The Credibility Measurement Model of Food Safety On-chain Data based on Blockchain

Hongwei Tao[1], Yinghui Hu[1], Hui Li[2*], Deqiang Fan[3], Haoran Chen[1]

[1] College of Computer and Communication Engineering, Zhengzhou University of Light Industry, China
[2] College of Economics, Qingdao Agricultural University, China
[3] Henan New Landmark Construction Engineering Limited Company, China
tthhww_811@163.com, hyingh6@163.com, leephil@163.com, 93613155@qq.com, chenhaoran@zzuli.edu.cn

Abstract

Food safety is related to the national economy and people's livelihood and has always been the focus of the people and the government. Blockchain technology has characteristics of being decentralized, tamper-free, and having underlying openness. It can record and trace product information, prevent data tampering, effectively enhance the transparency of product information, and provide new methods and ideas for food safety traceability. At present, research hotspots mainly focus on the design and construction of a trusted blockchain traceability system, but the provided blockchain traceability system cannot provide a way to verify the authenticity of the information. This paper studies the credibility evaluation model of the members involved in the blockchain and the on-chain data quality model and provides a method to solve the credibility of the on-chain data. Meanwhile, the effectiveness of the method is validated by a case study.

Keywords: Blockchain, Food safety traceability, Data credibility measurement, D-S evidence theory

1 Introduction

In recent years, frequent food safety problems have seriously threatened people's health and life safety. The World Health Organization estimates that more than 400,000 people die from food-borne diseases every year. Food safety issues have attracted great attention from countries around the world. For example, the Canadian government has mandated the use of labels and barcodes to identify the source of products [1]. Similarly, the European Union, the United States, Japan, Canada, and South Korea have all taken coercive measures to trace the source of food. As an important country for the production and export of agricultural products, China's food quality and safety has been the focus of the government and consumers in recent years, and food quality and safety traceability has been paid more and more attentions. In 2015, the Food Safety Law of the People's Republic of China incorporated the establishment of a full traceability food safety system into the law to realize the two-way traceability function of products from raw materials to finished products and from finished products to raw materials [2]. In 2018, the China Food Safety Development Report pointed out that China's food safety risks are still prominent, and they are characterized by concealment, durability, and complexity. Eliminating the information asymmetry between the participating parties in the food supply chain is one of the major research topics for society to solve the food safety problem.

At present, traceability systems at home and abroad mainly adopt centralized methods to store relevant data information in centralized servers. However, some enterprises are unwilling to share food traceability information from their own interests, and even forge product traceability information. Therefore, the existing food supply chain traceability systems cannot guarantee the reliability of traceability results. In addition, in today's food supply chain network, the process of achieving traceability is time-consuming and complex. All these have brought a certain degree of difficulty to the traceability of food safety.

Blockchain technology is considered to be one of the most disruptive and revolutionary innovations in recent years. It has emerged in many fields, such as finance, trade, the Internet of Things, and the sharing economy [3-4]. It is essentially a technology that provides trust and decentralization. The characteristics of decentralization, automation, and trustworthiness of blockchain technology can provide more effective data protection, which is in line with the needs of people's new food safety traceability system. The blockchain technology can improve the traceability efficiency of the traceability system and ensure the authenticity of the traceability results. Nowadays, the traceability system of the food supply chain has become an important application direction of blockchain technology.

2 Related Works

In 1991, Haber et al. published an article titled "How to Add Timestamps to Digital Documents" [5], which is considered to be the prototype of blockchain technology. In this article, time stamps are added to digital documents to ensure the immutability and irreproducibility of digital contents. After a series of studies [6-7], in 2009, Satoshi Nakamoto published "Bitcoin: A Peer-to-Peer Digital Cash System" [8], which proposed the blockchain technology and gave its practical applications. Blockchain technology is a model innovation technology integrating distributed ledgers, cryptography, smart contracts, and consensus mechanisms [9-10]. It has the characteristics of decentralization, being

*Corresponding Author: Hui Li; E-mail: leephil@163.com
DOI: 10.53106/160792642022072304007

tamper-free, and being bottom opening. Blockchain technology can provide new methods and ideas for food safety traceability. Scholars at home and abroad have begun to explore the research and development of traceability systems based on blockchain technology [11-12]. For example, Daniel Tse et al. believe that the source of products in the supply chain must be transparent, tamper-proof, and adaptable to the changing environment. Therefore, a system based on private and public chains was designed [13]. Reference [14] presented a solution based on blockchain and the Internet of Things, which is used in agricultural food supply chain to ensure information security. Jintao Hao et al. proposed the use of IPFS and auxiliary databases to achieve agricultural food storage and traceability [15]. Andreas Kamilaris et al. illustrated the advantages of blockchain technology in food safety and critically analysed and studied the maturity, challenges, and potential of existing or ongoing projects [16]. Yupin Lin et al. presented an e-agriculture system and assessment tool [17]. Affaf Shahid et al. gave a complete solution for blockchain-based agriculture and food supply chain, which was deployed over the ethereum blockchain network [18]. Sachin S. Kamble et al. built a blockchain model to ensure the traceability of the agricultural supply chain [19].

In summary, most studies have focused on highlighting the benefits and value of organizations using blockchain technology, but little research has been done on the credibility of the on-chain data itself. Moreover, although there are many blockchain traceability projects in the market, whether listed agricultural companies, traditional technology companies, or small start-up companies are in the stage of exploring blockchain traceability, and the provided blockchain traceability system cannot show a way to verify the authenticity of the information. Consequently, the general public can only rely on the information provided by the third-party quality inspection agencies to judge the authenticity of the information. Therefore, based on the theoretical research of traditional data quality analysis, this paper puts forward a blockchain-based, dynamic, and hierarchical approach to build a credibility model for food safety on-chain data, which can not only enable users to directly understand the current data state, but also provide credibility reference for future data application and analysis.

3 Measurement Models

In order to solve the credibility problem of on-chain data, this paper presents a method to construct a credibility measurement model for on-chain data. The model is divided into two parts: the credibility measurement model of on-chain member and the on-chain data quality model. The credibility measurement model of an on-chain member consists of the initial credible degree and the consensus degree of the uploaded data; the measures and synthesis of data quality attributes constitute the on-chain data quality model. Finally, these two models are fused through evidence theory to measure the credibility of on-chain data, as shown in Figure 1.

To facilitate the understanding of the model construction method proposed in this paper, the relevant definitions are given below.

The third-party quality inspection agencies refer to the organizations or organizations in the blockchain field that have reliable ability to implement the certification system and can independently, objectively, and impartially engage in certification activities in the whole process of certification.

On-chain members refer to all members who purchase and use things described by on-chain data (specifically food here).

Consensus network refers to the network composed of all related accounting nodes with consistent transaction data and block information.

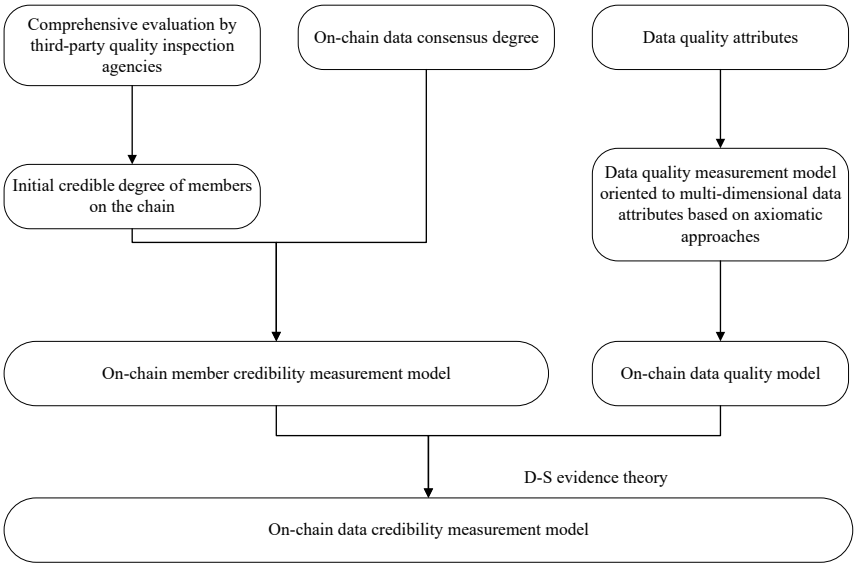

Figure 1. On-chain data credibility measurement model framework

3.1 On-Chain Member Credibility Measurement Model

In this subsection, we establish the on-chain member credibility measurement model.

Definition 1 (On-Chain Member Credibility Measurement Model) The on-chain member credibility measurement model consists of the initial credible degree of the member on the chain and the consensus degree of the uploaded data. Let $T_1(B,Data,t)$ be the credibility measurement model of on-chain

member *B* about uploaded data *Data* at time *t*, then it is defined as Equation (1).

$$T_1(B, Data, t) = \lambda \cdot InitT(B, t) + \xi \cdot ConsenT(B, Data, t). \quad \textbf{(1)}$$

where

1) λ, ξ are used to distinguish the contributions of initial credibility and consensus of the uploaded data *Data* to the credibility of member *B* at time *t*, which satisfy $\lambda + \xi = 1$, $0 < \lambda, \xi < 1$.

2) $InitT(B,t) = ThirdT(B,t)$. $ThirdT(B,t)$ represents the credibility evaluation of the third-party quality inspection agencies to the member *B* at time *t*.

3) *ConsenT* (*B*, *Data*, *t*) indicates the consensus degree of the uploaded data *Data* by *B* at time *t*, which is defined as Equations (2) and (3).

$$ConsenT(B, Data, t) = \begin{cases} InitT(B,t), t = 0 \\ ConsenT(B, Data, t-\Delta t).\sigma_T(t), \Delta Context(N,B,t) \neq 0, t - \Delta t > 0 \\ ConsenT(B, Data, t-\Delta t).\mu_T(t), \Delta Context(N,B,t) = 0, \Delta t > Time, \end{cases} \quad \textbf{(2)}$$

$$ConsenT(B, Data, t) = ConsenT(B, Data, t\text{-}\Delta t) * (1 + \theta_T(t))/2, \quad \textbf{(3)}$$

of which, $\sigma_T(t)$ represents the penalty coefficient used at time *t* due to inconsistent data, which is determined as Equation (4).

$$\sigma_T(t) = \begin{cases} 1, \Delta Context(N,B,t) = 0. \\ 0, \Delta Context(N,B,t) \neq 0. \end{cases} \quad \textbf{(4)}$$

N is a set composed of at least half of the members on the chain. $\Delta Context(N,B,t) \neq 0$ means that for the same data, the data hash of member *B* is inconsistent with the data hash of at least half of the members on the chain, $\Delta Context(N,B,t) = 0$ implies that for the same data, the data hash of member *B* is consistent with the data hash of at least half of the members on the chain. *Time* represents maximum time interval when data is not updated, which can be set by the user according to the actual situation. $\mu_T(t)\left(0 \le \mu_T(t) \le 1\right)$ represents the time attenuation coefficient at time *t*. If the data of member *B* is not updated between time *t*-Δt and *t*, then the time decay penalty will be imposed on it. $\theta_T(t)\left(0 \le \theta_T(t) \le 1\right)$ is the credibility restoration coefficient at time *t*. If the member *B* on the chain successfully participates in the consensus process for a specified number of times between *t*-Δt and *t*, Equation (3) is used to gradually restore its credibility.

3.2 On-Chain Data Quality Model

Data quality assessment is the scientific and statistical evaluation process of data to determine whether they meet the quality required by the project or business process and can truly support the correct type and quantity of their intended use [20]. Data quality assessment mainly includes the selection, measures, and synthesis of data quality attributes [21-23].

Different organizations and users choose data quality attributes differently, and their determination depends on specific businesses and user needs. The on-chain data quality model presented in this section supports both objective and subjective evaluation. Objective evaluation is aimed at the dataset itself, while subjective evaluation is mainly the feedback evaluation after users use the dataset. Since this article is oriented to food safety traceability, the data quality attributes are divided into critical attributes and non-critical attributes for the on-chain food data. The critical attributes are extracted from the attributes involved in the definitions related to data quality, including accuracy, completeness, accessibility, and integrity. In contrast, the non-critical attributes are given based on specific application scenarios, for example readability.

Suppose that there are *k* data sources to form a data source set denoted as $S=\{S_1, S_2, \ldots, S_k\}$, $k \in N^+$, indicating that they come from *k* different trusted blockchain nodes. $E_1, E_2, \ldots, E_m$, are *m* records of a certain data source in *S*, which form a data set $D=\{E_1, E_2, \ldots, E_m\}$, $m \in N^+$, and each data record has *n* attributes, expressed as $E_i = \{T_{i1}, \ldots, T_{in}\}$, $n \in N^+$, where T_{ij} represents the value of the *j*-th attribute of the record E_i. Let $R = \{R_{11}, R_{12}, \ldots, R_{mn}\}$ be the authoritative reference data source, where R_{ij} represents the correct or expected value of the record E_i on the attribute *j*. The following are the specific definitions of various attributes and their quantification methods.

Accuracy: Accuracy is represented by the symbol y_1. It refers to the accuracy of the description of each attribute value of the data on the chain. For example, timeliness is one of the considerations for ensuring food quality and safety. Food-related timestamps must be consistent and accurate with the time of the local server in the consensus network. Let $F_1(\cdot)$ be the mapping which is defined as Equation (5).

$$F_1(T_{ij}) = \begin{cases} 1, T_{ij} = R_{ij}. \\ 0, otherwise. \end{cases} \quad \textbf{(5)}$$

Then the accuracy of *D* on attribute *j*, as in:

$$Acc = \sum_{i=1}^{m} F_1(T_{ij}) / m. \quad \textbf{(6)}$$

The accuracy of *D* on all attributes is given in Equation (7).

$$y_1 = \sum_{j=1}^{n} \sum_{i=1}^{m} F_1(T_{ij}) / (m * n). \quad \textbf{(7)}$$

Completeness: Completeness is represented by the symbol y_2. It indicates the completeness of the description of any attribute of any record. For example, a null value in a record is an indication of incompleteness. Let $F_2(\cdot)$ be the mapping from the value result of the evaluation object T_{ij} to {0,1}, the value is 1 if meet the conditions. Otherwise, it is 0, as in

$$F_2(T_{ij}) = \begin{cases} 1, otherwise. \\ 0, T_{ij} = 0 \ or \ null. \end{cases} \quad \textbf{(8)}$$

The completeness of D on all attributes is presented in Equation (9).

$$y_2 = \sum_{j=1}^{n}\sum_{i=1}^{m} F_2\left(T_{ij}\right)/\left(m*n\right). \quad \textbf{(9)}$$

Accessibility: Accessibility is represented by the symbol y_3. It means that the on-chain data is public, allowing authorized users to easily obtain and use. This attribute is closely related to data disclosure. The higher the degree of data accessibility, the easier it is to obtain, and the lower the data will be tampered with. Let s represent the total number of nodes in the consensus network, and a represent the number of nodes that have lost connection. The accessibility of D is proposed in Equation (10).

$$y_3 = (s-a)/s. \quad \textbf{(10)}$$

Integrity: Integrity is represented by the symbol y_4. This attribute mainly evaluates non-numerical data. If the data has not been modified in any way, the hash value of the data stored in different nodes is the same. The integrity of D is given in Equation (11).

$$y_4 = \begin{cases} 1, & (\exists S_k \in S)\ such\ that\ D = S_k. \\ 0, & otherwise. \end{cases} \quad \textbf{(11)}$$

Readability: Readability is represented by the symbol y_5. It means that the data on the chain is expressed in a standard way with good terms, attributes, units, codes, or abbreviations, so that people can understand and interpret them correctly. Suppose k represents the number of records in D that contain annotated information, then the readability of D is given in Equation (12).

$$y_5 = k/m. \quad \textbf{(12)}$$

Our data quality model is based on axiomatic approaches for multi-dimensional data attributes [24].

Definition 2 (On-Chain Data Quality Model) The on-chain data quality model about data *Data* is defined as Equation (13).

$$T_2(Data) = {y_1}^{\alpha_1} * {y_2}^{\alpha_2} * {y_3}^{\alpha_3} * {y_4}^{\alpha_4} * {y_5}^{\beta_5}. \quad \textbf{(13)}$$

Where

1) y_1, y_2, y_3, y_4 are the degrees of critical attributes and y_5 is the degree of non-critical attributes, which are computed by Equations (7), (9), (10), (11), and (12) respectively. .

2) $\alpha_i (1 \le i \le 4)$ and β_5 are used to distinguish the contributions of critical attributes and non-critical attributes to data quality which satisfy that $\sum_{i=1}^{4}\alpha_i + \beta_5 = 1$, $0 \le \alpha_i, \beta_5 \le 1 (1 \le i \le 4)$.

The on-chain data quality model given here is a combination of power function products. On the one hand, it satisfies the nature of data quality measurement, which can be proved by simple calculation. On the other hand, it conforms to the "series rule", which reflects that each attribute is important. It also reflects the "barrel principle". In order to ensure the data quality, each attribute must be done well.

3.3 On-Chain Data Credibility Measurement Model

The deficiencies of the data itself are the primary problem faced in the credit analysis process. These deficiencies are manifested in many aspects, such as the "uncertainty" and "ignorance" of the data information. There are some mathematical theories that can effectively describe defective data, such as rough set theory, probability theory, possibility theory, Dempster-Shafer (D-S) evidence theory, fuzzy set theory, etc.

D-S evidence theory is a common method to deal with imprecision and uncertainty [25]. At present, this theory has been widely applied in the fields of fusion decision making, such as target recognition and fault diagnosis. In this section, we use D-S evidence theory to build the on-chain data credibility model based on the on-chain member credibility measurement model and the data quality model established previously.

We first introduce some symbols in the D-S evidence theory used here. X represents the recognition frame, 2^X is the set containing all the subsets of X, the function m is the mass function of frame X, as in

$$\begin{cases} m(\varnothing) = 0. \\ \sum_{A \in 2^X} m(A) = 1. \end{cases}$$

Because only two information sources need to be fused in this article, we consider the synthesis of two mass functions m_1 and m_2. The joint mass function $m_{1,2}$, is determined as Equation (14).

$$\begin{cases} m_{1,2}(\varnothing) = 0 \\ m_{1,2}(A) = (m_1 \oplus m_2)(A) = \dfrac{1}{K}\sum\limits_{\substack{E\cap C = A \neq \varnothing \\ E, C \subseteq X}} m_1(E) m_2(C) \end{cases}. \quad \textbf{(14)}$$

Where K is the amount of direct conflict between the two information sources m_1 and m_2, which is calculated by Equation (15).

$$K = 1 - \sum_{E\cap C = \varnothing} m_1(E) m_2(C). \quad \textbf{(15)}$$

Let m_1 credibility distribution follow $\{T_1(B,Data,t), 1\text{-}T_1(B,Data,t)\}$, m_2 credibility distribution follow $\{T_2(Data), 1\text{-}T_2(Data)\}$, then we can build the on-chain data credibility measurement model as follow.

Definition 3 (On-Chain Data Credibility Measurement Model) The on-chain data credibility measurement model of member B about uploaded data *Data* at time t is defined as Equation (16).

$$T(B,Data,t)=\frac{T_1(B,Data,t)\times T_2(Data)}{1-T_1(B,Data,t)\times(1-T_2(Data))-T_2(Data)\times(1-T_1(B,Data,t))} \quad (16)$$

4 Measurement Procedure

The on-chain data credibility measurement procedure is as follows. For the given data *Data*, time *t,* and the member *B*;

Step1: The initial credible degree of the member *B* is determined by third-party quality inspection agencies;

Step2: Equations (2) and (3) are used to calculate the consensus degree of uploaded data *Data*;

Step3: Based on the results of Step1 and Step2, the credibility of the member *B* about uploaded data *Data* at time *t* are calculated by using Equation (1);

Step4: The degrees of critical attributes y_1, y_2, y_3, y_4 are computed according to Equations (7), (9), (10), and (11) , the degree of non-critical attribute y_5 are calculated by Equation (12), the weights of critical attributes $\alpha_i (1 \le i \le 4)$ and the weight of non-critical attribute β_5 are determined, and the value of data quality is deserved by utilizing Equation (13);

Step5: The credible degree of the data *Data* is finally obtained according to Equation (16).

5 Case Study

To demonstrate the effectiveness of our measurement models, we use a set of data results to test these models in a simulated data environment. Table 1 shows the attributes and their weight distribution in the data quality model. Table 2 gives the default values of the parameters used in the experiment, which are selected by continuous debugging during the experiment and are somewhat empirical and subjective. Table 3 shows the credibility evaluation results of some representative data on the chain after D-S evidence theory.

According to the models established in section 3, for the given time *t* and the member *B*, the credibility evaluation result of the normal data in Table 3 is obtained as follows.

$$T_1(B,normal\ data,t)=0.800.$$

$$T_2(B,normal\ data,)=y_1^{\alpha_1}*y_2^{\alpha_2}*y_3^{\alpha_3}*y_4^{\alpha_4}*y_5^{\beta_5}$$
$$=0.800^{0.300}*0.900^{0.300}*0.800^{0.200}*0.700^{0.100}*0.800^{0.100}$$
$$=0.818$$

$$T(B,normal\ data,t)=\frac{T_1(B,normal\ data,t)\times T_2(normal\ data)}{1-T_1(B,normal\ data,t)\times(1-T_2(normal\ data))-T_2(normal\ data)\times(1-T_1(B,normal\ data,t))}$$
$$=\frac{0.800\times0.818}{1-0.800\times(1-0.818)-0.818\times(1-0.800)}=0.930$$

Similarly, we can deserve the credibility evaluation results of the data tampered, the data not updated, and the poor data, as shown in Table 3.

Table 1. Data quality attributes and their weight distribution

Attributes	Accuracy	Completeness	Accessibility	Integrity
Weights	0.300	0.300	0.200	0.100

Table 2. Default values of parameters in the simulation experiment

Parameter	Defaults	Description
Time	180(days)	Maximum time interval when data is not updated
θ_t	0.100	Confidence Restoration Coefficient
μ_t	0.100	Time decay coefficient of the consensus degree of on-chain data
λ	0.628	Credit weight coefficient of on-chain members
ξ	0.372	On-chain data quality weight coefficient
α_1	0.300	The weight coefficient of accuracy in critical attributes
α_2	0.300	The weight coefficient of completeness in critical attributes
α_3	0.200	The weight coefficient of accessibility in critical attributes
α_4	0.100	The weight coefficient of credibility in non-critical attributes
β_5	0.100	The weight coefficient of readability in non-critical attributes

Table 3. Results of credibility evaluation of data on the chain after D-S evidence theory

Data type	Credit of members on the chain	y_1	y_2	y_3	y_4	y_5	On-chain data quality	On-chain data credibility
Normal data	0.800	0.800	0.900	0.800	0.700	0.800	0.818	0.930
Data tampered	0.400	0.800	0.900	0.900	0.700	0.700	0.826	0.760
Data not updated	0.489	0.700	0.900	0.800	0.900	0.800	0.806	0.799
Poor data	0.700	0.500	0.500	0.400	0.500	0.600	0.487	0.689

In Table 3, we present four groups of data, namely, normal data group, data tampered group, data not updated group, and poor data group. By observing the evaluation results in Table 3, it can be seen that except the credibility of the normal data group can get a higher score (0.9297), the credibility of the data of other groups after D-S calculation is in the middle range of values. This is because the data itself has various flaws (due to data tampered or data not updated for a long time or poor data). These show that the model is practical and effective, and can reflect the relative differences between the data. Only when the credibility of the members on the chain and the evaluation value of the data quality are generally high, the value of the on-chain data credibility will be high. Therefore, the reasonable thresholds should be set to ensure the credibility of the data in practical applications.

6 Conclusion and Future Work

Aiming at the characteristics of blockchain technology and combining the traditional data quality analysis model, this paper proposes a hierarchical on-chain data credibility measurement model based on D-S evidence theory. The more blockchain nodes and the amount of data are provided, the higher the accuracy of the quantitative evaluation results of credibility given by this model. Finally, the simulation experiment results prove that the quantitative evaluation results given by this model are realistic.

There are several problems that are worth further study. First, we will further expand the attributes in the data quality model and build their measurement models. Secondly, we do not give methods for computing the parameters involved in the models. How to determine the values of these parameters is important future work. Lastly, we will study the application of the model in practice.

7 Acknowledgements

This work was financially supported by the National Key Research and Development Projects of China (2018YFB2101300); National Natural Science Foundation (61906175); the Science and Technology Project of Henan Province (202102210351, 212102210076); the Doctoral Research Fund of Zhengzhou University of Light Industry (2016BSJJ037).

References

[1] A. M. Turri, R. J. Smith, S. W. Kopp, Privacy and RFID Technology: A Review of Regulatory Efforts, *Journal of Consumer Affairs*, Vol. 51, No. 2, pp. 329-354, Summer, 2017.

[2] State Council, *Food Safety Law of the People's Republic of China*, Law Press, 2015.

[3] D. C. Li, W. E. Wong, S. Pan, L. S. Koh, M. Chau, Design Principles and Best Practices of Central Bank Digital Currency, *International Journal of Performability Engineering*, Vol. 17, No. 5, pp. 411-421, May, 2021.

[4] D. Li, W. E. Wong, M. Chau, S. Pan, L. S. Koh, A Survey of NFC Mobile Payment: Challenges and Solutions using Blockchain and Cryptocurrencies, *2020 7th International Conference on Dependable Systems and Their Applications*, Xian, China, 2020, pp. 69-77.

[5] S. Haber, W. S. Stornetta, How to Time-Stamp a Digital Document, *Journal of Cryptology*, Vol. 3, No. 2, pp. 99-111, January, 1991.

[6] H. Massias, X. S. Avila, J.-J. Quisquater, Design of a Secure Timestamping Service with Minimal Trust Requirements, *the 20th Symposium on Information Theory*, Haasrode, Belgium, 1999, pp. 1-8.

[7] A. Back, Hashcash - A Denial of Service Counter Measure, http://www.hashcash.org/hashcash.pdf, September, 2002.

[8] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, *https://bitcoin.org/bitcoin.pdf*, 2008.

[9] P. P. R, A. Tiwari, A. Pandey, S. Krishna, Identifying Video Tampering using Watermarked Blockchain, *International Journal of Performability Engineering*, Vol. 17, No. 8, pp. 722-732, August 2021.

[10] D. C. Li, W. E. Wong, J. C. Guo, A Survey on Blockchain for Enterprise using Hyperledger Fabric and Composer, *2019 6th International Conference on Dependable Systems and Their Applications*, Harbin, China, 2020, pp. 71-80.

[11] G. Y. Pan, Y. Yang, G. Q. Li, J. Wang, W. X. Huang, Blockchain for Collaborative Creation System, *International Journal of Performability Engineering*, Vol. 16, No. 10, pp. 1608-1616, October, 2020.

[12] D. C. Li, W. E. Wong, M. Zhao, Q. Hou, Secure Storage and Access for Task-scheduling Schemes on Consortium Blockchain and Interplanetary File System, *2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion*, Macao, China, 2020, pp. 153-159.

[13] Z.-J. Li, H.-Y. Wu, B. King, Z.-B. Miled, J. Wassick, J. Tazelaar, A Hybrid Blockchain Ledger for Supply Chain Visibility, *17th International Symposium on Parallel and Distributed Computing*, Geneva, Switzerland, 2018, pp. 118-125.

[14] D. Tse, B. W. Zhang, Y. C. Yang, C. L. Cheng, H. R. Mu, Blockchain Application in Food Supply Information Security, *2017 IEEE International Conference on Industrial Engineering and Engineering Management*, Singapore, 2017, pp. 1357-1361.

[15] J.-T. Hao, Y. Sun, H. Luo, A Safe and Efficient Storage Scheme Based on BlockChain and IPFS for Agricultural Products Tracking, *Journal of Computers*, Vol. 29, No. 6, pp. 158-167, December, 2018.

[16] A. Kamilaris, A. Fonts, F. X. Prenafeta-Boldú, The Rise of Blockchain Technology in Agriculture and Food Supply Chains, *Trends in Food Science & Technology*, Vol. 91, pp. 640-652, September, 2019.

[17] Y.-P. Lin, J. R. Petway, J. Anthony, H. Mukhtar, S.-W. Liao, C.-F. Chou, Y.-F. Ho, Blockchain: the Evolutionary Next Step for ICT E-Agriculture, *Environments*, Vol. 4, No. 3, pp. 1-13, September, 2017.

[18] A. Shahid, A. Almogren, N. Javaid, F. A. Al-Zahrani, M. Zuair, M. Alam, Blockchain-based Agri-Food Supply Chain: A Complete Solution, *IEEE Access*, Vol. 8, pp. 69230-69243, April, 2020.

[19] S. S. Kamble, A. Gunasekaran, R. Sharma, Modeling the Blockchain Enabled Traceability in Agriculture

Supply Chain, *International Journal of Information Management*, Vol. 52, Article No. 101967, June, 2020.
[20] D. McGilvray, *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information*, Morgan Kaufmann Publishers Inc., 2008.
[21] F. Sidi, P. H. S. Panahy, L. S. Affendey, M. A. Jabar, H. Ibrahim, A. Mustapha, Data Quality: a Survey of Data Quality Dimensions, *2012 International Conference on Information Retrieval and Knowledge Management*, Kuala Lumpur, Malaysia, 2012, pp. 300-304.
[22] P. Glowalla, P. Balazy, D. Basten, A. Sunyaev, Process-Driven Data Quality Management-An Application of the Combined Conceptual Life Cycle Model, *2014 47th Hawaii International Conference on System Sciences*, Waikoloa, HI, USA, 2014, pp. 4700-4709.
[23] C. Batini, C. Cappiello, C. Francalanci, A. Maurino, Methodologies for Data Quality Assessment and Improvement, *ACM Computing Surveys*, Vol. 41, No. 3, pp. 1-52, July 2009.
[24] Y. X. Chen, H. W. Tao, *Software Credibility Measurement Evaluation and Enhancement Specifications*, Science Press, 2019.
[25] A. Dempster, Upper and Lower Probabilities Induced by a Multivalued Mapping, *Annals of Mathematical Statistics*, Vol. 38, No. 2, pp. 325-339, April, 1967.

Biographies

Hongwei Tao received the Ph.D. degree in computer applications technology from East China Normal University in 2011. Currently, he is an Associate Professor with Zhengzhou University of Light Industry, China. He has published more than 30 papers in various journals and conferences. His current research interests include software trustworthiness measurement, big data analysis and formal methods.

Yinghui Hu received the B.S. degree from Zhengzhou University of Light Industry, China in 2018. He is currently pursuing the M.S. degree in Zhengzhou University of Light Industry, China. His current research interests include blockchain technology and artificial intelligence.

Hui Li received the Ph.D. in risk management and accurate calculation from Renmin University of China in 2014. Currently, he is a Lecturer with Qingdao Agriculture University of Faculty of Economics. He has published more than 10 papers in various journals and conferences, and are leading many projects. His current research interests include finance measurement, big data analysis and statistical methods.

Deqiang Fan received the B.S. degree in engineering management form East China University of Science and Technology in 2006. Currently, he is chief economic manager of HeNan New Landmark Construction Engineering Limited Company. His current research interests include intelligent building, the industrialization of construction and the prefabricated build.

Haoran Chen received the Ph.D. degree in computer science and technology from Beijing University of Technology in 2019. He is currently working at Zhengzhou University of Light Industry, China. He has published 7 papers in several journals and conferences, His current research interests include artificial intelligence, machine learning, pattern recognition and manifold optimization.