Figure 1.

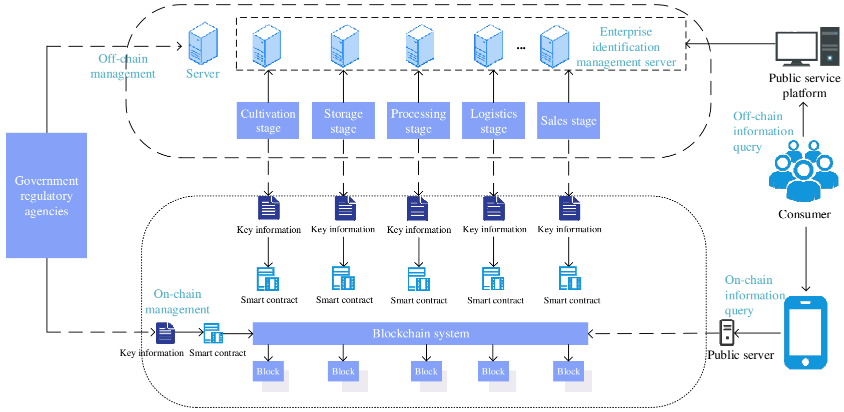


Figure 1: On-Chain and Off-Chain Management Model

3. Implementation

3.1. Database Module

In the agricultural product traceability system within the blockchain and IoT environment, all quality and safety-related data at each stage is stored on the blockchain, while non-quality and safety-related data is stored in a centralized database. This blockchain and IoT-based agricultural product traceability system is a management platform aimed at multiple nodes. To support the extensive data processing requirements, we choose to use MySQL, a relational database. MySQL is a commonly used enterprise-level database, known for its relative stability and ease of maintenance. The database ER diagram is shown in Figure 2.

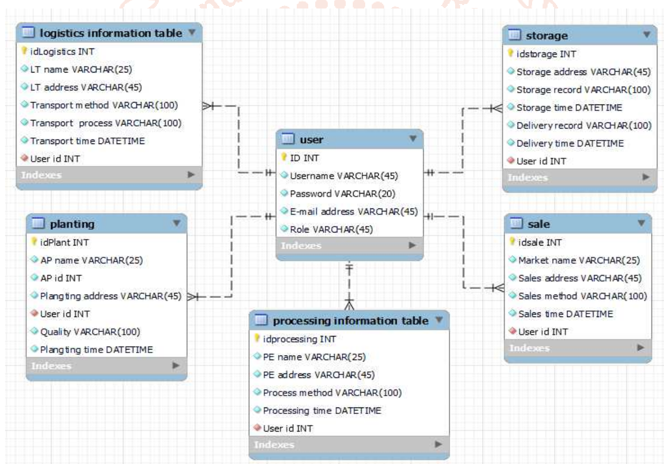


Figure 2: System Off-Chain Database ER Diagram

3.2. Blockchain Creation

1. **Block Implementation**: First, a **Block** class is created in the program, which represents the properties and methods of an agricultural product block. Then, a **Blockchain** class is created, which initializes a list in its constructor to store the blockchain of agricultural products. The code for the agricultural product block is shown in Figure 3.

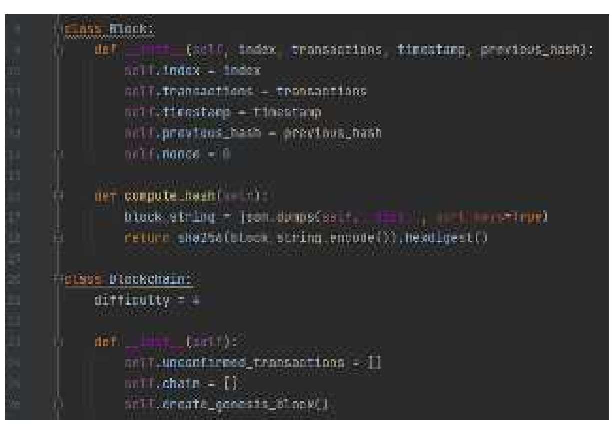


Figure 3: Agricultural Product Block Code Diagram

2. **Blockchain Implementation**: After instantiating the **Blockchain**, the first block, known as the genesis block, needs to be created and appended to the blockchain. The genesis block has an index of 0, a **previous_hash** of 0, and a valid hash. The code for creating the genesis block is shown in Figure 4.

```
def create_genesis_block(self):
    genesis_block = Block(0, [], time.time(), "0")
    genesis_block.hash = genesis_block.compute_hash()
    self.chain.append(genesis_block)

@property
def last_block(self):
    return self.chain[-1]
```

Figure 4: Creation of the Genesis Block

3. **Block Addition and Validation:** The process of adding a new block to the chain requires validation. The validation steps include checking if the proof of work is valid and if the previous_hash referenced in the block matches the hash of the latest block in the chain. The code for this process is shown in Figure 5.

```
def add_block(self, block, proof):
    previous_hash = self.last_block.hash

    if previous_hash != block.previous_hash:
        return False
    if not self.is_valid_proof(block, proof):
        return False

    block.hash = proof
    self.chain.append(block)
    return True
```

Figure 5: Block Addition

3.3. Web Services and Client Development Module

1. Node Database Implementation

The agricultural product traceability system designed in this paper involves a large amount of information, making the design of the node database crucial. In this project, the MySQL node database was designed based on the following principles:

1. Reduce Redundant Data: Due to the frequent operations of adding, deleting, querying, and modifying data, redundant data storage should be minimized to improve database performance and efficiency.

2. Ensure Completeness and Accuracy of Traceability Information: The completeness of traceability information means covering all stages of the agricultural product supply chain, from planting, processing, storage, and transportation to sales, with all key steps needing to be recorded and preserved. Accuracy means the information stored in the node database must be true and accurate.

During the database design process, the actual application requirements and business processes should be combined to define appropriate table structures, set primary keys and foreign keys, and establish indexes. By using Python to call the database, operations such as adding, deleting, querying, and modifying data can be implemented, thereby realizing the traceability function of the agricultural product supply chain.

The establishment of the node database needs to comprehensively consider factors such as data scale, performance requirements, and security to ensure the system can run efficiently and protect the data's security. Figure 6 shows the design diagram of the node database, illustrating the tables and relationships within the database.

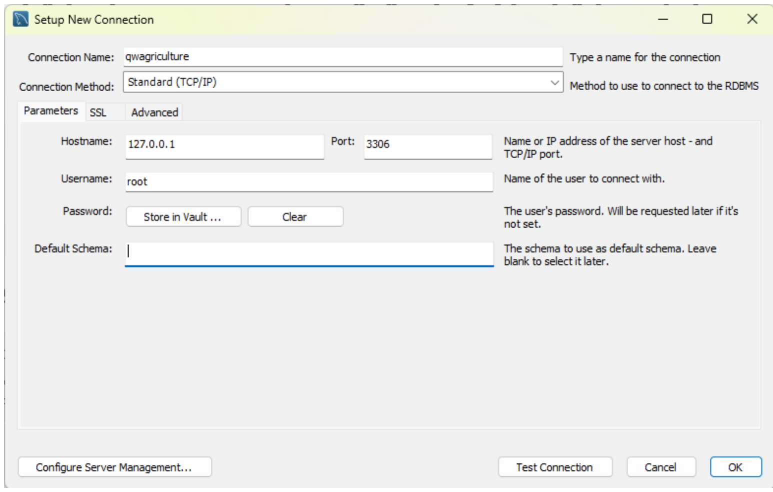


Figure 6: Establishing the Node Database

2. System Function Implementation

User Registration and Login Function Implementation: Users at various nodes in the agricultural product supply chain can register and obtain login accounts through the system's registration interface. The key code snippet for the registration interface is shown in Figure 7.

Figure 7: Key Code Snippet for Registration Interface

In summary, the agricultural product traceability system designed in this paper implements various functionalities using the Python programming language. These functionalities include user registration and login, agricultural product planting, warehousing, processing, logistics, sales, and traceability. By applying blockchain technology, the system achieves the upload, storage, and traceability of data at each stage of the agricultural product supply chain, thereby enhancing the efficiency and reliability of agricultural product quality traceability. Additionally, the use of tools such as PyCharm facilitates the design of the software interface and the implementation of programming logic, making the system more user-friendly and efficient.

3.4. Chaincode

Smart contracts are essentially computer programs deployed on a shared distributed database that are driven by time. They can automatically execute business logic when certain trigger conditions are met, and the behavior of nodes can be transmitted and confirmed in an information-based manner. In Hyperledger Fabric, smart contracts are referred to as chaincode, significantly reducing human involvement and ensuring data integrity.

The smart contracts in the traceability system designed in this paper mainly achieve two functions :

Fist, Automatically storing and verifying key quality information collected by sensor devices during operation and sent to the network service layer.

Second, Invoking corresponding chaincode based on different user requests and interacting with Peer nodes in the blockchain network via system chaincode to handle system configuration, endorsement, and verification.

Based on the analysis of these two functions of smart contracts, the smart contract model shown in Figure 8 was designed. When storing key quality information during the manufacturing process, the system smart contract (chaincode) calls the key information storage smart contract to process it. Hyperledger Fabric's consensus process then stores this information in all manufacturing nodes in the blockchain network, completing the ledger entry and achieving information on-chain. When a user submits traceability business information, the system smart contract calls the traceability business-related smart contract. During the endorsement policy execution at Peer nodes, the chaincode completes the traceability business by using methods such as Install, Instantiate, and Invoke to determine the business logic.

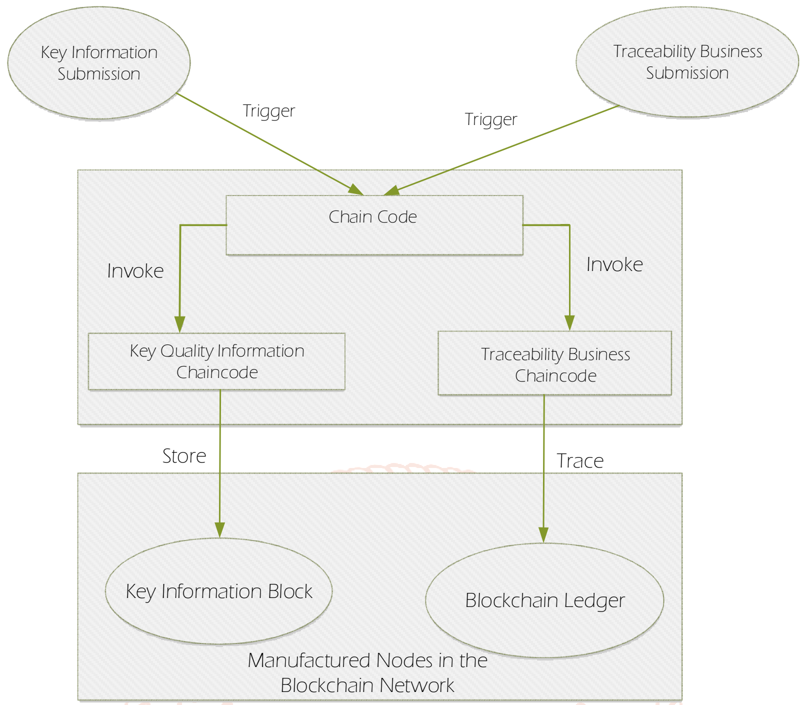


Figure 8: Traceability Smart Contract Model

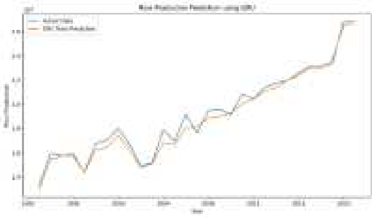
3.5. Demand Forecasting

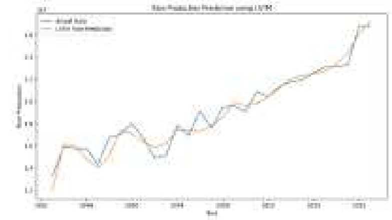
To enhance supply chain efficiency in business operations, we combine machine learning and blockchain technology to predict market supply and demand, helping enterprise managers better manage and operate their projects. Blockchain and IoT can collect and ensure the authenticity of information, while the use of machine learning models allows for data analysis and provides useful recommendations to enterprises based on the results.

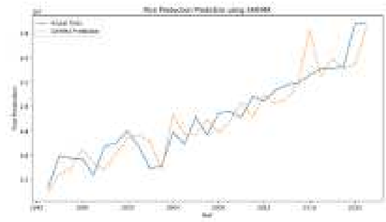
In this system, IoT collects information, and the blockchain ensures the information's authenticity and immutability. By leveraging historical supply information from the supply chain system, we train time series models to predict future demand, thereby improving supply chain efficiency. This helps avoid inventory overstock and supply-demand shortages. Among many machine learning models, we selected GRU, LSTM, and SARIMA models and used the same dataset for prediction. The results show that GRU and LSTM exhibited better prediction performance, while SARIMA's prediction results had significant differences from the actual data.

Dataset: We used the total rice production in Australia from 1961 to 2021 as the dataset for demand. We designed GRU, LSTM, and SARIMA models to predict the future demand for production volumes. The following figure shows the trend comparison between the predictions of these three models and the actual values.

Through this system that combines IoT, blockchain, and machine learning technologies, enterprises can more accurately forecast demand, improve supply chain management efficiency, ensure reasonable allocation of production and inventory, and ultimately enhance business operation efficiency and market competitiveness.







In the figure, the GRU model demonstrates a trend that is almost identical to the actual production, indicating that the GRU model has good accuracy in handling time series data. In comparison, although the LSTM model also shows decent predictive performance, it is slightly inferior to the GRU model in some details. The SARIMA model, on the other hand, shows less satisfactory predictive performance, particularly at some critical points where it fails to accurately reflect the actual situation. This indicates that for complex nonlinear data, GRU and LSTM models perform significantly better than traditional time series models.

If the GRU model is applied throughout the entire supply chain system and the remaining percentage of the forecast values is kept within 5%-10%, it can effectively achieve a balance between supply and demand, avoiding both surplus and shortages. This precise forecasting capability is of great significance for optimizing supply chain management, as it can significantly improve the system's efficiency and responsiveness.

However, due to data limitations, we can currently only predict supply and demand relationships using annual grain production data from individual countries. An ideal supply chain system requires not only macro-level data but also more granular data, such as regional or even town-level data. Such detailed data can extend the supply chain system to every corner, achieving point-to-point supply and demand relationships, thus meeting not only macro-level needs but also the needs of each region.

Fine-grained grain demand forecasting is of significant reference value for a country's production and distribution. High-precision data can significantly improve the accuracy of demand forecasting and help further address imbalances in supply and demand between regions. By optimizing production and distribution strategies, scheduling costs and time can be reduced, thereby enhancing the overall efficiency of the supply chain.

4. Conclusion

This study successfully implemented the design of an agricultural product information traceability system in a blockchain and IoT environment based on Hyperledger Fabric blockchain technology and frontend display technology. This system not only enables product traceability but also authenticates and authorizes IoT devices used for data collection. The system features product information traceability, data tamper-proofing, data confidentiality, and reliability.

The design allows users at various nodes in the agricultural product supply chain to easily register and log in, and to input and query information on agricultural product planting, warehousing, processing, logistics, and sales. The system employs a combined on-chain and off-chain storage approach, effectively mitigating the problem of blockchain data explosion, ensuring system performance and stability. By writing on-chain smart contracts, the regulatory mechanism is strictly managed, ensuring the security of information and the credibility of traceability data. The system can also flexibly respond to different user needs, improving efficiency and user experience. The automatic execution of smart contracts reduces human involvement and ensures data tamper-proofing, enhancing system security. The use of Hyperledger Fabric as a consortium chain enables modular design, endowing the system with high cohesion and low coupling, further enhancing its flexibility and scalability.

The application of the GRU model for predicting grain production and demand helps in precisely controlling the supply-demand balance. This is significant for optimizing supply chain management, improving production and distribution efficiency, reducing waste, and lowering costs. As data acquisition and processing technologies continue to advance, we have reason to believe that future supply chain systems will become more intelligent and efficient, better meeting the needs at various levels. This will not only address macro-level production needs but also resolve supply-demand imbalances between regions.

We will continue to explore the integration of blockchain with IoT and machine learning technologies, promoting the development of the agricultural product traceability field. Our goal is to achieve full traceability and security of agricultural products, as well as balance supply and demand more effectively.

Acknowledgment
This work was supported by Beijing Social Science Fund Project (18GLC066).

References

[1] Thomas Kelepouris, Katerina Pramatari, Georgios Doukidis. Rfid-enabled traceability in the food supply chain[J].Industrial Management & Data Systems, 2007, 107(2):183-200.

[2] Myo Min Aung, Yoon Seok Chang. Traceability in a food supply chain: safety and quality perspectives[J].Food Control, 2014, 39:172-184

[3] PELEKOUDASOIKONOMOU F, ZACHOS G, PAPAIOANNOU Maria, et al. Blockchain-based security mechanisms for iomt edge networks in iomt-based healthcare monitoring systems[J].Sensors, 2022, 22(7):2449-2449.

[4] Yanling Chang, Eleftherios Iakovou, Weidong Shi. Blockchain in global supply chains and cross border trade: a critical synthesis of the state-of-the-art, challenges and opportunities.[J]. CoRR, 2019.

[5] Rauchs, M., Blandin, A., Bear, K., & McKeon, S.B. (2019). 2nd Global Enterprise Blockchain Benchmarking Study. Banking & Insurance eJournal.

[6] Oscar N. Blockchain Meets IoT: An Architecture for Scalable Access Management in IoT[J]. IEEE Internet of Things Journal, 2018, 5(2).

[7] LI C, ZHANG L J.A blockchain based new secure multi-layer network model for Internet of things[C]//2017 IEEE International Congress on Internet of Things, 2017: 33-41.

[8] LIU Y, WANG K, LIN Y, et al. Light chain: a lightweight blockchain system for industrial Internet of things[J].IEEE Transactions on Industrial Informatics, 2019, 15 (6): 3571-3581.

[9] MONDAL S, WIJEWARDENA K P, KARUPPUSWAMIS, et al. Blockchain inspired RFID- based information architecture for food supply chain[J].IEEE Internet of Things Journal, 2019, 6 (3): 5803-5813.

[10] Shamim R, Bentalha B. Blockchain-Enabled Machine Learning Framework for Demand Forecasting in Supply Chain Management[M]// Integrating Intelligence and Sustainability in Supply Chains. IGI Global, 2023: 28-48.