






Research Article

Incentive Mechanism of Data Storage Based on Blockchain for Wireless Sensor Networks

Yongjun Ren ^{1,2}, Yepeng Liu ^{1,2}, Sai Ji ^{1,2}, Arun Kumar Sangaiah ³ and Jin Wang ⁴

¹School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, China

²Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET), Nanjing University of Information Science & Technology, Nanjing, China

³School of Computing Science and Engineering, Vellore Institute of Technology (VIT), Vellore, India

⁴School of Computer & Communication Engineering, Changsha University of Science & Technology, Changsha, China

Correspondence should be addressed to Jin Wang; jinwang@csust.edu.cn

Received 20 April 2018; Accepted 5 August 2018; Published 29 August 2018

Academic Editor: Yuh-Shyan Chen

Copyright © 2018 Yongjun Ren et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, the blockchain technology is utilized to build the first incentive mechanism of nodes as per data storage for wireless sensor networks (WSNs). In our system, the nodes storing the data are rewarded with digital money. The more the data stored by the node, the more the reward it achieves. Moreover, two blockchains are constructed. One is utilized to store data of each node and another is to control the access of data. In addition, our proposal adopts the provable data possession to replace the proof of work (PoW) in original bitcoins to carry out the mining and storage of new data blocks, which greatly reduces the computing power comparing to the PoW mechanism. Furthermore, the preserving hash functions are used to compare the stored data and the new data block. The new data can be stored in the node which is closest to the existing data, and only the different subblocks are stored. Thus, it can greatly save the storage space of network nodes.

1. Introduction

Wireless sensor network (WSN) has become very hot research topic recently in the field of microelectronics, communication, network, database, etc., because of its broad application prospects. It combines multiple technologies, such as sensing, computing, and wireless communication. The physical targets are monitored in real time through various types of microsensors, producing a large number of perceptual data at an unprecedented rate. Although the application scenarios and the deployment of hardware are different, the ultimate goal is to collect, transmit, and process the perceived data. Finally, users can achieve interesting information from the data [1, 2].

The wireless sensor network is a data-centric network. Therefore, the data storage of nodes is the fundamental problem in WSN, which should be solved. For the users, what they concerned are the perception of the data, rather than the sensor node itself and the networks they make up.

Furthermore, the wireless sensor networks support efficient and reliable data storage and access under the heterogeneous, unreliable environment. As the storage space and energy of each node are limited, how to effectively store data in the limited storage space has been an important research hot spot of data management in WSN.

The normal operations of WSN require the cooperation of network nodes. However, some network nodes may choose selfish behavior due to their limited resources, such as energy and storage space. If most network nodes take selfish behavior and do not forward packets, the entire network will not be able to provide normal service. Therefore, inciting selfish nodes to cooperate and ensuring the normal operation of the entire network are part of the important researches in WSN.

Traditionally, the solutions to the selfishness problem of nodes in WSN have based on the mechanisms of game theory and the mechanisms based on reputation. But the researches mainly focus on data transmission and packet

forwarding. Moreover, now there is no specific incentive mechanism of data storage for nodes in WSN.

The storage capacity of nodes in WSN is limited, and the data storage capacity is also an important resource. This paper focuses on the incentive of data storage in WSN. In this paper, the blockchain technology is adopted to construct the first incentive mechanism of nodes' data storage in WSN. In our system, the data set which is storing every node is considered as a block of the blockchain. If the nodes store the data, they will be rewarded with digital money (bitcoins, etc.). Additionally, if the nodes store more data, they will attain more rewards. When mining and storing new data blocks in progress, we apply the provable data possession instead of the proof of work (PoW) in original bitcoins. The method can greatly reduce the computing power of the miners. Apart from this, comparing the existing data in nodes with the new data block, we can take advantage of the preserving hash functions. The node stores the new data, which is closest to the existing data, and only the distinct subblocks need to be stored. So, it greatly saves the storage space of network nodes.

The rest of this paper is organized as follows. Section 2 introduces the related works of data storage strategy in WSN and incentive mechanism. In Section 3, we analyze the existing problem of data storage in WSN. Section 4 presents the building blocks of our scheme based on the blockchain. And in Section 5, the incentive mechanism of data storage based on blockchain in WSN is proposed. Finally, Sections 6 and 7 present the discussion and conclusions of this paper, respectively.

2. Related Work

2.1. Data Storage Strategy in WSN. At present, there are three main ways of data storage in WSN: external storage, local storage, and data-centric storage [3].

2.1.1. External Storage. Sink node is a special kind of storage node, and its storage space and energy are not restricted and do not need to consume another node energy. Other nodes will send the collected data to the sink node, which will consume a lot of energy. If all the nodes in the network send data to the sink node, it will cause the network block and the nearby sink nodes will be invalid.

The LEACH protocol is proposed to collect data from the hierarchical sensor network, in which a subset of nodes is randomly selected as cluster heads, and the other nodes added different cluster according to the calculated distances between the nodes and the cluster heads. During a period, the nodes transmit data to the cluster heads, and the cluster heads process the data and then sends them to the sink nodes. The PEGASIS protocol [4] improved the LEACH protocol, in which the sensor network was organized into a chain structure. Each node receives and forwards data by its neighboring nodes. The sink nodes only select one other node to communicate with it. The data are aggregated in the process of forwarding from a node to the next node and eventually reaching the sink node. Thus, the consumed

energy in the PEGASIS protocol is less than that in LEACH. Wang et al. also proposed a new protocol [5], which is an improvement to the LEACH protocol. The protocol establishes the soft and hard threshold, which can dynamically adjust and compare the collected data to reduce unnecessary data transmission. When the node data are above the hard threshold, the data are transmitted and they are taken as a new hard threshold.

The storage strategy of external storage is mainly focused on data acquisition, ignoring the data storage ability of WSN and the demand of nodes for data.

2.1.2. Local Storage. In local storage, the data are stored in nodes of the network, which consumes little energy. The query commands are only sent to the other nodes. After the node receives the query and processes it, the result is passed to the sink node. So, queries consume longer delays.

The directed diffusion protocol stores the data collected by the nodes in the local nodes. The sink nodes achieve their information by broadcasting the "interest message" to the other network nodes. The node that received the message creates a gradient within the network, pointing to the sink node. The node establishes one or more paths to the sink nodes, doing flood search and performing data transmission. The geographic and energy-aware routing (GEAR) protocol [6] is the improvement of the directed diffusion protocol. In GEAR protocols, when a query message is sent in the target area, the propagation of the "interest message" is limited to the target area because of the geographical location, which avoids flooding in entire network and reduces the cost of routing.

The storage process of local storage strategy is simple. And the strategy focuses on data query processing and has less description of information, which leads to a lot of energy in the query process.

2.1.3. Data-Centric Storage. The data-centric storage is a hot research direction in recent years. It mainly studies how to store the perceived data of sensor nodes so as to ensure the high efficiency, stability, and real-time performance of the later query.

The concept of data-centric storage (DCS) is proposed by Meyfroyt et al., and the data storage algorithm GHT is designed based on the geographic information mapping table [7]. Its core idea is that data are stored according to their attributes, and a specific data are defined as an event. The sensor detects the data, hashes the event through a hash function, then achieves a geographic location, and saves the data to the nearest node based on the geographic information [8]. The algorithm is conducive to data query, which is only based on the query event attributes. And the use of mapping function can be found in the storage node, which avoids flooding. The disadvantage of the algorithm is the lack of efficient storage hot spot processing mechanism. When the data storage overloads, it cannot be transferred to another node. Moreover, accessing geographic information needs GPS and consumes system energy. In the data storage algorithm ARI [9], adaptive ring index structure is used to

solve the hot spot problem of the DCS algorithm. And hash functions are utilized to hash a certain type of event to the event storage node. A ring is created around the event storage node, and events are dispersed and stored in the index nodes. In general, it is difficult to define clear demarcation of a wireless sensor network, which is not ideal for hot spot problems. In data storage algorithm of Reference [10], two-tier data storage structure is used to track the moving target of the mobile multisink node in the WSN. Data are transferred and stored through the creation of virtual grids in the algorithm. When the data collected by the grid storage nodes are queried, it is just needed to flood the request within the grid, which will save energy. In addition, some scholars have proposed a distributed index structure algorithm (DIFS) [11]. DIFS is an improvement of the TTDD algorithm. In DIFS, multilevel quadtree is constructed based on spatial decomposition technique and hash function, and the geography hash method is used as the index of data [12, 13]. The corresponding node stores the observed data through hash functions, and it can determine the range of the minimum number of index nodes by the query range.

2.2. Incentive Mechanism. At present, there are two main incentive mechanisms. One is based on game theory. The other is based on external incentives [14–20].

2.2.1. Incentive Mechanism Based on Game Theory. In the paper [14], the concept of multidomain wireless sensor network was first proposed, and the game theory was used to evaluate the impact of cooperative behavior. In the system, the participants in game analysis are the various individual wireless sensor networks, and it is assumed that each wireless sensor network has to make decisions: whether to help other networks to carry out data transfer and whether to request other networks to help its data transmission, which is the strategy of each participant in game analysis. On the basis of the above mechanism, the problem was continued to study the cooperative behavior among networks in multidomain wireless sensor networks [15]. The main differences in the game analysis are as follows: (1) the income function of the game is mainly expressed by the whole life cycle number of the network [16], rather than the calculation of the accumulated revenue of nodes and (2) the strategy of the sensor node is more intelligent. The choice of actions will be limited after many unsuccessful data transfers so that the network has minimal QoS guarantee. References [17, 18] analyzed the impact of different cooperation strategies on the life cycle of multidomain wireless sensor networks. The author proposes a linear design framework and uses the corresponding one-dimensional and two-dimensional linear models to assess the performance of different strategies. Based on the ideal conditions, the author adds various restrictions to observe the influence on the cooperation strategy. Simulation experiments have confirmed that cooperation can significantly extend the life cycle of the network. And under some special circumstances, some cooperative strategies can increase the life cycle to an order of magnitude.

2.2.2. Incentive Mechanism Based on External Incentives. In addition to the use of game theory to analyze the multidomain wireless sensor network, there are some researches of external incentive mechanisms. The main external incentive methods include virtual currency mechanism and honor incentive mechanism. In [19–21], an economic model of dynamic prices and incentive methods is proposed to study the cooperation in multidomain wireless sensor networks. And the proposed economic model and the traditional routing protocol AODV protocol [22] are merged into a hybrid protocol for simulation experiments. In the simulation experiment, the author compared the proposed NES method with other EES methods and PDM [23]. The experimental results confirm that the cooperation between the sensor networks will be enhanced, and the overall energy consumption in the network will be significantly reduced.

3. Problem Statement

The development of wireless sensor networks originated from military applications, such as battlefield monitoring. Nowadays, wireless sensor networks have been applied to many civilian applications, such as environmental and ecological monitoring, healthcare, home automation, and traffic control.

In the sensor network, nodes are deployed in a variety of ways within or around a perceived object. These nodes form a wireless network through self-organization method. And they can sense, collect, and process specific information in a cooperative way within the coverage area. Finally, it can realize the collection, processing, and analysis of any location information at any time. Each node of the sensor network is not only equipped with a radio transceiver but also a small microcontroller and an energy source (usually a battery), in addition to multiple sensors. The size of a single sensor node is as large as a shoe box, as small as dust. The size and complexity of the restrictions for sensor nodes determine the constraints of energy, storage, computing speed, and bandwidth. In large sensor networks, the sensor and network structure are different. Thus, the integration of heterogeneous networks often occurs in sensor networks. At the same time, the heterogeneous network structure also brings difficulty to data storage and sharing in WSN.

Moreover, the data storage capacity is also an important resource. But the storage capacity of nodes in a wireless sensor network is limited. Some network nodes give up storing data in order to save their own storage and energy resources, which are called selfish behavior. If the most network nodes behave selfishly and do not store data, then the entire network will not be able to provide normal service.

To solve the problem, we use incentive mechanisms based on blockchain to encourage network nodes to store data. The data storage based on the blockchain technology can not only provide the corresponding data storage function but also reward the digital currency to the network node that stores data. Therefore, data storage based on the blockchain technology in WSN is very suitable.

4. Building Blocks

4.1. Blockchain Technology. The blockchain system contains the following important components: underlying transaction data, distributed ledgers, important consensus mechanism, complete and reliable distributed P2P network, and distributed application on the network. And the framework is shown in Figure 1. The underlying data are organized into blocks, and each block is chained into a chain in the chronological order, which is called blockchain [24–26]. Each node of a fully distributed network stores a distributed ledger, that is, blockchain. The P2P protocol is used in the network to communicate with each other. All parties will reach agreement through consensus mechanisms. Advanced applications are generated based on these foundations. In the architecture, the nontampering blockchain data structure, the consensus mechanism in distributed network, the proof of work mechanism, and the increasingly flexible smart contracts are representative innovations [27, 28].

The underlying data are not stored in the blockchain. The raw data need further processing so that they can be written into the block. The underlying data are the most fundamental transaction records; the other data are only intended to encapsulate the message records. The network layer encapsulates the networking mode of the blockchain system, the message propagation protocol, and the data authentication mechanism. Combining with the practical application requirements and designing the specific propagation protocol and data verification mechanism, each node in the blockchain system can participate in the checksum accounting process of the block data. Only when the block data are verified by most nodes in the whole network, the block is recorded in the blockchain [29–31].

The PoW mechanism is an important innovation that closely integrates the functions of currency issuance, transaction payment, and verification. And the safety and decentric of the blockchain system are ensured through the competition of computing force. The core idea is to ensure the consistency of data and the security of the consensus by the computing force competition of distributed nodes. In the bitcoin system, the miners work together to solve a complex but easy-to-validate SHA-256 mathematical problems (i.e., mining) based on their respective computer forces. The nodes that solve the problem the fastest will get the right to account the block and bitcoin reward. The mathematical problem can be expressed as follows. Based on the current difficulty value, a suitable random number (Nonce) is sought so that the double SHA-256 hash of the metadata of the block header is less than or equal to the target hash value. However, the PoW consensus mechanism has a significant flaw: the waste of resources (such as electricity), caused by their strong computing power, has always been criticized by researchers [32–34].

The consensus process of the blockchain system realizes the data validation and accounting of shared blockchain ledgers by aggregating the computational power resources of large-scale consensus nodes, so it is essentially a task crowdsourcing process of consensus

nodes. In the decentralized system, the consensus nodes themselves are selfish, and maximizing its own revenue is the fundamental goal of its participation in data validation and accounting. Therefore, it is necessary to design a reasonable and well-conceived mechanism of incentive and compatibility so that the individual rational behavior of the consensus node maximizing its own income is consistent with the overall goal of guaranteeing the safety and effectiveness of the decentralized blockchain system. The blockchain system integrates large-scale nodes and forms a stable consensus on the history of the blockchain by designing a modest economic incentive mechanism and integrating with the consensus process.

The contract layer is business logic and algorithm based on the blockchain virtual machine, which is the basis for realizing the flexible programming and operation data of the blockchain system. The smart contract has important significance to the blockchain system, which not only provides the programmable capabilities to the underlying data of the blockchain but also encapsulates the complex behavior of each node in the blockchain network. And it provides a convenient interface for building an upper application based on blockchain technology. Thus, blockchain technology with smart contract is extremely broad prospects.

4.2. PDP Mechanism. Provable data possession (PDP) mechanism is used to determine whether the data on the remote node are damaged (Figure 2). The PDP mechanism was first used in grid computing and P2P networks. He et al. constructed the PDP mechanism using RSA-signed homomorphic properties, but this mechanism requires that the entire file is represented by a large number, which results in high computational costs. Wang et al. proposed a probabilistic strategy to complete the integrity verification, using the homomorphic properties of the RSA signature mechanism to aggregate the evidence into a small value, greatly reducing the communication overhead of the protocol [35–38]. Wang et al. realized another mechanism that supports full dynamic operation of the PDP mechanism. It considers the use of the Merkle hash tree in order to ensure the correctness of the data block in position, and data block value ensures its correctness through the BLS signature mechanism [39–42]. In order to reduce the burden on the user, the mechanism also introduces an independent third party instead of the user to verify the integrity of outsourced data. In this article, this algorithm is used to replace the PoW mechanism in the original blockchain.

The PDP scheme is as follows. At first, encode M into M' so that each data block m_i of M' contains s data segments, that is, $m_i = (m_{i,1}, m_{i,2}, \dots, m_{i,s})$. The metadata σ_i are calculated for each data block m_i as follows:

$$\sigma_i = \left(H(\text{name}||i) \times \prod_{j=1}^s u_j^{m_{i,j}} \right)^\alpha, \quad (1)$$

where α is the private key of the user and u_j ($1 \leq j \leq s$) is randomly selected from the bilinear group G . Similar to the

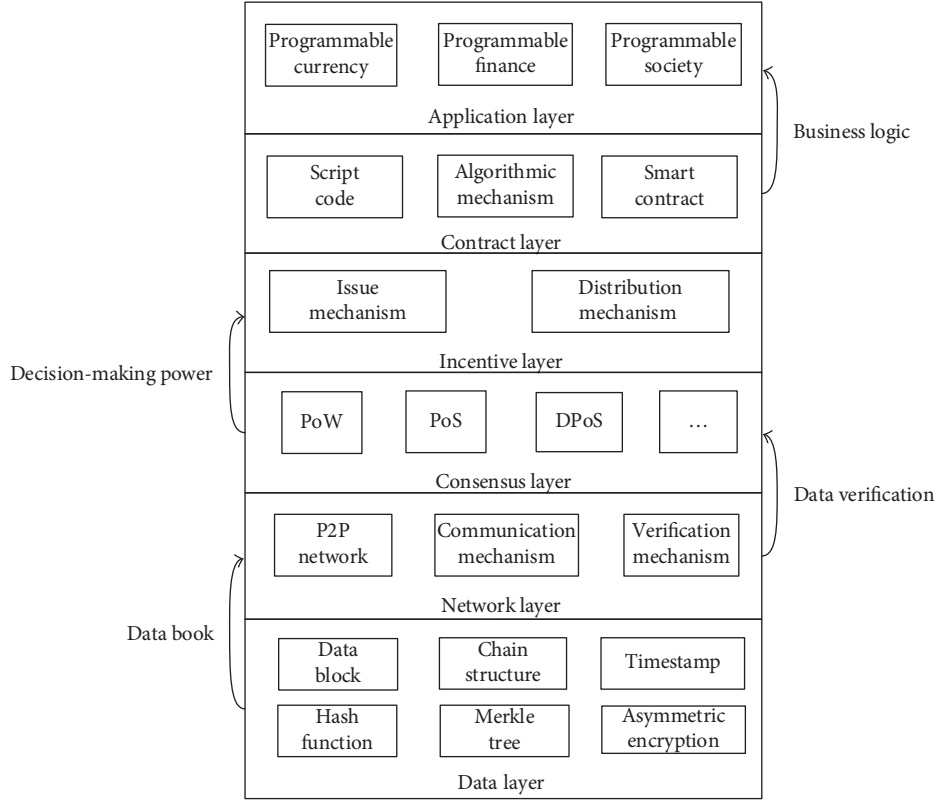


FIGURE 1: A basic framework of blockchain.

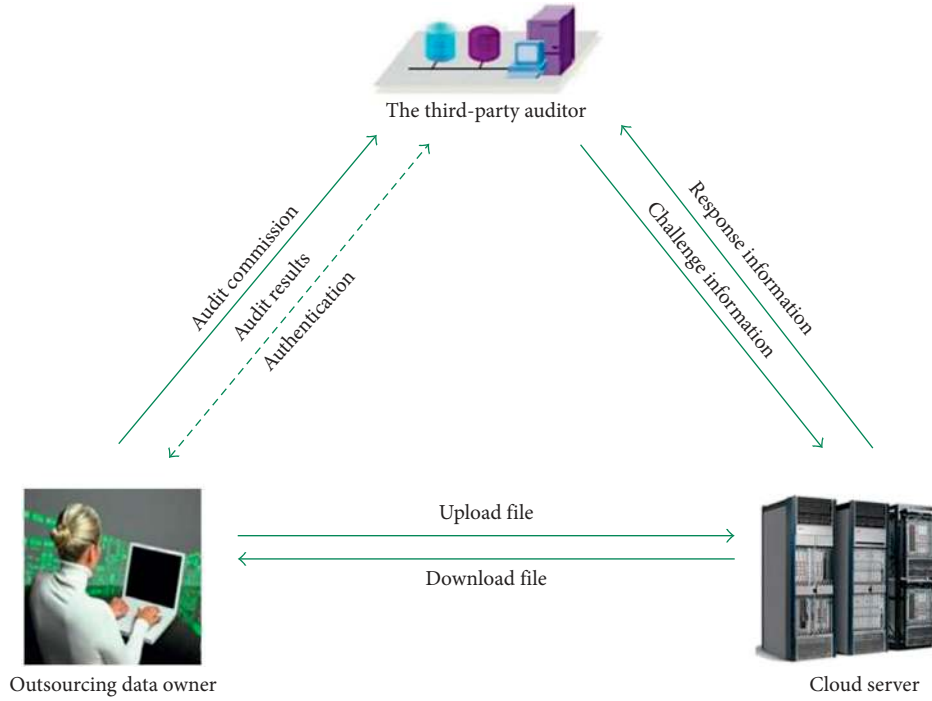


FIGURE 2: Provable data possession.

literature [4], the factor $(\prod_{j=1}^s u_j^{m_{i,j}})^{\alpha}$ contained in the metadata σ_i also supports the aggregation operation. So, the cloud storage server can generate the corresponding partial

aggregation in the integrity verification phase. The algorithm also signs the data name, the number of data blocks, and the parameter u_j to obtain a tag of data r .

To verify the integrity of the outsourced data, a query challenge $C = \{(i, v_i)\}$ is submitted by the verifier, including the block number i which is randomly selected and the corresponding coefficient v_i . The cloud server calculates aggregated data blocks $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ and metadata σ as the proof, that is, (μ, σ) , where $\prod_{(i,v_i) \in C} \sigma_i^{v_i}$ is the metadata, the aggregated data blocks is $\mu_i = \sum_{(i,v_i) \in C} v_i m_{i,j}$.

Verification is done by checking the following formula and performing two bilinear operations:

$$e(\sigma, g) = e\left(\prod_{(i,v_i) \in C} H(\text{name}||i)^{v_i} \times \prod_{j=1}^s u_j^{\mu_i}, v\right), \quad (2)$$

where v is the user's public key corresponding to α .

In the above scheme, Shacham and Waters double the data for the first time so that each data segment $m_{i,j}$ corresponds to a data block of the aforementioned scheme. This segmentation strategy has the obvious advantage that by generating metadata for a set of data segments, the size of the processed data can be reduced, thereby reducing the storage costs of the cloud server.

5. Incentive Mechanism of Data Storage Based on Blockchain in Wireless Sensor Network

In this paper, the blockchain technology is utilized to build the first incentive mechanisms of nodes' data storage in WSN. In our system, the data set stored by every node is treated as a block of the blockchain. The nodes storing the data are rewarded with digital money (bitcoin, etc.). Moreover, the more the data stored by the node, the more the reward it achieves. Our proposal adopts the provable data possession to replace the proof of work (PoW) in original bitcoin to carry out the mining and storage of new data blocks. The method can greatly reduce the computing power by PoW mechanism. Furthermore, the preserving hash functions are used to compare the stored data and the new data block. Thus, the new data can be stored in the node which is closest to the existing data, and only the different subblocks are stored. So, it can greatly save the storage space of network nodes.

5.1. Blockchain of Data Storage for Sensor Node. The sensor network is often composed of multiple heterogeneous subsystems, and various network nodes have different capabilities in computing, energy, communication, and storage. In addition, the network nodes which using different types of sensors make the types of collected data varied. Therefore, the shared data storage mechanism should be adopted to realize the storage and management of the data in wireless sensor network. The blockchain has the advantage of decentralization. Moreover, the data storage based on decentering credit can be realized in the WSN, where the node does not need to be trusted using the encryption algorithm, time stamp, tree structure, consensus mechanism, and reward mechanism. Each network node can use the Merkle tree in the blockchain to store its data. The data of the

nodes are stored in the leaves of the Merkle tree. Each stored datum can be a block, and all the data stored by the nodes are linked to form the data blockchain (Figure 3).

5.2. Trust Management of Network Node. In the system, the trust of network nodes is managed. When the network node is found to be fraud and with other behaviors, it is removed from the WSN network. We use the reputation system to manage the nodes in WSN. Once the network node is found cheating, it will be immediately excluded from the WSN.

In the system, the trust of the data initiator (node i) in the network to the data store (node j) can be obtained by calculating the number of success and failure of the node data storage in a certain period of time. After the k th data storage is successful, d_{ij}^k indicates the trust evaluation value of the data initiator node i to the data store node j ; δ ($0 \leq \delta \leq 1$) is the time attenuation coefficient of the trust, which is used to reflect the influence degree of trust for the network node in data storage procession. The larger the weight of the recent score record, the greater the weight of the calculation of the trust value, as shown in the following equation:

$$d_{ij}^k = \sum_{m=1}^k \delta_m d_{ij}^m. \quad (3)$$

In the system, the trust of data storage between node i and node j is divided into five levels according to the satisfaction degree, and 0, 0.25, 0.5, 0.75, and 1 are assigned in turn. The first level indicates that the data storage between the network node i and the network node j is failure, and the node i considers the node j is malicious. The second, third, fourth, and fifth levels of trusts are sequentially increased. The fifth level is the highest level, indicating that the data storage between the network node i and the network node j is successful, and that the node i fully trusts the node j . When there is a data storage relationship between the two nodes i and the node j , the degree of trust of node i to node j is calculated using Equation (1). When there is no direct transaction between the two nodes, use the following formula to calculate the average trust of the network as the recommended trust degree of node:

$$d_0 = \frac{\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^k d_{ij}^k}{n^2 \sum_{k=1}^k k}. \quad (4)$$

Most nodes in the network play dual role. One role is consumer, who is provided with storage service in the system. Another is the service provider, who provides storage service for other nodes. As a consumer, the trust evaluation of network nodes to other nodes is always considered accurate and deterministic. Therefore, the node modifies the data in the table with minimal possibility. Even if making a recommendation for a particular node, it does not make sense. In addition, it is safe to locally store the relevant calculated data of the trust value. As a service provider, it is the object to be evaluated. Any node i in the network cannot know the storage node which stores its reputation information, which avoids the possibility of the node to raise its reputation.

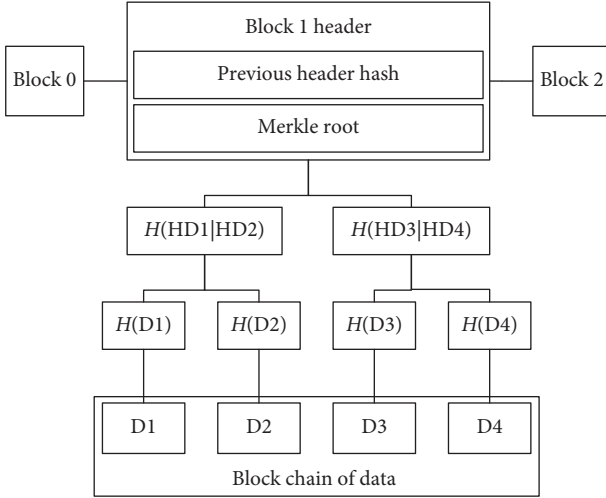


FIGURE 3: Data blockchain of the network node.

5.3. Access Control Based on Blockchain. We use blockchain to securely store access right to the data stored in the sink nodes. The data owner, the data visitor, and some additional metadata are included in the signed storage transaction. Each data block is set to access rights and is restricted in time. The data owner can extend or revoke the right to access the data. For any data retrieval request, another node first checks the access rights record through the corresponding distributed hash table (DHT). Theoretically, malicious nodes can share data without permission. Since the access rights of the data are monitored, unauthorized data access will be detected. In addition, if the malicious node is detected, it will be removed out of the network. Therefore, the possibility of such insecure data access is very small. It is shown in Figure 4.

Below we build a block-based DHT for distributed storage and management of index data. DHT is a huge hash table, which is shared by a large number of nodes. Each sink node is assigned to a hash block that belongs to itself and becomes the manager of the hash block. Through the hash function, any data can be mapped to a 160-bit hash value, and the network nodes are mapped to a space. DHT can adapt to the dynamic join and exit sink nodes and has the characteristics of balance and query accuracy. We use the DHT algorithm based on Chord network; through the SHA series hash function, the data are mapped to 160-bit hash value. For chord structure, we use the predecessor list positioning to improve the positioning fault tolerance, by selecting the node to reduce the positioning delay. That is, in the positioning process to select the next jump, those nodes which are a small delay and closer to the other nodes are selected in the bottom of the logistics. Take the above predecessor's search function as an example. Assuming that a node m returns to the predecessor list of the node n , in addition to the node location information, there is a delay of each node to m . Based on these delays, the node n evaluates each node in the list and selects the node that it considers the most reasonable.

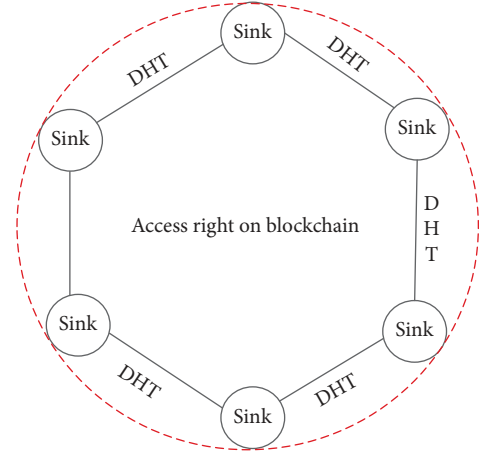


FIGURE 4: Access control based on blockchain.

5.4. Mining and Incentive Mechanism Based on PDP for Data Storage of Node in WSN. There is a significant flaw of the PoW consensus mechanism in traditional blockchain technology, which requires a lot of computation and causes serious waste of resources (such as electricity). That has always been criticized by academics and industry. In order to solve the problem, the PDP mechanism is used to replace the PoW mechanism to construct the mining and incentive mechanism for data storage of node in the resource-constrained WSN.

5.4.1. Scheme Description. A new data block, which will be stored in the sensor network, is broadcast. And each network node then calculates the challenge of PDP for the data block. If the PDP is verified correctly, the new data block will be stored by the node, and the node will receive a reward for storing the data block as a result, that is, a unit of the digital currency. The proposed scheme is as follows.

- (1) A new data block $M = \{m_1, m_2, \dots, m_n\}$ which will be stored. The public key of the data publisher is (g^x, u) , and the private key is x ; H_1 is a preserving hash function, and data publisher computes $\{H_1(m_i)\}$ and generates the authenticator $\sigma_i = (H(i)u^{m_i})^x$ for each subblock m_i ; The request information for the data is broadcast in the sensor network.
- (2) Each network node searches for the stored subblock m'_i closest to the value according to $\{H_1(m_i)\}$, that is, $|H_1(m_i) - H_1(m'_i)| \leq \text{dif}$. Then, the random number v_i will be selected for the subdata block i of the data block M , denoted as $Q = (i, v_i)$. The network node sends $\{H_1(m'_i)\}$ and Q to the data publisher.
- (3) The data issuer receives $\{H_1(m'_i)\}$ from each network node and compares them with the $\{H_1(m_i)\}$ value, selecting the $H_1(m'_i)$ value which is closest to each $H_1(m_i)$ value and adding the network node j that sent the $H_1(m'_i)$ value to the node set J .

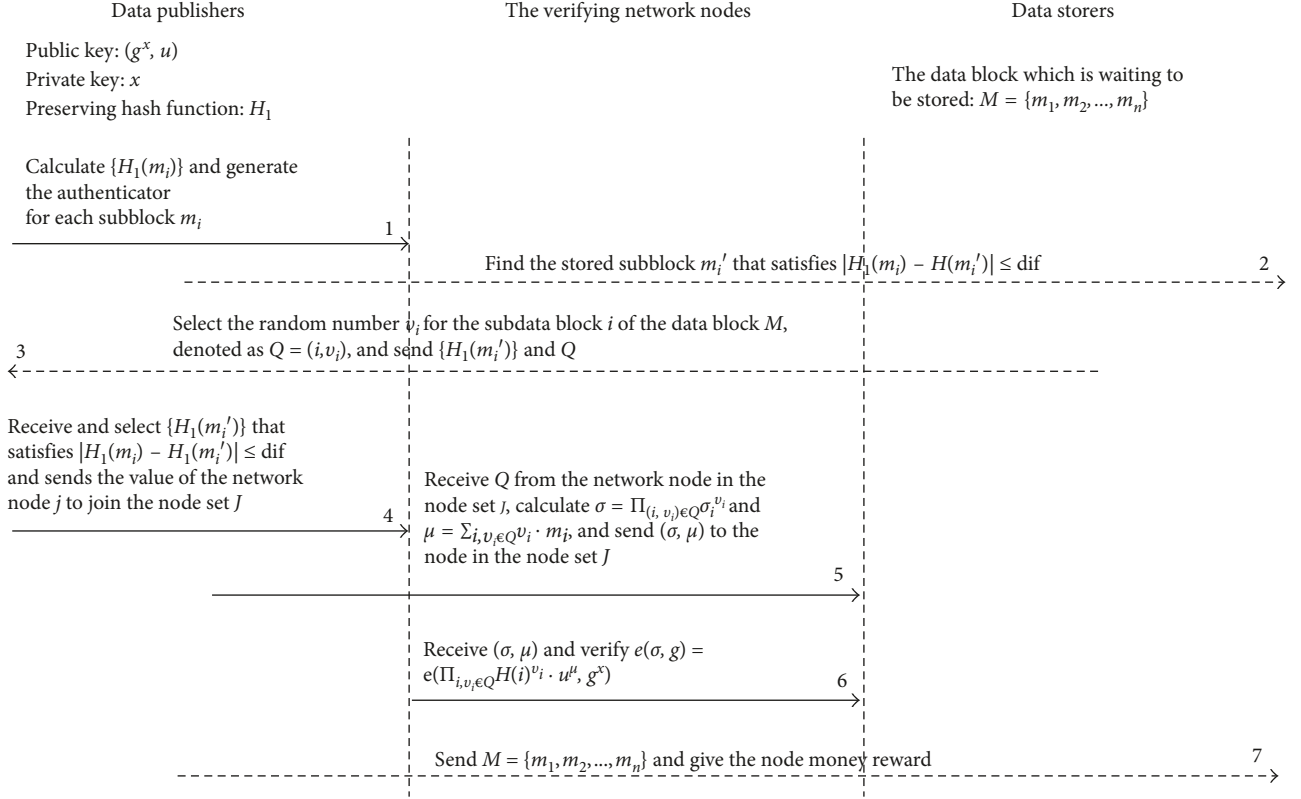


FIGURE 5: Incentive mechanism based on PDP.

Then, based on the Q received from the network node of the node set J , do the following calculation: $\sigma = \prod_{(i, v_i) \in Q} \sigma_i^{v_i}$ and $\mu = \sum_{(i, v_i) \in Q} v_i \cdot m_i$, and then send (σ, μ) to the network node of node set J .

- (4) The network node of the node set J receives (σ, μ) , verifying the following formula: $e(\sigma, g) = e(\prod_{(i, v_i) \in Q} H(i)^{v_i} \cdot u^x, g^x)$. If it is true, the data issuer will send the data block $M = \{m_1, m_2, \dots, m_n\}$ to each network node of the set J for storage and give the node the digital currency reward.
- (5) From the nature of the preserving hash function, it can be seen that the original data block of each network node in the set J contains data similar to the new data block $M = \{m_1, m_2, \dots, m_n\}$. It only needs to store the part that is not the same as the original data. Therefore, through the strategy, it can greatly reduce the required storage space. The scheme is shown in Figure 5.

5.4.2. Parameters in Our Scheme. In our scheme, we take the pairing function $e : G_1 \times G_1 \rightarrow G_T$, where $|G_1| = |G_T|$ and g, u are generators of the group G_1 . The practical constructions of pairings are done on hyperelliptic curves defined over a finite field. $E(F_q)$ is a set of points on an elliptic curve E defined over the finite field F_q . G_1 is taken as a subgroup of $E(F_q)$, and G_T is taken as a subgroup of $F_{q^k}^*$, where k' is the embedding degree. The hash function H hashes a binary string of arbitrary length into G_1 , and u is

a random element of G_1 . σ also is an element of G_1 , and μ belongs to Z_p . The Barreto–Naehrig (BN) curves are suitable for our scheme.

5.4.3. Efficient Storage. In our scheme, the network node stores l data segments $\{(m_j, \sigma_j)\}$, $j \in I$ and $|I| = l$. I is the set of indices of M corresponding to these l segments, and σ_j is the tag of the segment m_j . If SHA-256 is used to compute these hash values, then the size of each σ_j becomes 256 bits. This generates a small storage requirement for each of the segments, though the number of segments in the data M is huge in general. Instead of the Merkle proof, the nodes store a small tag and authenticator of size 256 bits along with each segment. Therefore, a network node in our scheme enjoys around 256 bits less storage overhead per segment.

6. Discussion

Compared with the PDP mechanism, the POR (proofs of retrievability) mechanism can effectively identify whether a file is damaged, and at the same time, it can recover the errors that have occurred in the data file through fault tolerance technology to ensure that the file is available. The POR mechanism can be further adapted in our scheme to improve the fault tolerance of the system.

The PDP mechanism can quickly determine whether the data on the remote node are damaged or not and pay more attention to efficiency. POR mechanism can not only identify whether the data are damaged but also recover the

damaged data. POR mechanism can not only detect data integrity but also further ensure data integrity. The publicly authenticated POR mechanism allows any third-party alternative user to initiate the integrity detection of data on a remote node. When the damage of the data was found less than a certain threshold ε , the error is recovered through the fault tolerance mechanism; otherwise, the data returned to the user fail.

Before the POR performs the initialization phase, it is needed to increase the redundant coded data preprocessing process to make the data file fault-tolerant, that is, to divide M into n blocks and then group n blocks. Then, for each group of data blocks, the Reed–Solomon error correction code can be used for fault tolerance coding to form a new data file. The same verification technology as PDP mechanism is adopted. For the POR mechanism, the assumption is that it is within the allowed error range (an error occurs once in 1000000 but passes the verification of the POR mechanism). Define $Y\omega = 1/\#B + (\rho n)^c / (n - c + 1)$, if $\varepsilon - \omega X$ is a negligible value, through $O(n/(\varepsilon - \omega))$ interactions, POR can recover the data with a failure rate of ρ . Here, B is the selection space of the random number when challenging the request, ρ is the data encoding rate, and c is the number of randomly selected data blocks.

7. Conclusions

In this paper, the first incentive mechanisms of nodes for data storage are built based on the blockchain technology in WSN. The data stored by every node are treated as a block of blockchain in our system. The reward for digital money will be obtained by the node who stored the data, and the reward for the node implementation increases as the data it store increases. In addition, it constructs two blockchains. One is to store data for each node, and the other for controlling the access of the data. Moreover, the provable data possession in the proposed scheme is used to substitute the proof of work (PoW) in primary bitcoins, which executes the mining and storage of the new data block. Compared with the PoW mechanism, it cuts down the computing power extremely. Furthermore, due to making use of the preserving hash functions, the new data can be stored in node which is nearest to the currently existing data. And only the different subblocks are stored. Therefore, the storage space of nodes in WSN can be highly saved.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Disclosure

The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

All the authors wrote the paper.

Acknowledgments

This work was supported by the NSFC (61772280, 61772454, 61702236, and 6171101570), the PAPD fund from NUIST, and Changzhou Science and Technology Program (CJ20179027).

References

- [1] J. Wang, Y. Cao, B. Li, H. Kim, and S. Lee, "Particle swarm optimization based clustering algorithm with mobile sink for WSNs," *Future Generation Computer Systems*, vol. 76, pp. 452–457, 2017.
- [2] B. Wang, X. Gu, L. Ma, and S. Yan, "Temperature error correction based on BP neural network in meteorological WSN," *International Journal of Sensor Networks*, vol. 23, no. 4, pp. 265–278, 2017.
- [3] L. Min, W. Fan, Z. Guo, and G. Fan, "Wireless sensor networks data storage strategy based on RCFfile," *Computer Science*, vol. 42, pp. 76–80, 2015.
- [4] S. C. Lindsey and S. P. Raghavendra, "Power efficient gathering in sensor information systems," in *Proceedings of IEEE Aerospace conference*, pp. 1125–1130, IEEE, Big Sky, MT, USA, March 2002.
- [5] J. Wang, C. Ju, H. J. Kim, R. S. Sherratt, and S. Lee, "A mobile assisted coverage hole patching scheme based on particle swarm optimization for WSNs," *Cluster Computing*, vol. 3, pp. 1–9, 2017.
- [6] N. Zaman, L. T. Jung, and M. M. Yasin, "Enhancing energy efficiency of wireless sensor network through the design of energy efficient routing protocol," *Journal of Sensors*, vol. 2016, Article ID 9278701, 16 pages, 2016.
- [7] T. M. Meyfroyt, S. C. Borst, O. J. Boxma, and D. Denteneer, "Data dissemination performance in large-scale sensor networks," in *Proceedings of International Conference on Measurement and Modeling of Computer Systems*, pp. 395–406, Austin, Texas, USA, June 2014.
- [8] X. Shen, W. Liu, I. W. Tsang, Q. S. Sun, and Y. S. Ong, "Multilabel prediction via cross-view search," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 99, pp. 1–15, 2018.
- [9] R. Huang, X. Chu, J. Zhang, and Y. H. Hu, "Scale-free topology optimization for software-defined wireless sensor networks: a cyber-physical system," *International Journal of Distributed Sensor Networks*, vol. 13, no. 6, pp. 1–12, 2017.
- [10] J. Wang, J. Cao, S. Ji, and J. H. Park, "Energy-efficient cluster-based dynamic routes adjustment approach for wireless sensor networks with mobile sinks," *Journal of Supercomputing*, vol. 73, no. 7, pp. 3277–3290, 2017.
- [11] Y. Liu, Q. Zhang, and L. Ni, "Opportunity-based topology control in wireless sensor networks," *IEEE Transactions on Parallel and distributed systems*, vol. 21, no. 3, pp. 405–416, 2010.

- [12] X. Shen, F. Shen, Q. S. Sun, Y. Yang, Y. H. Yuan, and H. T. Shen, "Semi-paired discrete hashing: learning latent hash codes for semi-paired cross-view retrieval," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4275–4288, 2018.
- [13] X. Shen, F. Shen, L. Liu, Y. H. Yuan, W. Liu, and Q. S. Sun, "Multiview discrete hashing for scalable multimedia search," *ACM Transactions on Intelligent Systems and Technology*, vol. 9, no. 5, pp. 1–21, 2018.
- [14] S. V. A. Jeba and R. S. Kumar, "Reliable anonymous secure packet forwarding scheme for wireless sensor networks," *Computers and Electrical Engineering*, vol. 48, pp. 405–416, 2015.
- [15] J. R. M. Dios, K. Lferd, A. D. S. Bernabe, G. Nunez, A. Torres-Gonzalez, and A. Ollero, "Cooperation between UAS and wireless sensor networks for efficient data collection in large environments," *Journal of Intelligent and Robotic Systems*, vol. 70, pp. 491–508, 2013.
- [16] D. Zeng, Y. Dai, F. Li, R. S. Sherratt, and J. Wang, "Adversarial learning for distant supervised relation extraction," *Computers, Materials and Continua (CMC)*, vol. 55, no. 1, pp. 243–254, 2018.
- [17] H. Yetgin, K. T. K. Cheung, M. Ei-Hajjar, and L. Hanzo, "Network-lifetime maximization of wireless sensor networks," *IEEE Access*, vol. 3, pp. 2191–2226, 2015.
- [18] R. I. Ogie, "Adopting incentive mechanisms for large-scale participation in mobile crowdsensing: from literature review to a conceptual framework," *Human-Centric Computing and Information Sciences*, vol. 6, no. 1, pp. 1–31, 2016.
- [19] Z. M. Nezhad and S. Khorsandi, "Cooperation enforcement based on dynamic pricing in multi-domain sensor network," in *Proceedings of Consumer communications and networking conference 2011 (CCNC 2011)*, pp. 1055–1060, IEEE, Las Vegas, Nevada, USA, January 2011.
- [20] S. Maity and J. Park, "Powering IoT devices: a novel design and analysis technique," *Journal of Convergence*, vol. 7, 2016.
- [21] D. Yasmine, K. Bouabdellah, and F. K. Mohammed, "Using mobile data collectors to enhance energy efficiency and reliability in delay tolerant wireless sensor networks," *Journal of Information Processing Systems*, vol. 12, pp. 275–294, 2016.
- [22] D. Goyal and M. R. Tripathy, "Routing protocols in wireless sensor networks: a survey," in *Proceedings of second international conference on advanced computing and communication technologies (ACCT 2012)*, pp. 256–275, IEEE, Rohtak, Haryana, India, January 2012.
- [23] M. Li, E. Kamioka, and S. Yamada, "Pricing to simulate node cooperation in wireless Ad hoc networks," *IEICE Transactions on Communications*, vol. E90-B, no. 7, pp. 1640–1650, 2007.
- [24] Y. Yuan and F. Wang, "Blockchain: the state of the art and future trends," *Acta Automatica Sinica*, vol. 42, no. 4, pp. 481–494, 2016.
- [25] Q. Shao, C. Jin, Z. Zhang, W. Qian, and A. Zhou, "Blockchain: architecture and research progress," *Chinese Journal of Computers*, 2017, <http://cjc.ict.ac.cn/online/cre/10xsqfs-2017127145754.pdf>.
- [26] Y. Ren, J. Shen, D. Liu, J. Wang, and J. Kim, "Evidential quality preserving of electronic record in cloud storage," *Journal of Internet Technology*, vol. 17, no. 6, pp. 1125–1132, 2016.
- [27] B. Christian, M. Ueli, T. Daniel, and Z. Vassilis, "Bitcoin as a transaction ledger: a composable treatment," in *Proceedings of 36th annual international cryptology conference—Advances in Cryptology (CRYPTO 2017)*, pp. 324–356, Santa Barbara, CA, USA, August 2017.
- [28] Y. Ren, J. Shen, Y. Zheng, J. Wang, and H. Chao, "Efficient data integrity auditing for storage security in mobile health cloud," *Peer-to-Peer Networking and Applications*, vol. 9, no. 5, pp. 854–863, 2016.
- [29] W. Qian, Q. Shao, Y. Zhu, C. Jin, and A. Zhou, "Research problems and methods in blockchain and trusted data management," *Journal of Software*, vol. 29, pp. 150–159, 2018.
- [30] Y. Tu, Y. Lin, J. Wang, and J. U. K. Kim, "Semi-supervised learning with generative adversarial networks on digital signal modulation classification," *Computers, Materials and Continua (CMC)*, vol. 55, no. 2, pp. 243–254, 2018.
- [31] L. Xiang, Y. Li, W. Hao, P. Yang, and X. Shen, "Reversible natural language watermarking using synonym substitution and arithmetic coding," *Computers, Materials and Continua (CMC)*, vol. 55, no. 3, pp. 541–559, 2018.
- [32] R. Qiao, S. Dong, Q. Wei, and Q. Wang, "Blockchain based secure storage scheme of dynamic data," *Computer Science*, vol. 45, pp. 57–62, 2018.
- [33] R. Meng, S. Rice, J. Wang, and X. Sun, "A fusion steganographic algorithm based on faster R-CNN," *Computers, Materials and Continua (CMC)*, vol. 55, no. 1, pp. 1–16, 2018.
- [34] P. He, G. Yu, Y. Zhang, and Y. Bao, "Survey on blockchain technology and its application prospect," *Computer Science*, vol. 44, pp. 1–8, 2017.
- [35] Y. Ren, J. Shen, J. Wang, J. Han, and S. Lee, "Mutual verifiable provable data auditing in public cloud Storage," *Journal of Internet Technology*, vol. 16, pp. 317–323, 2015.
- [36] D. He, N. Kumar, S. Zeadally, and H. Wang, "Certificateless provable data possession scheme for cloud-based smart grid data management systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 232–242, 2018.
- [37] H. Wang, K. Li, K. Ota, and J. Shen, "Remote data integrity checking and sharing in cloud-based health internet of things," *IEICE Transactions on Information and Systems*, vol. E99.D, no. 8, pp. 1966–1973, 2016.
- [38] Q. Jiang, F. Wei, S. Fu, J. Ma, G. Li, and A. Alelaiwi, "Robust extended chaotic maps-based three-factor authentication scheme preserving biometric template privacy," *Nonlinear Dynamics*, vol. 83, no. 4, pp. 2085–2101, 2016.
- [39] Z. Faheem, A. Khan, S. U. R. Malik et al., "A survey of cloud computing data integrity schemes: design challenges, taxonomy and future trends," *Computers and Security*, vol. 65, pp. 29–49, 2017.
- [40] N. Gargn and S. Bawa, "Comparative analysis of cloud data integrity auditing protocols," *Journal of Network and Computer Applications*, vol. 66, pp. 17–32, 2016.
- [41] K. Gu, W. Jia, and J. Zhang, "Identity-based multi-proxy signature scheme in the standard model," *Fundamenta Informaticae*, vol. 150, no. 2, pp. 179–210, 2017.
- [42] Y. Wang and Q. Wu, "A survey on cryptographic technologies for data integrity checking in clouds," *Journal of Cyber Security*, vol. 2, pp. 23–35, 2017.

