Hindawi
Security and Communication Networks
Volume 2022, Article ID 5842371, 12 pages
https://doi.org/10.1155/2022/5842371

Research Article

Design of a Blockchain-Based Traceability System with a Privacy-Preserving Scheme of Zero-Knowledge Proof

Yudai Xue and Jinsong Wang

School of Computer Science and Engineering, Tianjin University of Technology, Tianjin, China

Correspondence should be addressed to Jinsong Wang; jswang@tjut.edu.cn

Received 6 March 2022; Accepted 3 June 2022; Published 29 June 2022

Academic Editor: Jiewu Leng

With the fast development of the industrial Internet, its interconnectivity poses new challenges for the cooperation of industrial entities. Cooperation among these entities is built on trust, and trust is based on high-quality industrial products at reasonable prices. A traceability system can play an essential role in objectively reflecting the production process and promoting this trust. However, traditional traceability systems often have data privacy issues. Because traceability data are collected or generated during the production process (namely, production-related data), they could be considered privacy data. Several researchers have introduced privacy protection schemes into the traceability system, such as authentication or encryption. Nevertheless, when a privacy protection scheme is established, the original data are disclosed to the legal user of the system, but the data may still be leaked intentionally or unintentionally. Except for data privacy issues, a traditional traceability system can be vulnerable to network attacks, data unavailability, and reliability issues. The authors conducted a study to overcome these shortcomings, and this paper reports the results. We built a traceability prototype system using a blockchain protocol and a zero-knowledge proof method. First, we built a blockchain to record key production process data, aiming to maintain data reliability and availability. Second, through an analysis of traceability purpose using production knowledge, the traceability purpose could be divided into multiple provable statements. By introducing privacy protection through a zero-knowledge proof, the traceability process was converted to proving relative statements. Finally, the statements were validated by a smart contract that provided openness and reliability during the traceability process. Analysis has shown that our approach could meet the requirements for high security and privacy. In addition, the paper also discusses the calculation cost of the traceability process to show our work's viability. The traceability system described in this paper creates new possibilities for constructing a healthy and reliable trust relationship between production entities to provide further support in the development of the industrial Internet.

1. Introduction

With the application of information technology (IT) to traditional industrial production, production power has significantly increased, and senior automation and information technology have optimized the production process. However, as the scale of production expands, the industrial production model may be overwhelmed by high production levels, the bullwhip effect, or biased pricing [1]. These issues impair the trust and cooperation between industry entities [2]. To break the production limit, Industry 4.0 [3] and the industrial Internet [4] were introduced by Germany in 2010 and General Electric in 2012. The primary purpose of this structure is to connect people, data, and machines with an open and globalized network and to achieve a high degree of integration of industrial systems with computing, analysis, and IT systems [5]. In particular, the industrial Internet may create a data corridor between each element in production scenarios. By integrating the traditional manufacturing technology with big data analysis, artificial intelligence, and advanced semiconductor technology, industrial Internet could reform the entire production and cooperation model [6].

In traditional industrial production, cooperation among entities depends on the supply chain [7], and the traceability data are one of the key parts of the supply chain system to ensure the high efficiency and stable operation. Traceability

data are also an important constituent for improving the quality of industrial products [8] and optimizing the supply chain and the production process [9]. As an important part of the supply chain, a traceability system can significantly affect product quality control, order management, and production service. However, with the continuous expansion of the supply chain and development of industrial Internet, the traceability system may face challenges due to the requirement of interconnection and open cooperation between industrial entities [10]. Current traceability system-related data are not sensitive. However, some field data related to the production process can be unsafe for a company to publish due to industrial confidentiality. Against an industrial background, sensitive production data can objectively reflect the quality and even advanced technology of industrial products, so traceability requirements for these types of sensitive data do exist. Furthermore, tracing these production data can promote and encourage interindustry cooperation and interactions. This new type of cooperation and production may transform the dominant industrial model.

For regular scenarios, traceability can be treated as a process that satisfies a particular purpose related to demand. In this situation, the initiator of traceability should have a clear purpose, such as locating quality problems in the production process [11] or verifying whether a particular process in the production process satisfies required standards or specifications. No matter the purpose of the traceability, it is ultimately up to the traceability initializer to judge the relevant data obtained from a traceability system. However, acquirers of traceability data might lack knowledge about production in an actual traceability process. After the related data are received, a third party is needed to interpret the result using traceability data, which may eventually cause production data to leak. Therefore, the traceability process for sensitive production-related data should avoid data transmission.

To achieve the above, this study adopts a privacy protection mechanism based on a zero-knowledge proof and realizes the traceability of the target with no need for the traceability data owner to provide any original data. In addition, the traceability system still needs to solve the problem of original data availability and reliability, and the proof should be open and without the possibility of repudiation. To satisfy the above requirements, this paper introduces blockchain technology into the design of traceability system, which could provide features that are tamper-resistant, unable to be repudiated, and open to supervision [12]. To design a complete system, this study made the following assumptions:

(1) Raw industrial production data are stored in the respective production domains, while the related data digest is stored in the blockchain.

(2) The traceability process was initialized by the traceability data acquirer with a clear purpose or an expected traceability result.

(3) There is a correlation between the traceability purpose and the industrial traceable production data.

Under the above premise, this paper introduces a zero-knowledge proof for raw data production. Through a purpose analysis of traceability data acquirer, the production process was converted into multiple statements and adopted a zero-knowledge proof engine generating a validator to prove those statements. Finally, through publishing a smart contract, the traceability process is completed in an open and fair manner. The innovations of this paper are as follows:

(1) An abstraction of the industrial production process into several traceability features according to their traceability is developed.

(2) A privacy-preserving traceability system architecture for production data is constructed and the algorithm flow involved in the architecture is explained.

(3) Availability and security issues are discussed through a comparative analysis.

2. Background Knowledge

2.1. Blockchain and Distributed Ledgers. A blockchain is a type of distributed ledger system designed on a cryptography algorithm, peer-to-peer (P2P) network, and distributed consensus algorithm [13]. Blockchain has attracted much attention since its creation. Many researchers have shown increasing interest in the application of blockchain. Generally speaking, blockchain can be divided into two categories: unauthorized blockchain and authorized blockchain [14]. The former category is usually used to build payment systems instead of centralized banks, such as Bitcoin or Ethereum [15, 16], and the latter is designed for specific application scenarios such as medical, agrifood, or other fields. No matter what type of blockchain is being used, it can help to build trust between participants in an exchange. Authorized blockchain supported with smart contracts could be applied to financial, medical, and logistics scenarios [17]. Moreover, its features of tamper-resistance and non-repudiation may provide highly reliable data that could be the basis for industrial entities creating cooperative relationships. Especially in the research area of combining blockchain with industrial Internet or industry 4.0, the security and trust features of blockchain [18] may help industrial entities create healthy cooperative relationships and promote the production level.

The main idea of the traceability system designed in this paper is to trace the privacy data generated in the production process. In industrial production scenarios, such as the industrial Internet of things or the industrial Internet, the data generated from billions of sensors, controllers, and data collectors makes it possible for entities to make production more intelligent, optimize production plans, and realize cooperative production [19]. To create a reliable traceability system, the first step is to guarantee the reliability and availability of production data [20]. In traditional centralized traceability approaches, there are many potential information security issues, including denial-of-service attacks, spoofing attacks, and data leaking and tampering, while a blockchain-based approach may be immune to the security issues above and make the data both tamper-resistant and highly available [21]. In this situation, the circulation of data is treated as a transaction, and data digests can be recorded

in a transaction for confirmation by all participants. Supported by blockchain, a traceability system able to fully record industrial production could be created.

Nevertheless, the openness of blockchain may also create privacy issues [22]. Production-related data may be tightly bonded with industrial secrets and sensitive data, making it impossible for entities to share their production data for traceability. Therefore, a privacy-preserving scheme should be deployed.

2.2. Zero-Knowledge Proofs. Zero-knowledge proofs can enable one subject to verify the correctness of a statement put forward by another subject without involving any raw data or relying on a third party [23]. Therefore, zero-knowledge proofs can be used as effective privacy protection mechanisms. There are two leading roles involved in the zero-knowledge proof process. The first is the prover, who declares a statement and generates a proof with raw data. The second role is the verifier, or the proof receiver, who has the ability to verify the proof. Zero-knowledge proofs are widely used in privacy-preserving schemes due to their completeness, soundness, and zero-knowledge [24]. Completeness means the statement can be verified by the prover and convince the verifier of its veracity. Its soundness provides an environment in which the prover cannot cheat the verifier with a false proof. Zero-knowledge ensures that the raw data is never revealed to the public. The prover can always maintain their ownership of the raw data during the proving and verifying processes to protect their privacy.

In the real-world usage of zero-knowledge proof schemes, a toolkit based on zero-knowledge succinct non-interactive arguments of knowledge (zkSNARK) is introduced to build corresponding systems [25]. zkSNARK allows the prover to prove its statement with low process complexity using a simple message. Therefore, the toolkit based on zkSNARK is widely used in the design of blockchain-based applications [26]. According to the application scenario, a zkSNARK-based toolkit offers a flexible and effective way to create a smart contract for automatic verification and generate the proof with raw data.

In this situation, the application scenarios of zero-knowledge proof are significantly expanded and provide the possibility for the implementation of traceability in this paper.

2.3. Related Work. Research on the combination of blockchain and traceability systems has shown that the data recorded in a blockchain can provide reliable support for data traceability so long as production data can be stored with high reliability. Previous work has successfully connected the traceability process with the data produced during production [27]. The traceability of production data is beneficial for tracking production drawbacks and raising production quality [28], and, in the research area of traceability, Chen et al. [29] demonstrated the relationship between quality control and traceability and then designed a quality control model based on traceability. On this basis, Tsai and Wang et al. [30] then designed a cooperative production method based on the production data to improve and optimize the production process. By importing blockchain into traceability system design, more researchers have concentrated on the design of blockchain-based decentralized traceability systems. Helo et al. [31] designed a high-performance traceability model for the supply chains based on blockchain, Radio Frequency Identification (RFID), and Internet of Things (IoT) technology. Zhu et al. [32] optimized the supply chain by using blockchain-based traceability system to trace production processes and coordination. Xiao et al. [33] designed a traceability model for the agrifood industry, providing a safe and traceable environment for food production quality control and anti-counterfeiting. Tarun [34], based on blockchain, constructed a traceability system for textile manufacturing and improved production efficiency. Uddin [35] built a blockchain-based traceability framework for the pharmacy industry that provided reliability verification for the circulation of medicine. Patelli et al. [36] built a traceability system for the supply chain management of food industry based on blockchain, which is immune to several network attacks compared to the traditional traceability mechanism. The above studies have proven that blockchain technology can be effectively integrated with traceability mechanisms to improve production efficiency and product quality to ensure data reliability. Except for studies on the reliability of traceability data, privacy-preserving schemes for traceability data were also discussed by researchers. Yang [37] used RFID encryption to provide production data privacy during data collection. Wang [38] treated the traceability process as transactions in the blockchain and divided the traceability process into three actions: demanding, pricing, and trading.

Privacy-preserving mechanisms have also been introduced to avoid privacy leakage issues formed by the openness of blockchain. The above research shows that privacy preservation methods mainly depend on access control, identity authorization, or data encryption. Although the traceability process is protected, raw traceability data can still be leaked by the traceability data receiver.

In this paper, a zero-knowledge proof is used to protect the raw traceability data. Zero-knowledge proofs can provide proof for satisfying specified conditions in a specific scene without disclosing any private information. Currently, zero-knowledge proofs are widely used in digital currency. Zcash realized a privacy-protected digital currency system by applying a zero-knowledge proof. In addition, Eberhardt [39] combined a zero-knowledge proof with an Ethereum smart contract by constructing ZoKrates, realizing a zero-knowledge proof mode of offline computing and online verification, thus expanding the possible applications of zero-knowledge proofs. Based on this, Westerkamp [40] designed a side chain proof mechanism using ZoKrates. In addition to the field of digital currency, Ibrahem [41] applied zero-knowledge proofs and homomorphic encryption to an anonymous voting system. Rasheed et al. [42] realized an anonymity authentication method based on the premise of protecting user privacy by applying a zero-knowledge proof in the area of Internet of vehicle. Jeong et al. [43] applied smart contracts and zero-knowledge proofs in online real

estate transactions to provide a transaction process with a privacy protection mechanism. Qi et al. [44] applied a zero-knowledge proof to the auto insurance industry, achieving an effective insurance evaluation method based on usage habits while preserving privacy. Umar et al. [45] combined zero-knowledge proofs with wireless body sensors to effectively avoid privacy leakage caused by malicious attackers monitoring communication channels. Huang et al. [46] proposed an auditable information-sharing mechanism for the industrial Internet based on a zero-knowledge proof mechanism, which effectively avoided the leakage of sensitive information into the industrial environment. The above research results show that zero-knowledge proofs can provide a proof process with a privacy protection mechanism for different fields according to their needs, and so it is feasible and beneficial to attempt to apply it to the privacy protection process of production data traceability.

3. System Architecture

3.1. Architecture Review. Based on industrial production data, as industrial production entities record production data digests in the blockchain, these production data can objectively describe the production process and reflect the technologies of production and the flow of the production process. Traceability processes are often used to show the high quality or advanced technology of production [47]. In this situation, the traceability process needs to be open and fair. However, the traceability data acquirers may lack production-related knowledge and cannot judge whether acquired data could reach their purpose. Therefore, the current solution relies on a trusted third party, which means that the traceability data acquirers need to inform their purposes to a trusted third party, entrust it as an agent to receive the traceability data, and make a final judgement [48]. However, this solution still exposes the production data to a third-party verifier. At the same time, the verification process would not be open and transparent. To establish an open, fair, and privacy-preserving traceability mechanism, this paper designed a new traceability system based on a zero-knowledge proof combined with a smart contract and blockchain, which is called a zero-knowledge-based traceability system (ZKTS), to create a privacy data traceability method independent from any third party in order to protect the privacy of producers. The architecture diagram of the traceability system is shown in Figure 1.

The traceability system in this paper contains three layers. The first is the physical data layer, which contains the data produced by each independent entity according to the actual production, the physical data generated by the transportation and sales process, and the raw data saved in the database and constructed by an independent production entity. At the same time, the digest of raw data was published in an authorized blockchain system called a data chain. Data chain is open for all certificated users to access the data digest of raw data. The second is the data privacy layer, composed of a zero-knowledge engine and its related external interfaces. The data privacy layer is mainly responsible for receiving the traceability features and the related data digest from the upper layer, building the smart contract, and interacting with the data provider to generate the proof for traceability using a zero-knowledge proof engine. The data privacy layer is responsible for the core function of the traceability system, privacy preservation. The third layer is the application layer; in this paper, we define the purpose of the traceability data acquirer as the traceability purpose, and the application layer maintains the authentication process and the traceability purpose analysis process or the traceability feature-generating algorithm. This algorithm is used to analyze the primary purpose of traceability and generate related features and could satisfy the traceability purpose. As shown in Figure 1, the entire system contains two relatively independent blockchains. One is used to ensure the integrity and reliability of relevant production data, and the other is used for traceability verification in public scenarios. The two blockchains cannot interact directly but can be accessed through authentication with supervision. The specific functions of each layer in the traceability model are as follows.

3.1.1. Physical Data Layer. The physical data layer includes sensors, controllers, data collection devices, and other production-related devices. In the physical data layer, production data can be mapped to devices using a unique label through RFID or other technologies. The physical data layer collects the data generated by those production devices, and these collected data can be used to generate traceability features. In this study, we assumed that no fake data had been created in the physical data layer and that the data digest would be published in a particular blockchain system (data chain).

3.1.2. Data Privacy Layer. The data privacy layer is the key component ensuring the protection of privacy in the traceability system. The data privacy layer has a data extraction and privacy processing engine. First, the data privacy layer collects the data digest of production from the data chain. According to the analysis of the purpose of the traceability data acquirer, the privacy processing engine then generates a smart contract with need-to-proof issues. A witness is then generated to interact with the owner of the raw production data to generate the proof.

The processing of the data privacy layer involves the privacy information of production, so the process is offline. After the proof is generated, the smart contract, related proof, and traceability features are disclosed to verify whether traceability is achieved.

3.1.3. Traceability Application Layer. The traceability application layer directly interacts with the traceability data acquirer and production data owner. The primary function of the application layer is to provide a traceability interface for both sides with an authentication mechanism and also to provide a platform for mutual traceability negotiation. The traceability application layer was established by industrial entities and traceability-related individuals or entities. The traceability application was developed with a smart contract

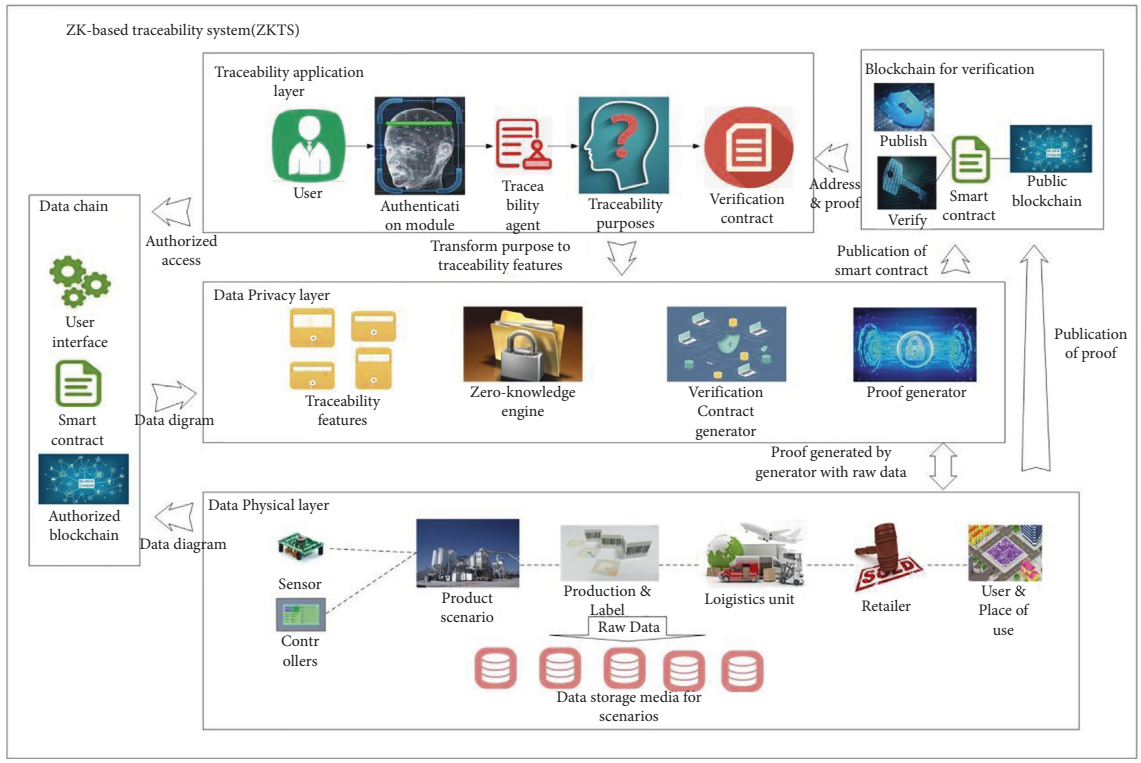

Figure 1: Overview of the ZKTS system.

on a public or light-authenticated blockchain system. The traceability application layer can publish the smart contract generated by the data privacy layer and provide a verification interface for the proof. A blockchain-based application layer is an essential part of the traceability system, allowing the system to be both fair and transparent.

3.2. Participant Design. In the traceability system we designed, given the ownership of the traceability data, there were three roles: the initializer of the traceability process, called the traceability data acquirer (TDA); the traceability data owner, called the traceability data provider (TDP); and a third party, called the traceability agent (TA).

The TDP was the original holder and provider of the production data to be traced. In production-related traceability, the data has production technology privacy that needs to be protected during the production and circulation of the relevant products. The TDA was the initializer of the tracing process and could be a partner or a consumer with either a cooperative or a transaction-based relationship with the TDP. The TDA initiated the tracing process with a clear purpose such as judgement of production quality, making it a sign of traceability completion. The TA introduced in this paper is a new role. The TA held the production knowledge of the TDP, which means that the TA could analyze the primary purposes of the TDA. According to the features generated by the TA, smart contracts with related need-to-proof issues were created and transferred to TDP. According to those need-to-proof issues, the TDP generates the proof with raw production data and publishes it to the public. Since the behavior of the TA did not involve the privacy of both sides of the traceability process, the whole traceability process can be conducted transparently to ensure the openness and effectiveness.

3.3. Traceability Process Design. The traceability process designed in this study contains four phases: the authentication phase, preprocessing phase, construction phase, and verification phase. In different stages, different roles may execute corresponding workflows, and a complete process of our approach together with the relationship between roles and workflows is shown in Figure 2.

A typical traceability process is shown in Figure 2. In the authentication phase, an authentication scheme is implemented based on the data recorded in the data chain. A traceability request is then sent from the TDA to the TA. In the preprocessing phase, the TA extracts the tracing proposal from the request, generates the traceability features, and transfers them to the TDP. In the construction phase, the TDP receives the traceability features and generates the proof using the raw production data. Meanwhile, a smart contract is published by the TA for verification. Finally, in the verification phase, the TA receives the generated proof and passes it to the TDA. After the complete process, the TDA determines whether traceability has been achieved. A more detailed explanation of the algorithm and its process is fully explained in what follows.

3.3.1. Authentication Phase. In this phase, due to the openness of the traceability system, the TDA and TDP both

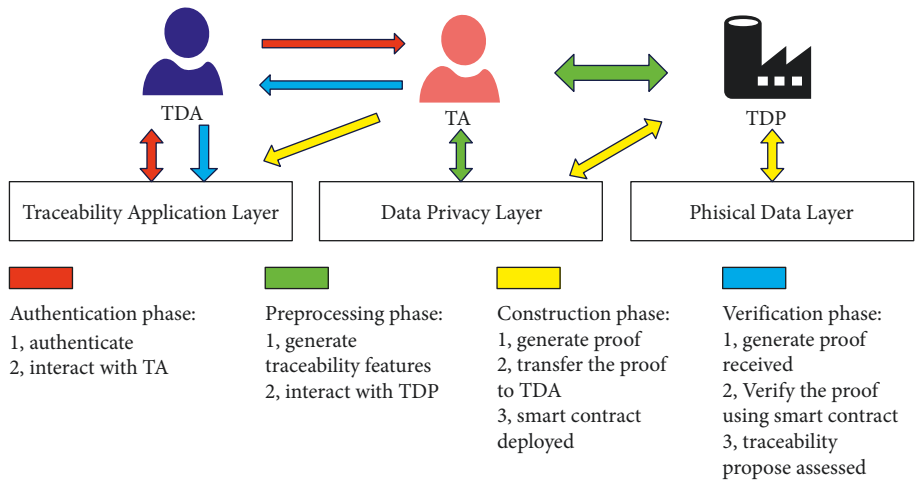

Figure 2: Overview of the entire traceability process.

have the privilege of accessing the data digest stored in the data chain, which means the TDA and TDP each have a public and private key pair: PKa/SKa and PKp/SKp, respectively. Furthermore, in the process of production trading, a relationship between the TDA and TDP is established and the related transaction data are also recorded in the data chain; these data could be used to generate ProductData (PD) using the following formula:

$$\text{ProductData} = \text{PID}|\text{TxTS}|\text{TxHash}|\text{random}(r). \quad (1)$$

The ProductData contains relevant information about production trading between the TDA and TDP, including product identification (PID), transaction timestamp (TxTS), transaction digest (TxHash), and a random number (r). The ProductData are generated in the process of product circulation. According to the relevant data, SKp and PKa are used to encrypt and sign PD, while ProductSign (PS) generated using formula (2) is used for authority authentication of TDA:

$$\text{ProductSign} = \text{sign}(\text{sign}(\text{ProductData}, \text{SKp}), \text{PKa})|\text{PID}. \quad (2)$$

The PS is then passed to the TDA, who can verify the product according to SKa and PKp. After PD is obtained, the TDA calculates the PD digest, simultaneously generates an authentication request, and then submits the request to the TDP to prove the traceability permission of the corresponding product. This completes the authentication process between the TDA and TDP. The authentication algorithm is shown in Algorithm 1.

3.3.2. Preprocessing Phase. After the authentication phase, the system workflow enters the preprocessing phase. In this phase, the traceability purpose of the TDA is analyzed by the traceability feature generating algorithm. The core of the algorithm is based on public industrial production standards or key technical knowledge. According to these standards, knowledge, and related production data, traceability features can be generated.

As the production process progressed, data corresponding to the production process is generated in industrial production. Here, we define all the production processes as P, and we obtain the following equation:

$$P = P_1 \cup P_2 \cup_3 \cup \cdots \cup P_n, \quad (3)$$

where P_I represents one of the production steps in industrial production process, and, for any I, we use the following formula:

$$\{D_1, D_2, \ldots, D_{k-1}, D_k\} \longrightarrow P_i, \quad (4)$$

D_i Refers to the relevant data generated in a single production process; that is, there is a mapping relationship between the production process and the data, and each data point generated in the production process can be defined by the following formula:

$$D_j \in \{\text{bool}, \text{num}, \text{hash}\}. \quad (5)$$

In other words, the production process data can be defined as a Boolean (bool), numeric (num), or hash type in a production record. Boolean data show the state of industrial devices, such as controllers or switches. Numeric data are used to record the data generated by the production devices like sensors. Hash data are used in industrial production to record relevant signatures. In this situation, a production process can be defined by the following equation:

$$P_i = D_1^i \cup D_2^i \cup \cdots \cup D_k^I, \quad (6)$$

where D_k^I is defined as the kth data generated in the ith production process. Therefore, for a single production product, its entire production process can be defined by the following equation:

$$\text{Product} = \left\{D_1^1 \cup D_2^1 \cup \cdots \cup D_m^1\right\} \cup \left\{D_1^2 \cup D_2^2 \cup \cdots \cup D_n^2\right\} \cup \cdots \cup \left\{D_1^k \cup D_2^k \cup \cdots \cup D_p^k\right\}. \quad (7)$$

On this basis, each industrial production process can be mapped to a set of data that holds all the data generated by

```
    Input: ID, PKt, SKte, Tx
    Output: AuthResult
(1) ProductData = PID|TxTS|TxHash|random(r)
(2) ProductSign = sign(sign(ProductData,SKp),PKa)|TxHash
(3) TDP.send(ProductSign)
(4) recvData = TDA.recv()
(5) ProductData = decrypt(recvData,SKa).decrypt(recvData,PKp)
(6) if verify(PID):
(7)     authToken = hash(ProductData)
(8) else:
(9)     deny()
(10)    TDA.send(authToken|PID)
(11)    authToken = TDP.recv()
(12)    if authToken = = hash(ProductData)
(13)        IndentityConfirmed()
(14)    Else
(15)        deny
```

Algorithm 1: Authentication algorithm.

the production process. Considering the production data and the correlation of the traceability process, a function, dataParser, was introduced to filter the key data out of the set. The key data are then generated using the following equation:

$$\text{KeyData} = \text{dataParser}(\text{Product}). \tag{8}$$

KeyData obtained through this process can be further processed and combined with relevant knowledge of the production field to generate traceability features. In this situation, the traceability purpose is defined as TP, and the traceability feature can be generated using the following equation:

$$\text{TraceabilityFeature} = \text{parse}(\text{KeyData}, \text{knowledge}). \tag{9}$$

In the above equation, knowledge means the production knowledge that can be obtained by the relevant production experts or by a public production standard. TraceabilityFeature is generated as a producer of zero-knowledge proof. On this basis, the TA will use TraceabilityFeature to generate a smart contract and also transfer the TraceabilityFeature to TDP for generating the proof with raw data.

In addition to traceability features generated in the preprocessing phase, high-availability production data are provided by the TDP. Sensors generate raw production data during production, transportation, trading, and other activities, among which different production entities are distributed. These data can be marked as rawData, and the rawData digest is marked as HashData. HashData and the identity of the data source form a transaction record, Tx. Finally, Tx is published to the data chain, which creates a relationship between production data and transactions. The on-chain data digest ensures the tamper-resistant and antirepudiation characteristics of the production data. The specific algorithm of data preprocessing is shown in Algorithm 2.

The data are first collected by a data collection device in the data physical layer and are used to extract the traceability

```
    Input: Data
    Output: Tx
(1) Data = [sensor, collector,controller, etc].collect()
(2) ID = [sensor, collector,controller, etc].PID
(3) Hashdata = hash(Data)
(4) Tx = TXgenerator(Hashdata|ID)
(5) Tx.submit()
```

Algorithm 2: Data preprocessing algorithm.

features with production knowledge. The blockchain-based data record model has been previously discussed in the literature [34, 37]. In our study, we only refer to those conclusions. The high-availability record of the production data provides a basis for the privacy protection approach to be adopted in the later construction phase.

3.3.3. Construction Phase. After the two phases above are finished, the traceability features are provided with the available production data. The construction phase may import the zero-knowledge engine to generate related smart contracts and the proof. In the construction phase, the TA uses the traceability features generated by the preprocessing phase to create need-to-proof issues, which is called the witness, and then transfers it to the TDP. The TDP receives and generates the proof using a witness and the related raw data. After the proof is generated, the TDP submits the proof to the TA or the public, and then the TDA receives the data and verifies the proof.

At first, the generation of witness should be discussed. We define the witness as verifyKey, and three main issues should be proven. First, the available data must prove that the data that the TDP used to generate the proof are the same as those recorded in the data chain. Second, those data must be from the production being traced. Third, the traceability feature proof must show that the proof of data is satisfied

Input: PID, TraceabilityFeature(TF), offlineVolume
Output: witness
(1) Tx = Datachain.search(PID)
(2) Hash = Tx.DataHash
(3) Data = offlineVolume. Find ("PID")
(4) For each in TF
(5) verifyKey.append(Hash(Data) = = Tx.hashData)
(6) verifyKey.append(PID = = Tx.PID)
(7) verifyKey.append(Data.satify(TF))
(8) return verifyKey

Algorithm 3: VerifyKey generation algorithm.

among the traceability features. This ensures the correctness of the traceability process. The verifyKey generation algorithm is shown in Algorithm 3.

After generating verifyKey, a zero-knowledge engine (ZKe) is introduced to the phase. The TA first uses ZKe to compile and set up with verifyKey to generate specific keys and transfer them to the TDP for creating a witness and proof. Simultaneously, the TA generates the smart contract using the genContract() function to provide an interaction interface for the traceability application layer. The TDP receives the keys to generating the witness using the related raw production data and then creates a proof that can be published to the public. The proof and contract generation algorithm is shown in Algorithm 4.

In the above algorithm, the TA publishes a smart contract address to the public to verify the proof. Through the TA, the TDP can first receive the traceability features and verification keys and then generate and publish the proof using raw production data. Meanwhile, the TDA can verify the proof through the smart contract in an open manner and finally reach traceability purpose in the verification phase.

3.3.4. Verification Phase. After the completion of the above phases, the TA and TDP can provide the contract address, proof, and traceability features. The TDA can analyze the traceability features with common knowledge and judge the expected result. If the TDP fails to generate a related proof, the traceability failed, and the TDA and TA might negotiate new traceability features and restart the traceability process.

It was worth noting that traceability is based on traceability features that are generated with expert or public knowledge. Simultaneously, the related smart contract was permanently deployed in the blockchain, which means that the smart contract with the related proof could be reused to simplify the traceability process. The algorithm designed in the verification stage is shown in Algorithm 5.

In the verification phase, according to the judgement of traceability features, the TDA, TDP, and TA may complete the traceability process in an open and privacy-preserving way. The traceability system for this paper could be effectively applied to typical traceability scenarios such as anti-counterfeiting verification [49], standard execution proofs [50], or abnormal investigations [51]. Furthermore, in the

Input: verifyKey, PID, TF
Output: Proof, contractAddr
(1) TA.compile(verifyKey)
(2) TA.setup()
(3) SC = TA.genContract()
(4) verifyKey,TF,PID = TDP.recv()
(5) rawData = TDP.getRawDataByID()
(6) witness = ZK.genWitness(rawData, verificationKey)
(7) proof = Verify.genProof(Witness)
(8) TDP.send(proof,TA)
(9) TA.publish(proof|contractAddr)

Algorithm 4: Proof and contract generation algorithm.

Input: proof, TF
Output: TP, result
(1) TDA.subscribe(TA,TDP)
(2) TraceabilityResult = TDA.verify(ProofAddr, Proof| ProofFeature)
(3) If TraceabilityResult = = satisfied:
(4) Finish tracing
(5) If TraceabilityResult = = not satisfied
(6) Restart traceability process()
(7) Return tracing failed

Algorithm 5: Verification algorithm.

above relevant traceability process, the TA could be industrial entities or even production experts. Those TA may create a competitive environment for traceability feature generation. As more new technologies have been imported into the production process, a competitive relationship between TAs may be beneficial in raising the availability of the traceability system.

4. Analysis with Discussion

4.1. Security Analysis. Security issues in network attacks and privacy protection are discussed in this section. For network attacks, in the traditional traceability system, malicious intruders may launch Distributed Denial of Service (DDoS) attack or Advanced Persistent Threat (APT) attacks in the centralized server [52]. Furthermore, the traceability data may be tampered with to destroy the traceability system's authority. However, the blockchain-based traceability system designed in this paper is immune to DDoS attacks and APT attacks [53]. All data digests stored in the blockchain for privacy protection are mapped with raw production data. With the importing of the zero-knowledge proof scheme, the traceability process can be transferred into a proof of related traceability features. The system described in this paper can avoid transferring the raw production data between any entities or individuals and so finally realize the traceability of privacy data. With the traceability system based on smart contracts, the entire traceability process could be monitored by the public, enhancing the traceability compliance and openness of the process. The system designed in this paper

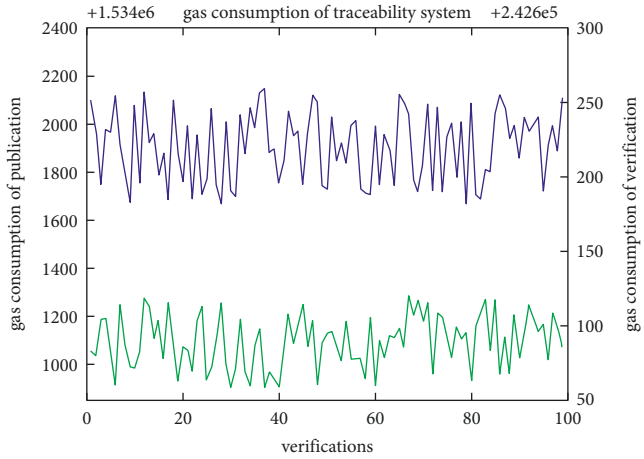

Figure 3: Gas consumption of publication and verification.

can be used to promote the industrial Internet, help it achieve a broader range of joint production, and provide an essential basis for trust [54].

Despite the network attack launched by malicious intruders, the blockchain itself may also have security issues such as transaction-related attacks (e.g., double spending attack), consensus failure, and smart contract-oriented attacks. In our work, we deployed the whole application on Ethereum environment to realize the availability of traceability process and assumed that the consensus process, smart contract execution, or other application process logic are operating without malicious behavior. Further research was deeply discussed in related works [55]; we just make the assumption above in order to reach the availability of traceability system.

4.2. Available Analysis. In this paper, zkSNARK toolkit was introduced to form a privacy-preserving scheme. According to the generated traceability features, a smart contract could be published in the Ethereum blockchain. Our experimental environment included an i7-6700 CPU and 8 GB of memory; considering the highly available and widely used blockchain system, a widely used IDE with Ethereum testnet [56] was also used. We designed the test with 1 to 100 traceability features and generated a related proof and smart contract. However, the calculation of Ethereum is not free; the verification and the publication of a smart contract could require a computing fee (gas). In this situation, gas could be treated as a cost of traceability process, and the gas consumption of contract publication and verification is shown in Figure 3.

As shown in the figure above, the gas consumption does not significantly increase as the number of traceability features increases. The gas consumptions of smart contract publication and validation are approximately 1535000 and 242800. With the current price of the Ethereum, the costs are approximately 0.15 ETH and 0.02 ETH. As shown above, the publication and verification may increase the cost of traceability. In contrast, in traceability scenarios, the gas consumption could be optimized, such as through the production of a continuous linear process that could be implemented in a contract for multiple data validations and could finally reduce the gas consumption of publication and verification. At the same time, the system designed in this paper is built in the Ethereum testnet environment, and the traceability system in this paper can also be deployed into mainstream authorized blockchain systems, such as Hyperledger. Through corresponding smart contracts, the extra cost of publication and verification may effectively be avoided.

4.3. Comparative Analysis. In the comparative analysis, this section compares our traceability system with the traceability systems constructed in related studies to demonstrate the advantages of our approach. In this paper, we choose three typical models of traceability systems: a traceability system designed without blockchain [11] (Centralized), a blockchain-based traceability system without a privacy-preserving scheme [30–35], and a blockchain-based traceability system with a privacy-preserving scheme [36, 37]. Although the traceability systems we choose to compare were designed for different purposes and industrial environments, the issues in each typical model are common; therefore, these traceability system models were selected for comparative analysis of data reliability, traceability target, privacy protection scheme, attack resistance, and cost of traceability. The analysis results are shown in Table 1.

Unlike other related traceability systems, this paper constructed a traceability system using a decentralized architecture based on blockchain; on this basis, unlike a centralized traceability system, data reliability and availability could be fully guaranteed. At the same time, our approach can defend against DDoS attacks, and, due to the privacy protection scheme of zero-knowledge proofs together with authentication, our approach realizes a traceability of privacy data in industrial production that is superior to other decentralized traceability systems and could effectively build trust relationships between industrial entities. For the comparison of traceability cost, a centralized traceability system needs to spend much to build and maintain a traceability system, that is, the traceability system

Table 1: Comparative of traceability systems.

	Traceability target	Attack resistance	Privacy protection	Data reliability	Traceability cost
Our approach	Privacy data/ nonprivacy data	√	Certification + zero-knowledge proof	√	According to the contract consumption of calculation
Centralized	Nonprivacy data	x	Certification	x	Determined by the traceability system creator
Blockchain-based without privacy-preserving	Nonprivacy data	√	None	√	According to the contract consumption of calculation
Blockchain-based with privacy-preserving	Privacy data	√	Certification	√	According to the contract consumption of calculation

manager may have the right to fix the price of traceability service. It may influence the cost of traceability system users. In contrast, in our approach, pricing was determined by the calculation cost of the smart contract. The TA was introduced to make the price of traceability more flexible and open. The TDA and TDP would also benefit from the traceability process.

5. Conclusions and Future Work

With the development of industrial production, the industrial Internet may create many more opportunities for industrial entities of all sizes. As a key part of building industrial cooperation between entities, the traceability system plays a significant role in the development of industrial Internet. The traceability system designed in this paper achieves traceability for privacy production data, which makes it possible to objectively judge the quality of products in order to raise production quality or optimize supply chain structure. On the other hand, the privacy-preserving scheme designed in this paper raises the willingness of parties to share data, which may take good effect to raise the production capacity or reduce the resource consumption.

In order to make a stronger trust relationship between industrial entities in industrial Internet, our work could be combined with anticounterfeiting system [57] to reach a higher data privacy level or applied in production lifecycle management to reach sustainable manufacturing [58] with privacy. We would research the feasibility of those application scenarios with our work in the future and our work could make contribution to meeting the privacy demand in industrial Internet to some extent.

It is worth mentioning that our work still needs to be improved. Firstly, in this paper, the traceability feature generation process still depends on traceability knowledge held by a third party or public production standard [59]. Future studies should focus on areas such as artificial intelligence or industrial big data. Traceability features will be an important research direction in our future work. Secondly, security issues also need to be further researched, especially the security-oriented in blockchain system and smart contract. Our work is based on the premise of nonmalicious blockchain nodes; security protection schemes should be further discussed in order to provide more secure environment for blockchain [12].

Data Availability

The gas consumption data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China under Grant no. 2021YFB3300900, the National Natural Science Foundation of China (no. 62072336), and the New Generation Artificial Intelligence Technology Major Project of Tianjin (no. 19ZXZNGX00080).

References

[1] Y. Wang and M. Singgih, J. Wang, M. Rit, "Making sense of blockchain technology: how will it transform supply chains?" *International Journal of Production Economics*, vol. 211, pp. 221–236, 2019.

[2] M. Balog and L. Knapíková, "Advances of intelligent techniques used in Industry 4.0: proposals and testing," *Wireless Networks*, vol. 27, 2019.

[3] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent manufacturing in the context of industry 4.0: a review," *Engineering*, vol. 3, no. 5, pp. 616–630, 2017.

[4] D. Li, E. L. Xu, and L. Li, "Industry 4.0: state of the art and future trends," *International Journal of Production Research*, vol. 56, no. 8, pp. 2941–2962, 2018.

[5] S. Cisneros-Cabrera, G. Pishchulov, P. Sampaio, N. Mehandjiev, Z. Liu, and S. Kununka, "An approach and decision support tool for forming Industry 4.0 supply chain collaborations," *Computers in Industry*, vol. 125, Article ID 103391, 2021.

[6] F. Kerschbaum, A. Schroepfer, A. Zilli et al., "Secure collaborative supply-chain management," *Computer*, vol. 44, no. 9, pp. 38–43, 2011.

[7] G. B. Zhang, Y. Ran, and X. L. Ren, "Study on product quality tracing technology in supply chain," *Computers & Industrial Engineering*, vol. 60, no. 4, pp. 863–871, 2011.

[8] xxxx.

[9] R. Naderi, M. Shafiei Nikabadi, A. Alem Tabriz, and M. S. Pishvaee, "Supply chain sustainability improvement using exergy analysis," *Computers & Industrial Engineering*, vol. 154, no. 1, Article ID 107142, 2021.

[10] J. Q. Li, F. R. Yu, G. Deng, C. Luo, Z. Ming, and Q. Yan, "Industrial internet: a survey on the enabling technologies, applications, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1504–1526, 2017.

[11] S. Ammendrup and L. O. Barcos, "Aplicación de los sistemas de trazabilidad," *Revue Scientifique et Technique de l'OIE*, vol. 25, no. 2, pp. 763–773, 2006.

[12] H. N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for internet of things: a survey," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8076–8094, 2019.

[13] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A Survey on the Security of Blockchain Systems," *Future Generation Computer Systems*, vol. 107, 2018.

[14] S. Johar, N. Ahmad, W. Asher, H. Cruickshank, and A. Durrani, "Research and applied perspective to blockchain technology: a comprehensive survey," *Applied Sciences*, vol. 11, no. 14, p. 6252, 2021.

[15] A. Manimuthu, R. Sreedharan, and D. Marwaha, "A literature review on bitcoin: transformation of crypto currency Into a global phenomenon," *IEEE Engineering Management Review*, vol. 47, no. 1, pp. 28–35, 2019.

[16] S. S. Kushwaha, S. Joshi, D. Singh, M. Kaur, and H.-N. Lee, "Systematic review of security vulnerabilities in ethereum blockchain smart contract," *IEEE Access*, vol. 10, pp. 6605–6621, 2022.

[17] S. Wang, L. Ouyang, Y. Yuan, X. Ni, X. Han, and F. Y. Wang, "Blockchain-Enabled smart contracts: architecture, applications, and future trends," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 11, pp. 2266–2277, 2019.

[18] J. Leng, S. Ye, M. Zhou et al., "Blockchain-secured smart manufacturing in industry 4.0: a survey," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 237–252, 2021.

[19] A. Bedin, M. Capretz, and S. Mir, "Blockchain for Collaborative Businesses," *Mobile Networks and Applications*, vol. 26, pp. 1–8, 2020.

[20] K. Demestichas, N. Peppes, T. Alexakis, and E. Adamopoulou, "Blockchain in agriculture traceability systems: a review," *Applied Sciences*, vol. 10, no. 12, p. 4113, 2020.

[21] J. F. Galvez, J. C. Mejuto, and J. Simal-Gandara, "Future challenges on the use of blockchain for food traceability analysis," *TRAC Trends in Analytical Chemistry*, vol. 107, pp. 222–232, 2018.

[22] S. Soni and B. Bhushan, "A Comprehensive survey on Blockchain: working, security analysis, privacy threats and potential applications," in *Proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, July 2019.

[23] J. Brandt, I. Damgard, P. Landrock, and T. Pedersen, "Zero-knowledge authentication scheme with secret key exchange," *Journal of Cryptology*, vol. 11, no. 3, pp. 147–159, 1998.

[24] D. Catalano and I. Visconti, "Hybrid commitments and their applications to zero-knowledge proof systems," *Theoretical Computer Science*, vol. 374, no. 1-3, pp. 229–260, 2007.

[25] J. Kim, J. Lee, and H. Oh, "Simulation-extractable zk-SNARK with a single verification," *IEEE Access*, vol. 8, Article ID 156569, 2020.

[26] Y. Zhang, Y. Long, Z. Liu, Z. Liu, and D. Gu, "Z-channel: Scalable and Efficient Scheme in Zerocash," *Information Security and Privacy*, Springer, Cham, Switzerland, 2018.

[27] I. Weber, X. Xu, R. Riveret, G. Governatori, A. Ponomarev, and J. Mendling, "Untrusted Business Process Monitoring and Execution Using Blockchain," *Business Process Management*, Springer International Publishing, Berlin, Germany, 2016.

[28] D. Berdik, S. Otoum, N. Schmidt, D. Porter, and Y. Jararweh, "A survey on blockchain for information systems management and security - ScienceDirect," *Information Processing & Management*, vol. 58, no. 1.

[29] X. C. Chen, D. Peng, Y. Nai-qing, and M.-X. Bi, "Study on discrete manufacturing quality control technology based on big data and pattern recognition," *Mathematical Problems in Engineering*, vol. 2021, Article ID 8847094, 10 pages, 2021.

[30] T. P. Tsai and F. C. Wang, "Improving supply chain management: a model for collaborative quality control advanced semiconductor manufacturing," in *Proceedings of the 2004. ASMC '04. IEEE Conference and Workshop IEEE*, Boston, MA, USA, May 2004.

[31] P. Helo and A. Shamsuzzoha, "Real-time supply chain—a blockchain architecture for project deliveries," *Robotics and Computer-Integrated Manufacturing*, vol. 63, Article ID 101909, 2020.

[32] X. N. Zhu, G. Peko, D. Sundaram, and S. Piramuthu, "Blockchain-based agile supply chain framework with IoT," *Information Systems Frontiers*, pp. 1–16.

[33] G. Zhao, S. Liu, C. Lopez, H. Lu, S. Elgueta, and B. M. Boshkoska, "Blockchain Technology in Agri-Food Value Chain Management: A Synthesis of Applications, Challenges and Future Research Directions - ScienceDirect," *Computers in Industry*, vol. 109, 2019.

[34] T. K. Agrawal, V. Kumar, R. Pal, L. Wang, and Y. Chen, "Blockchain-based framework for supply chain traceability: a case example of textile and clothing industry," *Computers & Industrial Engineering*, vol. 154, Article ID 107130, 2021.

[35] M. Uddin, "Blockchain Medledger: hyperledger fabric enabled drug traceability system for counterfeit drugs in pharmaceutical industry," *International Journal of Pharmaceutics*, vol. 597, Article ID 120235, 2021.

[36] M. P. Caro, M. S. Ali, M. Vecchio, and R. Giaffreda, "Blockchain-based Traceability in Agri-Food Supply Chain Management: A Practical Implementation," in *Proceedings of the 2018 IoT Vertical and Topical Summit on Agriculture - Tuscany (IOT Tuscany)*, pp. 1–4, Tuscany, Italy, May 2018.

[37] K. Yang, D. Forte, and M. Tehranipoor, "ReSC: an RFID-enabled solution for defending IoT supply chain," *ACM Transactions on Design Automation of Electronic Systems*, vol. 23, 2018.

[38] Z. Wang, Z. Zheng, W. Jiang, and S. Tang, "Blockchain-Enabled data sharing in supply chains: model, operationalization, and tutorial," *Production and Operations Management*, vol. 30, no. 7, pp. 1965–1985, 2021.

[39] J. Eberhardt and S. Tai, "ZoKrates - scalable privacy-preserving off-chain computations," in *Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data*, July 2018.

[40] M. Westerkamp and J. Eberhardt, "zkRelay: facilitating Sidechains using zkSNARK-based Chain-Relays," in *Proceedings of the IEEE SECURITY & PRIVACY ON THE BLOCKCHAIN (IEEE S&B 2020)*, September 2020.

[41] M. K. Ibrahem, *Robust Electronic Voting System Using Homomorphic Encryption Protocol and Zero-Knowledge Proof*, vol. 5, no. 1, 2016.

[42] A. A. Rasheed, R. N. Mahapatra, and F. H. Lup, "Adaptive Group-Based Zero Knowledge Proof-Authentication Protocol (AGZKP-AP) in Vehicular Ad-Hoc Networks," *IEEE*

Transactions on Intelligent Transportation Systems, vol. 21, 2019.

[43] S. H. Jeong and B. Ahn, "Implementation of real estate contract system using zero knowledge proof algorithm based blockchain," *The Journal of Supercomputing*, vol. 77, no. 10, Article ID 11881, 2021.

[44] H. Qi, Z. Wan, Z. Guan, and X. Cheng, "Scalable decentralized privacy-preserving usage-based insurance for vehicles," *IEEE Internet of Things Journal*, vol. 8, p. 1, 2020.

[45] M. Umar, Z. Wu, and X. Liao, "Channel characteristics aware zero knowledge proof based authentication scheme in body area networks," *Ad Hoc Networks*, vol. 112, no. 9, Article ID 102374, 2021.

[46] C. Huang, D. Liu, J. Ni, R. Lu, and X. Shen, "Achieving accountable and efficient data sharing in industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1416–1427, 2021.

[47] X. Yang, "Design and application of safe production and quality traceability system for vegetable," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 24, no. 3, pp. 162–166, 2008.

[48] M. Thakur and C. R. Hurburgh, "Framework for implementing traceability system in the bulk grain supply chain," *Journal of Food Engineering*, vol. 95, no. 4, pp. 617–626, 2009.

[49] Y. Lu, P. Li, and H. Xu, "A Food anti-counterfeiting traceability system based on Blockchain and Internet of Things," *Procedia Computer Science*, vol. 199, pp. 629–636, 2022.

[50] Z. Ge, P. Li, W. Ren, S. Hu, Q. Yin, and Q. Zhang, "Quantity traceability method of 1000 kV standard voltage transformers," in *Proceedings of the 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, September 2020.

[51] Y. Cai, X. Li, M. Li et al., "Traceability and quality control in traditional Chinese medicine: from chemical fingerprint to two-dimensional barcode," *Evidence-based Complementary and Alternative Medicine*, vol. 2015, Article ID 251304, 6 pages, 2014.

[52] R. Mahmoud, T. Yousuf, F. Aloul, and I. Zualkernan, "Internet of Things (IoT) Security: Current Status, Challenges and Countermeasures," in *Proceedings of the 2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, London, UK, December 2015.

[53] C. S. Kouzinopoulos, G. Spathoulas, K. M. Giannoutakis et al., "Using Blockchains to Strengthen theSecurity of Internet of Things," *Security in Computer and Information Sciences*, Springer, Cham, Switzerland, 2018.

[54] D. Li, C. Li, and R. Gu, "Evolutionary game analysis of promoting industrial internet platforms to empower manufacturing SMEs through value cocreation cooperation," *Discrete Dynamics in Nature and Society*, vol. 2021, Article ID 4706719, 14 pages, 2021.

[55] J. Leng, M. Zhou, J. L. Zhao, Y. Huang, and Y. Bian, "Blockchain Security: A Survey of Techniques and Research Directions," *IEEE Transactions on Services Computing*, vol. 99, 2020.

[56] D. Vashisth, P. Khandelwal, R. Johari, and V. Gaur, "BlockChain technology based smart contract agreement on REMIX IDE," in *Proceedings of the 8th International Conference on Signal Processing and Integrated Networks (SPIN 2021)*, Noida, India, August 2021.

[57] J. Leng, P. Jiang, K. Xu et al., "Makerchain: a blockchain with chemical signature for self-organizing process in social manufacturing," *Journal of Cleaner Production*, vol. 234, pp. 767–778, 2019.

[58] J. Leng, G. Ruan, P. Jiang et al., "Blockchain-empowered sustainable manufacturing and product lifecycle management in industry 4.0: A survey," *Renewable and Sustainable Energy Reviews*, vol. 132.

[59] L. U. Tie, "On technical standardization and industrial standard strategy," *China Industrial Economy*, vol. 7, pp. 43–49, 2005.