Design and Implementation of Blockchain-based Commodity Anti-counterfeit Traceability System

Xiangyin Meng[1,a*], Bingbing Liu[1,b], Miao Zhang[1,c]

mengxiangyin@nuaa.edu.cn[a], liubb0602@nuaa.edu.cn[b], miaozhang@nuaa.edu.cn[c]

Nanjing University of Aeronautics and Astronautics, Nanjing, China[1]

Abstract. With the booming development of big data technology, people pay more and more attention to the importance of data, and various big data analysis technologies have emerged; blockchain is widely used in anti-counterfeit traceability of commodities with its unique advantages, such as decentralization and non-tamper ability. Due to the popularity of anti-counterfeit traceability systems, more and more commodities support the function of checking authenticity; however, most of the current commodity checking systems do not make use of the data of historical checking, thus leading to unsatisfactory anti-counterfeit effects. In this paper, big data and blockchain technology are applied to the anti-counterfeiting and traceability system to give a unique code to each commodity and generate a corresponding QR code; when consumers scan the code to check, the system uses big data technology to calculate all the historical check records of the commodity and derive the reliability of this check. By setting the reliability threshold, we can achieve the purpose of anti-counterfeiting of commodities.

Keywords: big data; blockchain; anti-counterfeit traceability; QR code; checking reliability

1. Introduction

In recent years, with the advancement of technology, e-commerce platforms have developed rapidly, and the volume of online merchandise transactions has shown rapid growth. While enjoying the convenience brought by online shopping, the dangers of counterfeit goods cannot be underestimated [1], the consumers around the world have become increasingly concerned about product safety. [2] Although the national regulatory authorities are also conducting strict checks on goods and cracking down on counterfeiting, fake and inferior products have caused serious harm to the world economy and consumers. [3] One possible solution to this never-ending problem is the adoption of a blockchain based anti-counterfeit system. [4] Even though there is no direct correlation between counterfeited products and sales loss, research has shown that companies that are targeted by counterfeiters incur larger operating costs to combat counterfeits. [5] Therefore, the establishment of commodity anti-counterfeit traceability system is of great significance in safeguarding market order, protecting consumers' rights and interests, and improving enterprise brand value.

With the deepening of research, anti-counterfeit traceability technology has flourished, in response to the growing concern, innovative, fully-functional, integrable and affordable product anti-counterfeiting solutions, utilizing cutting-edge technologies, have been widely

ICBBEM 2023, May 19-21, Hangzhou, People's Republic of China

DOI 10.4108/eai.19-5-2023.2334270

and urgently demanded. [6] Lu et al [7] combined blockchain technology with IoT technology to develop a food anti-counterfeit traceability system, where food traceability data are stored through a federated chain, which has lower transaction latency and communication cost than most anti-counterfeit traceability systems using private chains; meanwhile, the authenticity and reliability of source data in the blockchain are guaranteed by using IoT technology. Qiu et al [8] built a traceability anti-counterfeiting system based on edge computing for ownership on blockchain, proposed an efficient and secure ownership data sharing based on blockchain, and introduced a hybrid architecture based on blockchain. Yan et al [9] added 2D code visual features to the system, using the texture of fiber paper to generate natural texture features and also manually generated microscopic features; the features are generated and registered when the 2D code is printed on the fiber paper; the verification phase obtains the features by a feature extraction algorithm and compares them with the records to calculate the similarity. Liu et al [10] formed a unique unforgeable ceramic fingerprint by extracting the size and location of the underglaze bubbles formed in ceramics during firing, and used a blockchain to record the correspondence between the ceramic fingerprint and the ceramics. Chen et al [11] proposed a traceability system based on Hyperledger Fabric that addresses performance scalability and privacy issues, and uses a chain code algorithm to maintain and constrain the ledger. Although the existing systems can realize the function of anti-counterfeit traceability, none of them consider the situation that the commodity package is stolen, i.e., the problem at the source of checking. In this paper, we introduce big data analysis technology to analyze and calculate the inspection records and derive the reliability of this inspection, which can solve the above problems to a certain extent. At the same time, a set of anti-counterfeit traceability system about commodities is developed by using technologies such as QR code and block chain, which has good anti-counterfeit traceability effect.

2. Preliminaries

2.1. QR Code Technology

Quick Response Matrix Code, or QR Code (Quick Response Code), is a type of two-dimensional barcode where the inventor wanted the content of the code to be decoded quickly. Generally, two-dimensional code has advantages of large storage capacity, high recognition rate and high confidentiality, which can effectively improve the efficiency of information collection. [12] With the widespread use of smart phones, two-dimensional code is known to the public, and the application is becoming more and more widely. [13] The QR codes mentioned in this system all refer specifically to QR codes, which are used as carriers of commodity codes, as shown in Fig.1, while QR codes are combined with anti-counterfeit verification codes to achieve better anti-counterfeit effects.

Mksmie

Fig.1 Commodity code carrier-QR code

2.2. Blockchain Technology

Blockchain is a decentralized and decentralized network ledger, and the cryptography is used to secure the identity of participants and ensure that the ledger is tamper-proof. [14] The network ledger constitutes a system in which the lowest and most fundamental layer is the data structure, the structure of which is shown in Fig.2. This structure is where information and data are organized in a certain way and format, fed into and processed by the blockchain system. As a future database technology, blockchain is open, digital, and peer-to-peer in that every node maintains the integrity and authenticity of the ledger. [15] Once the data is input in a unified way, the network layer starts to connect, broadcast and verify among the nodes in the whole network, and then the block is constructed by the consensus of the whole network in the consensus layer.

Blockchain records transactions from the very beginning. Nodes can access any block in order to get insight into public information within the block. [16] Once a block is created and authenticated, it is almost impossible to change the data in the block as it would necessitate all subsequent blocks to be changed. Altering the blocks would be extremely challenging computationally and requiring enormous amounts of power. [17] Therefore, the use of blockchain technology can ensure the authenticity and reliability of data and realize the need for traceability.

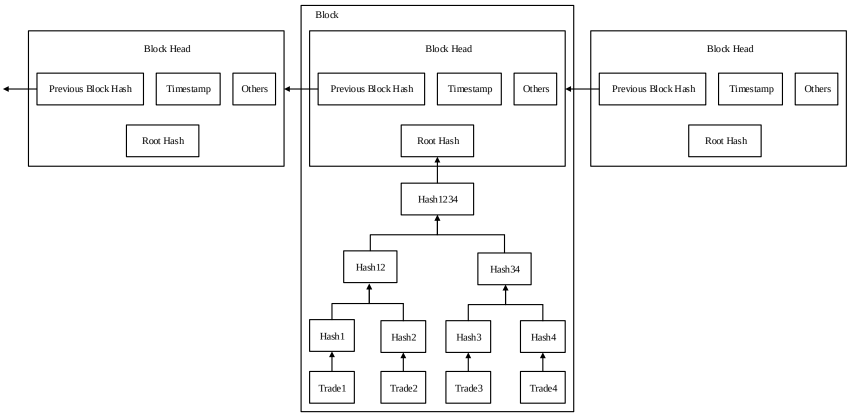

Fig.2 Blockchain structure diagram

2.3. Big Data Analytics Technology

With digital transformation, big data technologies have also found their general place in everyday business. [18] Big data analytics refers to the methodologies and techniques, as well as the software and hardware, that are used to receive, manage, and analyze enormous amounts of unstructured and structured information in actual time. [19] Hadoop is a technology that meets the needs of big data. It is horizontally scalable and designed as a software framework for processing very large amounts of data. [20] MapReduce is a programming model to express a distributed computation on a massive scale. MapReduce is a way of breaking down each request into smaller requests that are sent to many small servers to make the most scalable use of the CPU possible. [21] This system uses the big data analysis technology to operate the inspection history records, as shown in Fig.3 (see formula 1~4 for details), and finally derives the reliability pi of this inspection, and achieves the function of identifying the authenticity of goods by setting the reliability threshold.

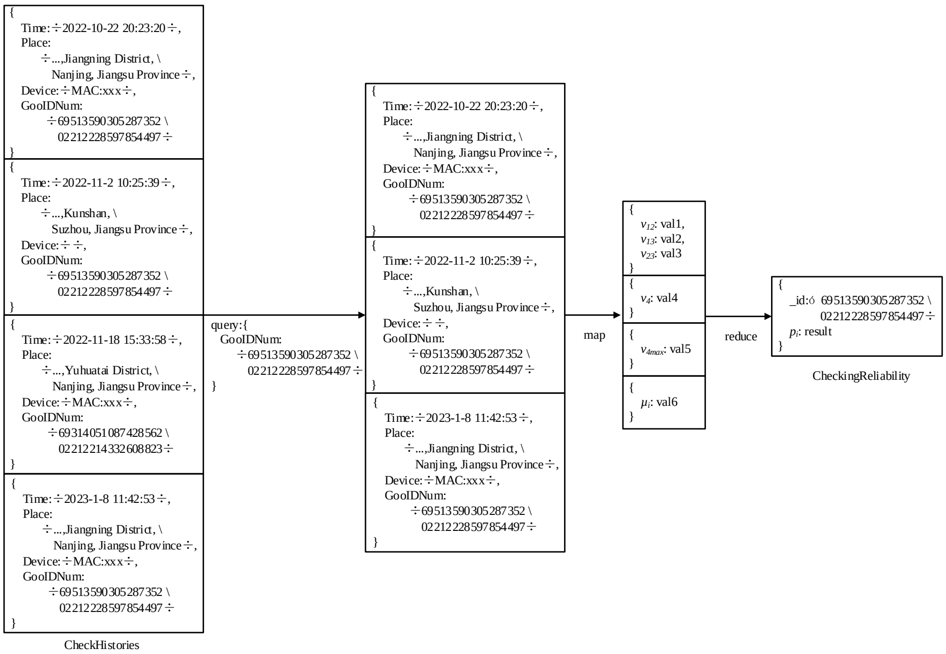

Fig.3 Calculation of pi using Map/Reduce technology

3. System Design and Implementation

The anti-counterfeit traceability system serves between merchants and consumers, and develops a commodity management system for merchants and a commodity sales system for consumers. The system architecture diagram is shown in Fig.4, and the overall design is as follows:

When a merchant increases the quantity of a certain commodity in the commodity management system, the system will generate an equal and unique code, and generate a QR code according to the code, while each QR code is used with an anti-counterfeiting verification code (referred to as verification code), and the merchant can print the QR code and verification code on the commodity package, as shown in Fig.1, and use the anti-counterfeiting layer to cover the verification code; with the unique code, the merchant can enter the details of this commodity into the system. With the unique code, the merchant can enter the detailed information of the production, sales and transportation of this commodity into the system, and realize the decentralized storage of the data and the characteristics of non-tamper ability through blockchain technology; after the commodity information is registered and released, it will be displayed in the commodity sales system for consumers to view and purchase; After the system receives the order, it will generate the order information, and after the merchant ships the goods, it will record the logistics information; when the consumer receives the purchased goods, he or she can make the request of scanning the code check, and input the verification code under the coating according to the system prompt; after the system verifies that the commodity code and the verification code correspond correctly, it will use the big data technology to analyze the time, place, number of checks and other factors of all the historical check records of the code goods When the reliability is lower than the threshold value set by the system, the user will be prompted to enter the piracy risk, otherwise, the detailed information of the commodity will be displayed to the user, and the user can also view all the historical inspection records of the commodity.

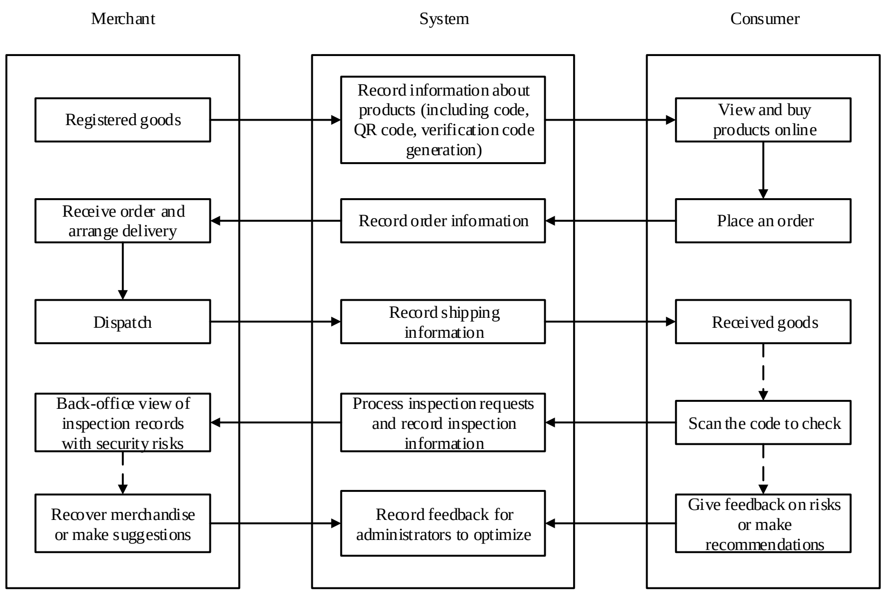

Fig.4 System architecture diagram

In summary, the core of the system is the unique coding of goods and the calculation of the reliability of inspection, which is explained in detail below.

3.1. Unique code of the product

One of the most effective ways to achieve the anti-counterfeiting function is to assign a unique code to each product, i.e. "one item, one code". The purpose of coding is to ensure the uniqueness of each product, and each product has a unique code, which is stored in the system to ensure that the forged code cannot be effectively identified; at the same time, for the regularity of coding, the product code used in the system refers to the Global Trade Item Number (GTIN), whose coding structure is shown in Fig.5, and is expanded accordingly.

The expanded coding structure according to GTIN is shown in Fig.6, the rules are as follows: use 3 bits to identify the country (directly use the country code in GTIN, such as Chinese country code is 690-699); use 13 bits to identify the company code (13 decimal digits can indicate 1 trillion different goods); use "batch + random" coding form for product coding. The product code is in the form of "batch + random" code, with 17 bits of identification, including 8 bits of batch + 9 bits of random code, the batch directly using the date, 9 bits of random code can identify 100 million different numbers, calculated: each company can code up to 100 million different goods per day; the last 1 bit is the check code, to prevent errors in the transmission process of the code.

	Global Trade Item Number(GTIN)														
	GS1 Application Identifier	GS1-8 digit prefix code or manufacturer identification code							Manufacturer Product Code						Validation Code
(GTIN-8)	0 1	0	0	0	0	0	0	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
(GTIN-12)	0 1	0	0	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}
(GTIN-13)	0 1	0	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}
(GTIN-14)	0 1	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}

Fig.5 GITN coding structure

Country Prefix Code	Company Prefix Code	Batch	Random Code	Validation Code
$N_1 N_2 N_3$	$N_4 N_5 N_6 N_7 N_8 N_9 N_{10} N_{11} N_{12} N_{13} N_{14} N_{15} N_{16}$	$N_{17} N_{18} N_{19} N_{20} N_{21} N_{22} N_{23} N_{24}$	$N_{25} N_{26} N_{27} N_{28} N_{29} N_{30} N_{31} N_{32} N_{33}$	N_{34}

Fig.6 Example of the expanded code

In summary, the 34-digit decimal number is used to encode the products to meet the "one item, one code" requirement. In addition, the system generates a 6-digit verification code for each code, and the purpose of the verification code is to minimize the possibility of direct reuse of the code. The verification code is covered with an anti-counterfeit coating, and the consumer receives the product and carefully observes whether the coating is damaged. The 6-digit verification code is obtained by scratching the coating and used for inspection.

3.2. Calculation of check reliability

The overall checking process of the system is shown in Fig.7, after the user scans the QR code and enters the verification code, if the QR code and the verification code correspond correctly, the system will use big data technology to calculate and analyze all the historical checking records of the coded goods, including the "movement rate v_{ij}" and "maximum movement rate

$v_{i\max}$ " of any two adjacent checking records, as well as the "average movement rate $\bar{v}_i$ " and "variance movement rate v_{dev} " of all historical checking records; v_{ij}, $\bar{v}_i$, $v_{i\max}$ and v_{dev} are used as the influence factors of check reliability, which are defined as shown in Equation (1)~(4), where T_i indicates the check time, S_i indicates the check location (the location of logistics sign-off is S_0 and time is T_0).

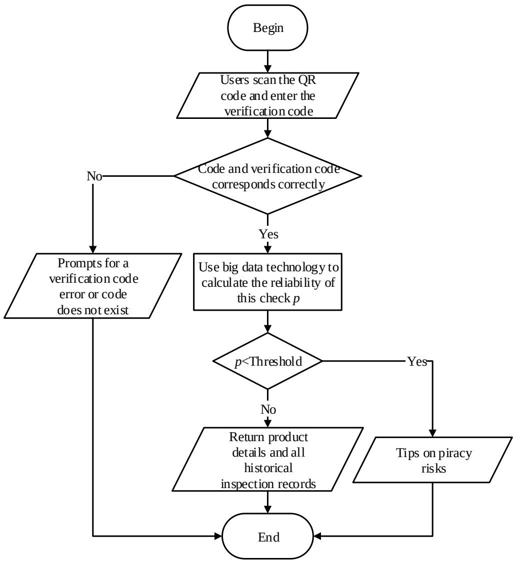

Fig.7 Checking flow chart

The latitude and longitude of the check address are stored in the check information, and the distance calculated using Equation (5), whose accuracy is similar to that of Google Maps, is used as the distance between the two check records, where: R is a constant indicating the radius of the Earth (*R≈6371000m*), $a = lat1 - lat2$, $b = (lon1 - lon2) \times \frac{\pi}{180}$, $lati$ indicates the dimension, $loni$ indicates the longitude, and $i \in \{1, 2\}$.

$$v_{ij} = \begin{cases} \left| \frac{|S_j S_i|}{T_j - T_i} \right| & (i \neq j) \\ 0 & (i = j) \end{cases} \tag{1}$$

$$\bar{v}_i = \frac{1}{i} \sum_{k=1}^{i} v_{k(k-1)} \tag{2}$$

$$v_{i\max} = \max_{0 \le k < m \le i} \sum_{k=0}^{i} \sum_{m=k+1}^{i} v_{km} \tag{3}$$

$$v_{dev} = \frac{1}{i} \sum_{k=1}^{i} (v_{k(k-1)} - \bar{v}_i)^2 \tag{4}$$

$$\left|S_j S_i\right| = 2R \cdot \arcsin\sqrt{\sin^2\frac{a}{2} + \cos lat1 \cdot \cos lat2 \cdot \sin^2\frac{b}{2}} \tag{5}$$

The number of checks N is also an important influence factor, the more checks, the lower its reliability will be; in addition, N_{addr} indicates the total number of different check locations (different at district/county level), the more the number of different check locations, the lower its reliability will be.

Further, the reliability of this check p_i $(p_i \in [0,100\%])$ is calculated based on the above check factors, as shown in Equation (6)~(7), where $\sigma_k \in (0,1)$ and $\sum_{k=1}^{6}\sigma_k = 1$. Two points need to be clarified: first, the check first stores this check record and then takes out all historical check records, so $N \in N^*$, the that is, positive integers (excluding 0), and $N_{addr} \in N$, that is, natural numbers (including 0), need to be judged; second, the first check, the value of *i-1* in the formula at this time is 0, while S_0 and T_0 in the formula indicates the location and time of the logistics sign-off, so the first check is not simply to make the reliability equal to 100%, and its also needs to meet the reliability of the formula calculation.

$$p_i' = 1 - (\sigma_1(1 - \frac{1}{\bar{v}_i + 1}) + \sigma_2(1 - \frac{1}{v_{\max} + 1}) + \sigma_3(1 - \frac{1}{v_{i(i-1)} + 1}) + \sigma_4(1 - \frac{1}{v_{dev} + 1}) + \sigma_5(1 - \frac{1}{N})) \tag{6}$$

$$p_i = \begin{cases} p_i' & ,(N_{addr} = 0) \\ p_i' - \sigma_6(1 - \frac{1}{N_{addr}}) & ,(N_{addr} \neq 0) \end{cases} \tag{7}$$

Set the corresponding weights for the above influence factors, and get the reliability p_i of this check by calculating the average value; when p_i is less than the set threshold, prompt the user that the goods do not have legality, otherwise, return the user's detailed goods information and historical check information.

4. Experiment

The previous section elaborated the calculation principle of the checking reliability in detail. The influence factors affecting the reliability of the check are mainly $v_{i(i-1)}$, $\bar{v}_i$, $v_{i\max}$, v_{dev}, N and N_{addr}, where $v_{i(i-1)}$, $\bar{v}_i$, $v_{i\max}$ and v_{dev} take values in the range of [0,+∞), and the relationships with the reliability p_i are shown in Fig.8, respectively; N takes values in the range of N^*, the set of positive integers, and N_{addr} takes values in the range of N, the set of natural numbers, and the relationships with the reliability p_i are shown in Fig.9, respectively.

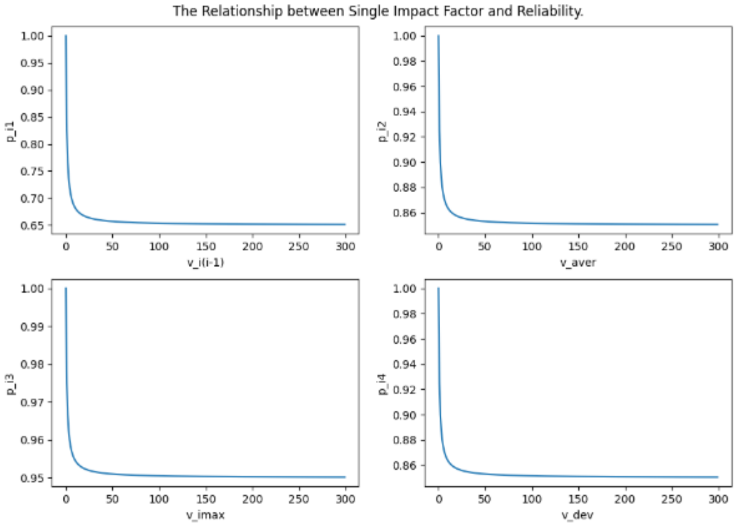

Fig.8 The relationship between $v_{i(i-1)}$、 $\bar{v}_i$ 、 $v_{i\max}$ 、 v_{dev} and p_i

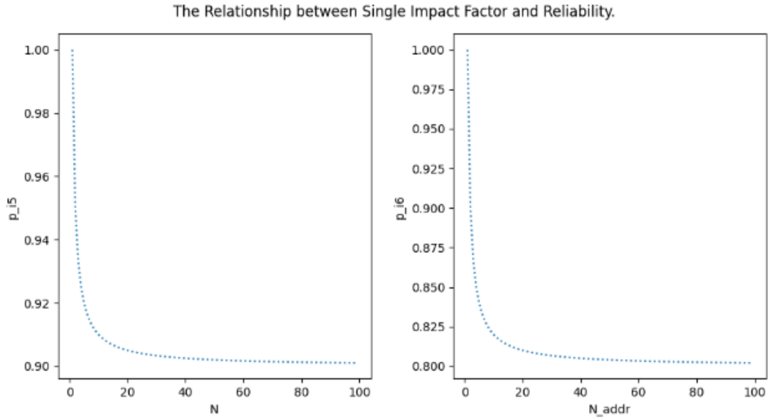

Fig.9 The relationship between N、 N_{addr} and p_i

After analyzing the relationship between individual influence factors and reliability, we try to add some constraints to test the relationship between all influence factors and the reliability of the checks. As shown in Fig.10 for a commodity in six months 50 times fixed position check, there exist 7 times according to this fixed position within 500km range on the random position, each time the value of check reliability, where the red point indicates that the check reliability is lower than the threshold value, from the figure can be seen that the system detects the existence of 8 times check abnormalities.

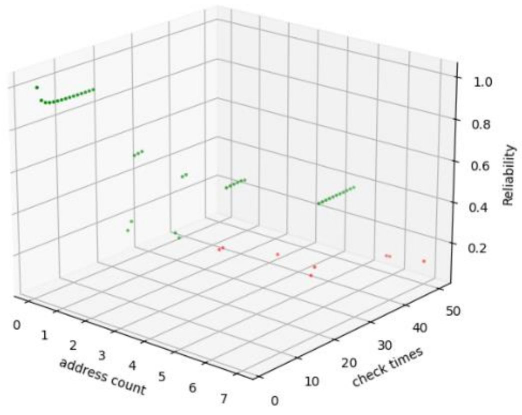

Fig.10 Simulation of the reliability of a commodity for 50 checks

Further, we increase the scale of the experiment, in a period of 1 year, the number of checks for a commodity reaches 140 times, of which 25 times the check location is within 9,600km from the fixed location, and the value of the reliability of each check is shown in Fig.11, where the red point indicates that the check reliability is below the threshold, and it can be seen from the figure that the system detects the existence of 24 check abnormalities.

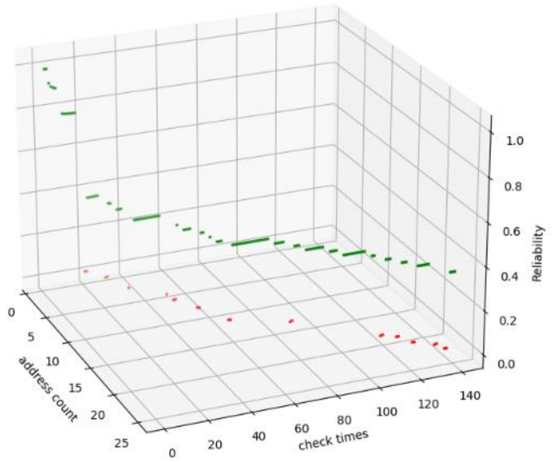

Fig.11 Simulation of the reliability of 140 checks for a commodity

As shown in Fig.12, the effect of the anti-counterfeiting check system is shown. Compared with the traditional check, there is an additional probability of genuine goods, i.e., check reliability, which can be achieved by controlling the threshold value of reliability to identify the authenticity of goods.

Fig.12 Commodity inspection results chart

5. Conclusion

This paper introduces in detail the use of big data technology for anti-counterfeiting of commodities and develops a set of anti-counterfeiting traceability system for commodities based on big data technology and blockchain technology, which ensures the reliability and traceability of commodities from several aspects and injects a new idea for anti-counterfeiting. This system not only provides a friendly checking experience for consumers, but also can safeguard the legitimate rights and interests of enterprises, and has good application prospects. However, there are shortcomings in the system, such as over-dependence on the time and place of checking and resale problems, etc. Later, the system will be further improved by adding machine learning-related knowledge and predicting the probability of QR code being reused, so as to achieve better anti-counterfeiting effects.

References

[1] Zheng Xin, Zhu Yangbin, Liu Yang, Zhou Linpeng, Xu Zhongwei, Feng Chen, Zheng Chunbo, Zheng Yueting, Bai Jieyu, Yang Kaiyu, Zhu Dongyan, Yao Jianmin, Hu Hailong, Zheng Yuanhui, Guo Tailiang, Li Fushan. Inkjet-Printed Quantum Dot Fluorescent Security Labels with Triple-Level Optical Encryption. [J]. ACS applied materials & interfaces,2021,13(13).

[2] Li J, Yu Y, Hu S, et al. An authority management framework based on fabric and ipfs in traceability systems[C]//Blockchain and Trustworthy Systems: First International Conference, BlockSys 2019, Guangzhou, China, December 7–8, 2019, Proceedings 1. Springer Singapore, 2020: 761-773.

[3] Zheng Zhaohui, Zheng Hong, Ju Jianping, Chen Deng, Li Xi, Guo Zhongyuan, You Changhui, Lin Mingyu. A system for identifying an anti-counterfeiting pattern based on the statistical difference in key image regions[J]. Expert Systems With Applications,2021,183.

[4] Anthony, Lee Michael Christian, Pearl Rafaelle Richel, Edbert Ivan Sebastian, Suhartono Derwin. Developing an anti-counterfeit system using blockchain technology[J]. Procedia Computer Science,2023,216.
[5] Butticè V, Caviggioli F, Franzoni C, et al. Counterfeiting in digital technologies: An empirical analysis of the economic performance and innovative activities of affected companies[J]. Research Policy, 2020, 49(5): 103959.
[6] Yiu Neo C. K.. Decentralizing Supply Chain Anti-Counterfeiting and Traceability Systems Using Blockchain Technology[J]. Future Internet,2021,13(4).
[7] Lu Yi, Li Peng, Xu He. A Food anti-counterfeiting traceability system based on Blockchain and Internet of Things[J]. Procedia Computer Science,2022,199.
[8] Zehuan Qiu, YiFan Zhu. Traceability anti-counterfeiting system based on the ownership of edge computing on the blockchain[J]. Journal of Ambient Intelligence and Humanized Computing,2021(prepublish).
[9] Yan Yulong, Zou Zhuo, Xie Hui, Gao Yu, Zheng Lirong. An IoT-Based Anti-Counterfeiting System Using Visual Features on QR Code[J]. IEEE INTERNET OF THINGS JOURNAL,2021,8(8).
[10] Liu X, Cheng X, Liao C, et al. Ceramic Anti-Counterfeiting Technology Identification Method Based on Blockchain[C]//2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom). IEEE, 2021: 36-41.
[11] Chen C L, Shang X, Tsaur W J, et al. An anti-counterfeit and traceable management system for brand clothing with hyperledger fabric framework[J]. Symmetry, 2021, 13(11): 2048.
[12] Jing Rao. Operation mode of electric business logistics based on the application of two-dimensional code technology[J]. International Journal of Metrology and Quality Engineering,2020,11.
[13] Yue Zhang, Yang Zhang, Shiwei Fan, Kai Wang, Pan Jiang. Indoor positioning and navigation based on two-dimensional code and database[J]. The Journal of Engineering,2019,2019(23).
[14] Li Xiaoming. Inventory management and information sharing based on blockchain technology[J]. Computers & Industrial Engineering,2023,179.
[15] Kumar Singh Atul, Kumar V.R. Prasath, Dehdasht Gholamreza, Mohandes Saeed Reza, Manu Patrick, Pour Rahimian Farzad. Investigating the barriers to the adoption of blockchain technology in sustainable construction projects[J]. Journal of Cleaner Production,2023,403.
[16] Tišma Sanja, Škrtić Mira Mileusnić. Blockchain Technology in the Environmental Economics: A Service for a Holistic and Integrated Life Cycle Sustainability Assessment[J]. Journal of Risk and Financial Management,2023,16(3).
[17] Allareddy Veerasathpurush, Rampa Sankeerth, Venugopalan Shankar Rengasamy, Elnagar Mohammed H, Lee Min Kyeong, Oubaidin Maysaa, Yadav Sumit. Blockchain technology and federated machine learning for collaborative initiatives in orthodontics and craniofacial health. [J]. Orthodontics & craniofacial research,2023.
[18] Meier Thomas, Makyšová Helena, Pauliková Alena. Evaluation of the Economic, Ecological and Ethical Potential of Big Data Solutions for a Digital Utopia in Logistics[J]. Sustainability,2023,15(6).
[19] Abd Algani Yousef Methkal, Vinodhini G Arul Freeda, Isabels K. Ruth, Kaur Chamandeep, Treve Mark, Kiran Bala B., Balaji S., Devi G. Usha. Analyze the anomalous behavior of wireless networking using the big data analytics[J]. Measurement: Sensors,2022,23.
[20] Stuardo C A, Leesatapornwongsa T, Suminto R O, et al. ScaleCheck: A Single-Machine Approach for Discovering Scalability Bugs in Large Distributed Systems[C]//FAST. 2019: 359-373.
[21] Azeroual Otmane, Fabre Renaud. Processing Big Data with Apache Hadoop in the Current Challenging Era of COVID-19[J]. Big Data and Cognitive Computing,2021,5(1).