

Received October 29, 2021, accepted December 22, 2021, date of publication December 28, 2021, date of current version January 13, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3139176

# Data Sharing Network Model and Mechanism of Power Internet of Things in Virtualized Environment

SHUNLI ZHANG 

Department of Information Technology and Engineering, Jinzhong University, Jinzhong, Shanxi 030600, China

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

e-mail: 544891117@qq.com


This work was supported by the National Key Research and Development Program of China under Grant 2018YFB1800704.

**ABSTRACT** To solve the problems of low security and low reliability in data sharing, this paper proposes data sharing network model and mechanism of power Internet of Things in virtualized environment. Due to the different coverage of the power Internet of Things, this paper proposes a node model based on the node network theory of blockchain technology. This model divides network nodes into data consumer nodes, data storage nodes, routing node, and coordination nodes according to business requirements. Through the cooperation of the four types of nodes, data sharing between multiple power Internet of Things can be realized efficiently. To solve the problem of low security in data sharing, this paper constructs a data transmission contribution analysis model of network nodes and a data access authorization model. The data transmission contribution degree model can fairly evaluate the behavior of network nodes based on the contribution degree, so as to quickly find malicious nodes. The data access authorization model calculates the entropy weight of each Internet of Things according to the trust level, thereby calculating the credibility of data sharing. To solve the problem of low transmission reliability of network nodes, a set of alternative links is constructed for data sharing routes. With load balancing as the goal, a relative cost value evaluation model is constructed for each route. The shortest route with the best relative cost value is regarded as the optimal data sharing route. In the simulation experiment, the related mechanisms are first improved based on the mechanism of this paper, which verifies that the mechanism of this paper improves the utilization and reliability of network resources in terms of data sharing routing. Secondly, it is verified that the mechanism of this paper improves the success rate and availability rate of network resources in terms of data security sharing.

**INDEX TERMS** Power Internet of Things, data sharing, network model, sharing mechanism.

## I. INTRODUCTION

The Internet of Power Things is a key technology for smart grid. More and more power companies have established their own Internet of Things, and the types and numbers of services provided have increased rapidly. In order to improve the use efficiency of power Internet of Things resources, Network Function Virtualization (NFV) technology has become a key technology for the construction and operation of the Internet of Things. When NFV technology is applied to the power Internet of Things, traditional physical network resources are divided into underlying networks and virtual networks.

The associate editor coordinating the review of this manuscript and approving it for publication was Rentao Gu .

The underlying network is responsible for allocating network resources for the virtual network, and the virtual network is responsible for carrying power services.

Although NFV technology improves the utilization of underlying network resources. However, in order to improve the scale and quality of services, each power company needs to share data with the Internet of Things of other power companies. In this context, the trust mechanism, secure connection, and efficient management of a large number of distributed smart devices in the power Internet of Things have become more and more important. Therefore, how to realize an efficient and safe data sharing mechanism between power companies has become an urgent problem to be solved.

In order to solve the problem of low security in the data management of the Internet of Things, more research is based on blockchain technology to improve the security of data [1]–[4]. In order to avoid the problems of low security and low performance caused by the unified storage of data in blockchain technology, there have been studies using federated learning technology to store data in a distributed manner [5]–[9]. In order to further reduce the problem of low data transmission efficiency caused by federated learning, existing studies have adopted model compression and structure optimization strategies to improve data transmission efficiency [10]–[20]. In order to increase the enthusiasm of data owners to participate in data sharing, existing studies have adopted game theory and auction mechanisms to enhance the incentive mechanism of data sharing enthusiasm [21]–[25].

It can be seen from the existing research and analysis that the research to solve the problem of data sharing mainly adopts two strategies of centralized storage and distributed storage. Ideally, electricity data should be stored on a decentralized hardware and software platform, but current solutions are mainly based on a centralized infrastructure. The centralized architecture has high maintenance costs, low intrusion costs, high possibility of data being intercepted and tampered with, and single points of failure for security threats. The use of blockchain technology to propose a decentralized architecture for the Internet of Things system can effectively reduce the hidden danger of a single point of failure in the system and make the Internet of Things data more secure and credible. However, the use of blockchain to construct the Internet of Things data management and control network, the data relationship between each blockchain is independent, which causes greater difficulties in data sharing.

In order to realize the safe sharing of data between multiple data sources, this article contributes as follows: (1). Data management and control network model: propose a data sharing network model for multiple domains, including data use nodes, data transmission nodes, data storage nodes, and data sharing coordination nodes. (2). The authority of data sharing: In order to solve the security problems of different security level domains when crossing authentication, a trust transfer relationship model is proposed to realize the transfer and evaluation of the trustworthiness of different IoT systems. (3). The malicious node aspect of data sharing: In order to solve the problem that malicious nodes report false data resources and false network resources to the controller in the domain, resulting in unavailability of data or failure of resource allocation algorithms due to insufficient underlying resource capacity, the node credit evaluation mechanism are proposed. (4). Network congestion of data sharing: In order to solve the problem of data transmission failure caused by unbalanced network resource load, the resource utilization of the underlying nodes and underlying links is analyzed, so as to achieve load balance between the underlying nodes and underlying links.

## II. RELATED WORK

With the explosive growth of the data collected by the Internet of Things, the research on the data sharing of the Internet of Things has received extensive attention from the academic community.

The decentralization, anonymization, traceability and non-tamper ability of blockchain technology make it a very attractive technology to be introduced into the Internet of Things data sharing to solve the trust problem of Internet of Things users. For example, in order to solve the problem of sharing medical data in a trustless environment, literature [1] uses blockchain technology to provide data sources, auditing and control for the shared medical data between big data entities. Aiming at the problem of data security sharing in multi-cloud platforms, literature [2] proposed a reliable collaboration model based on blockchain and smart contracts. Literature [3] proposed an efficient data collection and secure sharing scheme based on blockchain, combining Ethereum blockchain and deep reinforcement learning to create a reliable and secure environment. Literature [4] combines blockchain and support vector machines to propose a data training program that can protect the privacy of IoT data.

In the blockchain-based data sharing scheme, the data owner encrypts the data that needs to be shared and sends it to the blockchain for storage. If the amount of data contained in the transaction is too large, it will also have a negative impact on the performance of the entire blockchain network. Therefore, in the blockchain-based IoT data sharing, storing the shared data on the blockchain is not a feasible solution. Federated Learning (FL) was proposed by Google in 2017 and follows the principle of informed collection or data minimization [5]. In many fields, federated learning has broad research value and application prospects [6].

For example, in order to solve the problem of unbalanced air quality data distribution and waste of computing resources, literature [7] established a regional model of weighted value in federal learning by dividing air quality data into regions. In order to solve the problem of data privacy protection and the inability of large-scale collaborative training, literature [8] proposed a credit card fraud detection method with privacy protection based on federated learning, which uses federated learning to construct a global shared fraud detection method. Literature [9] uploads the data features and model parameters to the federated learning central server at the same time, and provides a feature fusion strategy for each client at the same time, and uses the echo state network to achieve precise trend tracking.

In order to mobilize data owners to actively participate in data sharing, it is an effective measure to encourage users to actively participate in IoT data sharing through an effective incentive mechanism. At present, there have been related researches that introduce incentive mechanisms into mobile group perception or resource trading and other Internet of Things application scenarios, and use this to encourage users to participate in data sharing behavior [10]–[12].

Existing incentive mechanisms are divided into centralized IoT data incentive sharing and distributed IoT data incentive sharing.

In the centralized Internet of Things data incentive sharing party, literature [13] proposed a recruitment strategy for vehicle-mounted ad hoc network data sharing participants suitable for vehicle trajectory prediction, which minimizes the total recruitment cost. Literature [14] proposes a mobile group perception framework based on fog computing, which solves the security and privacy issues between task requesters and workers. Literature [15] studied the data sharing problem of the Internet of Vehicles based on game theory, and used the Q-Learning algorithm to realize the vehicle compensation payment strategy.

In terms of incentives in distributed scenarios, the decentralized nature of the blockchain determines that data storage is a distributed storage method. Therefore, the incentive mechanism in distributed scenarios is a research focus. Literature [16] uses Bayesian inference model to design a blockchain-based trust management system. Literature [17] implements a blockchain-based mobile group awareness system that supports task requesters to directly send tasks to workers, avoiding the involvement of traditional centralized trusted third-party platforms. Literature [18] proposes a real incentive mechanism that can meet the different resource allocation needs of IoT users in a dynamic and distributed P2P environment. Literature [19] proposed a local P2P power resource sharing model to support local power trading transactions between hybrid electric vehicles.

In the exploration process of building a secure distributed sharing system based on blockchain, the data transmission rate in the network is a necessary condition for data sharing. In order to efficiently realize the large-scale data sharing market composed of billions of Internet of Things devices, we must consider improving network performance as much as possible [20]. During the data sharing process, the data node transmits local data to the server, and the central server transmits global data to each local node. In the two-way communication process, the potentially uncontrollable network state (network delay, communication cost, etc.) of the local data node can easily cause communication to become a bottleneck for data sharing.

In order to increase the data sharing rate, two strategies are usually adopted: optimization model and optimization system architecture. In terms of optimization model, literature [21] calculates the average value from low communication frequency to save communication delay and improve convergence speed. In terms of optimizing the architecture, literature [22] proposed a federated learning framework based on cloud-side collaboration in order to reduce the communication overhead between the backbone network and the central server, which reduced the data transmission volume of the backbone network. Literature [23] proposes a decentralized federated learning framework combined with blockchain. Blockchain smart contracts help train the computing nodes of the network to reach a consensus without

using a central server to complete the global model aggregation. Literature [24] proposed a decentralized federated learning framework for peer-to-peer networks to solve the problem of participants sharing the work of the central server aggregation model. Literature [25] proposed a framework for decentralized federated learning and a decentralized model aggregation algorithm based on a 5G network that supports device-to-device communication.

It can be seen from the existing research and analysis that the research to solve the problem of data sharing mainly adopts two strategies of centralized storage and distributed storage. The centralized architecture has high maintenance costs, low intrusion costs, high possibility of data being intercepted and tampered with, and single points of failure for security threats. The use of blockchain technology to propose a decentralized architecture for the Internet of Things system can effectively reduce the hidden danger of a single point of failure of the system. However, the use of blockchain to construct the Internet of Things data management and control network, the data relationship between each blockchain is independent, which causes greater difficulties in data sharing. This paper aims to solve the problems of low security and low reliability in power Internet of Things data sharing in a multi-domain environment, and proposes a power Internet of Things data sharing network model and mechanism in a virtualized environment. By constructing a data management and control network model and a trust transfer relationship model, a heuristic data transmission routing strategy and a network node credit evaluation mechanism are proposed to improve the security and reliability of the data in the data sharing process.

### III. PROBLEM DESCRIPTION

In order to facilitate the description of the data sharing network model and mechanism. This section explains the concepts related to the data sharing network model and mechanism of the power Internet of Things in a virtualized environment.

The user requesting to share data is called the data consumer. The user who provides the data is called the data provider. Considering that blockchain technology is an important technology and development trend of current data storage, this article uses blockchain technology to store data. The end-to-end communication path from data consumer node to data provider node is called data sharing route. The equipment in the data sharing route includes data consumer nodes, data provider nodes, network transmission equipment, network load balancing equipment, and network security equipment. In the network function virtualization environment, these devices can be constructed by general-purpose servers using virtualization technology.

The network equipment composed of general servers is called the underlying network, which is represented by  $G = (V, E)$ . The underlying network includes underlying nodes and underlying links. Use  $V$  to represent the bottom-level node set, and use  $u \in V$  to represent

a bottom-level node. Use  $E$  to represent the underlying link collection. Use  $e_{uv} \in E$  to represent the underlying link with  $u \in V$  and  $v \in V$  as the endpoint. Considering that node resources can be dynamically migrated and expanded through virtualization technology, node resources are not restricted. This paper mainly studies the resource management model under the condition of limited network link resources. The bandwidth resource of the underlying network link  $e_{uv} \in E$  is represented by  $C_{uv}^{bw}$ , and the link resource availability rate is represented by  $r_{uv}^{bw}$ .

Through NFV technology, the underlying nodes can generate multiple NFV instances (Network Function Virtualization Instance, NFVI) according to data management and sharing requirements. Each bottom node can carry multiple NFVIs. Use  $m_{NFVI} \in M_{NFVI}$  to represent a type of NFVI.  $M_{NFVI}$  represents the set of NFVI carried on the underlying network. For ease of description, the nodes and links included in Data Sharing Routing (DSR) are called virtual nodes and virtual links, respectively, and represented by  $G^{DSR} = (V^{DSR}, E^{DSR})$ .  $V^{DSR}$  represents a collection of virtual nodes on the data sharing route, and  $u^{DSR} \in V^{DSR}$  represents a virtual node on the data sharing route.  $E^{DSR}$  represents the collection of virtual links on the data sharing route.  $e_{uv}^{DSR} \in E^{DSR}$  represents the virtual link on the data sharing route. The start node and end node of the data sharing route  $G^{DSR}$  are represented by  $s^{DSR}$  and  $t^{DSR}$  respectively. Use  $Q_{e_{uv}^{DSR}}^{bw}$  to indicate the bandwidth resource that needs to be used in data sharing route  $e_{uv}^{DSR} \in E^{DSR}$ .

#### IV. DATA SHARING NETWORK MODEL AND MECHANISM

##### A. DATA SHARING NETWORK MODEL

In order to solve the security and efficiency issues in data sharing in a cross-domain environment, this paper proposes a network model for data sharing as shown in Fig. 1. As can be seen from the figure, the network model of data sharing includes multiple domain networks. The domain network refers to the network of each power company. Each domain network has an independent security domain and management domain. Each power company can have one or more domain networks.

Each domain network in the model is implemented using blockchain technology. Each domain network includes four types of nodes: data user nodes, data storage nodes, network nodes, and coordination nodes. The main function of the data user node is the submission of data requirements or information query. The data storage node is responsible for data user authority judgment and data provision. In order to achieve data consistency, data storage nodes need to synchronize all blockchain data information, and generally have strong storage capabilities.

The network node is responsible for data transmission. The network node is responsible for connecting the data user node with the data storage node. In order to realize data sharing among multiple domain networks, each domain network elects a data storage node as a coordinating node. The main

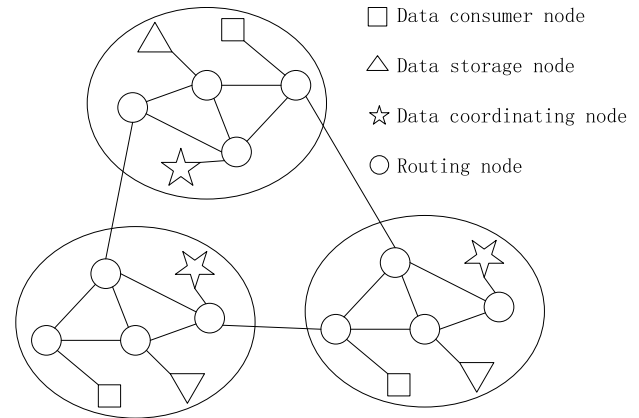


FIGURE 1. Data sharing network model.

function of the coordinating node is to remove the barriers to the isolation of the domain network. The coordinating node is a node trusted by an independent domain network and operated by the domain network. Multiple coordination nodes are connected to each other, and multiple independent domain networks are grouped into a network that can realize data sharing. The coordination node can implement routing information and identity authentication information of all domain networks in the domain network. Whenever a new domain network is added, the domain network selects a coordinating node and connects to the coordination node of the adjacent domain network. After initialization, the domain network can communicate with other domain networks. In summary, the coordination node is responsible for receiving data user data requests, data user authority calculation, data sharing routing arrangement, and network node trust management. The data information can be expressed in an existing DNS mode, or can be realized in an IP address mode. The information update mechanism in the coordination node can learn from the OSPF routing mechanism.

##### B. DATA SHARING NETWORK MECHANISM

The data sharing mechanism is shown in Table 1. The mechanism includes seven steps: submitting a data sharing request, querying the location of the data in the domain, querying the location of the data provider in the external domain, making a data sharing request to the external domain, determining the authority of the data requester, sending the data to the data requesting party, and evaluating the credibility of the node.

(1). The data consumer node submits a data sharing request: the data consumer node submits a data sharing request to the data storage node, and the content of the request is the basic information of the data that needs to be shared.

(2). The data storage node executes the query: the data storage node judges whether it belongs to the local network according to the basic information of the required data. If it does not belong to the local area network, the request needs to be forwarded to the coordinating node through the network node.

**TABLE 1.** Data sharing mechanism.

- 
- (1). The data consumer node submits a data sharing request.
  - (2). The data storage node searches for the location of the data.
  - (3). The coordination node finds the shortest and most reliable route to the data provider.
  - (4). The coordination node sends a data sharing request to the target domain network node.
  - (5). The target domain network coordination node judges the access authority of the data user.
  - (6). The data provider node sends the data to the data requester node.
  - (7). The coordinating node evaluates the credibility of the network nodes participating in the data sharing based on the degree of contribution.
- 

(3). The coordination node finds the shortest and most reliable route to the data provider. The coordination node finds the shortest and most reliable route to the data provider based on the basic information of the required data, the existing routing and cross-domain interaction status.

(4). The coordinating node sends a data sharing request to the target domain network node: First, the coordinating node finds the optimal path according to the routing protocol. Second, the coordination node sends a data sharing request to the target domain network node, including its own identity information and required data.

(5). The target domain network coordination node judges the access authority of the data user. After the coordinating node of the domain network where the data provider is located receives the data sharing request, it determines whether the data user has the authority to obtain the data; if there is no data access authority, it returns no authority. If there is data access authority, according to the existing routing, information such as the data access request and the routing to the data user will be forwarded to the data storage node.

(6). The data provider node sends the data to the data requester node. After receiving the data access request sent by the coordinating node, the data is sent to the data user node according to the information such as the route to the data user.

(7). Coordinating nodes evaluate the credibility of participating data sharing nodes based on their contribution. The coordination node in the blockchain where the data user is located evaluates the data availability, completeness and accuracy of the data providing node; evaluates the reliability and availability of the network node. The coordinating node in the blockchain where the data providing node is located evaluates whether the data requesting node leaks data and whether the data is used legally; evaluates the reliability and availability of the network node. The coordinating node stores the evaluation results in the attributes of each node.

When coordinating nodes find the shortest and most reliable route to the data provider, they first need to delete malicious nodes from the optional network nodes, and

secondly, select the route with the shortest route and meet the trust level as the link used for the final data sharing route resource. From the process of the data sharing mechanism, it can be seen that in the data sharing mechanism, the factors that have a greater relationship with data sharing include the shortest route search, data access authority management, and malicious node management. The shortest path search mainly solves the problems of data transmission speed and reliability. Data access authority management mainly solves the problems of data leakage, loss, and tampering caused by illegal access to data. Malicious node management mainly solves the damage of malicious nodes in the network to the network environment, causing the data sharing process to fail. The three processes are described in detail below.

## V. THE KEY PROCESS OF DATA SHARING

### A. DATA TRANSMISSION ROUTING

Data transmission speed and reliability play a key role in the speed of data sharing. If the speed and reliability of the data transmission network are poor, network congestion will occur in data sharing routing. In order to solve the problem of network congestion during data sharing, it is necessary to arrange the shortest and most reliable route for each data sharing route. Excessive resource utilization can easily lead to network congestion. This article takes load balancing as the goal and arranges data transmission routes.

In order to select the best data sharing route, the relative cost is used to measure the pros and cons of each data sharing route, so as to select the best data sharing route. By evaluating the relative cost of each link, the relative cost value of the data sharing route can be easily calculated. Use  $c_{i,uv}^{bw}$  to represent the relative cost value of the underlying link  $e_{uv} \in E$ . When the bandwidth required by the data sharing route is  $Q_{e_{uv}}^{bw,DSR}$ , formula (1) can be used to calculate the relative cost value of each underlying link in the potential data sharing route. In the formula,  $\max_{uv \in E} C_{uv}^{bw}$  represents the maximum bandwidth resource of the link resource in the current network.  $r_{uv}^{bw} C_{uv}^{bw}$  represents the amount of available bandwidth of the underlying link  $e_{uv} \in E$ . It can be seen from formula (1) that the relative cost value of the underlying link resources is inversely proportional to its resource availability. The greater the link resource availability rate, the smaller the relative cost value of link resource allocation. Therefore, the smaller the relative cost value of data sharing routing, the lower the resource utilization rate of the network, the higher the reliability, and the stronger the load balancing ability of network resources.

$$c_{i,uv}^{bw} = \frac{\max_{uv \in E} C_{uv}^{bw}}{r_{uv}^{bw} C_{uv}^{bw} - Q_{e_{uv}}^{bw,DSR}} \quad (1)$$

In order to achieve the optimization of data sharing routing, this paper adopts heuristic data transmission routing strategy. This strategy can achieve the goal of optimizing the overall cost value of data sharing routing by solving the link strategy with the smallest relative cost value one by one.

The heuristic data transmission routing strategy includes the following three steps. (1) Select all the underlying links that meet the bandwidth requirements of the data sharing route to form a set of candidate links; (2) From the set of candidate links, use the shortest path algorithm to solve the first  $n$  shortest path allocation strategies; (3) Choose the bottom path with the smallest relative cost among the  $n$  shortest paths as the optimal data sharing routing scheme. The relative cost value  $C_{i,u}^{bw,DSR_v,DSR}$  of the data sharing route DSR is calculated using formula (2).  $z_{uv}^{u,DSR_v,DSR}$  represents the variable of whether the virtual link  $e_{uv}^{DSR} \in E^{DSR}$  passes through the underlying link  $e_{uv} \in E$ .  $z_{uv}^{u,DSR_v,DSR} = 1$  means passing,  $z_{uv}^{u,DSR_v,DSR} = 0$  means no passing.

$$C_{i,u}^{bw,DSR_v,DSR} = \sum_{uv \in E} C_{i,uv}^{bw} z_{uv}^{u,DSR_v,DSR} \quad (2)$$

### B. DATA ACCESS AUTHORITY MANAGEMENT

In data sharing, data leakage, loss, tampering and other issues are the main issues that must be resolved. Only by solving these problems can data be shared safely. This article solves this problem from the trust level (TL) of each domain network.

This article adopts a third-party organization certification strategy to assign identity permissions to each domain network that joins data sharing. Data sharing cannot be carried out on a domain network that has not obtained identity authority. The identity authority assigned by the third-party organization to the domain network includes five levels. Considering that the trust degree needs to be calculated when multiple domains share data, this paper assigns a trust interval for each trust degree level. The five levels of trust are  $TL_1 = 0$ ,  $0 < TL_2 \leq 0.25$ ,  $0.25 < TL_3 \leq 0.5$ ,  $0.5 < TL_4 \leq 0.75$ , and  $0.75 < TL_5 \leq 1$ .

As the number of Internet of Things increases, the data of users may be stored in a remote domain network. At this time, data users and data providers need to communicate through two or more domain networks. When data users and data providers pass through multiple domain networks, use  $TL_i$  to represent the cross-domain credibility of the  $i$ -th path. To simplify the solution of cross-domain credibility  $TL_i$ , the entropy weight method is used to calculate the entropy value  $e_j$  of each domain network, as shown in formula (3).

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n \frac{TL_{ij}}{\sum_{i=1}^n TL_{ij}} \ln \frac{TL_{ij}}{\sum_{i=1}^n TL_{ij}} \quad (3)$$

Among them,  $TL_{ij}$  represents the identity authentication credibility of mutual trust between domain network  $i$  and domain network  $j$ . The value of  $TL_{ij}$  is the lowest credibility value of the two domain networks. Based on this, the adjacent domain network credibility matrix  $TL$  composed of all domain networks is shown in formula (4). The number of rows  $m$  is the number of data users who have cross-domain network data sharing requests, and the number of

columns  $n$  is the number of domain networks.

$$TL = \begin{bmatrix} TL_{11} & \dots & TL_{1n} \\ \dots & \dots & \dots \\ TL_{m1} & \dots & TL_{mn} \end{bmatrix} \quad (4)$$

The credibility weight of  $n$  domain networks is denoted as  $W = [w_1, \dots, w_j, \dots, w_n]$ . At this time, use formula (5) to calculate the weights  $w_j$  between the two domain networks, where  $0 \leq w_j \leq 1$ , and  $\sum_{j=1}^n w_j = 1$ .

$$w_j = \frac{1 - e_j}{\sum_{j=1}^n (1 - e_j)} \quad (5)$$

Therefore, the cross-domain credibility  $TL_i$  of each data sharing route is shown in formula (6).

$$TL_i = \sum_{j=1}^n w_j \frac{TL_{ij}}{\sqrt{\sum_{j=1}^m (TL_{ij})^2}} \quad (6)$$

### C. MALICIOUS NODE IDENTIFICATION MECHANISM

As IoT devices are distributed in a very wide area, some network nodes are vulnerable to attacks by viruses, hackers, etc., and thus become malicious nodes. In the data sharing network, these malicious nodes pose a threat to the security of the data sharing network. In order to quickly identify malicious nodes, this paper uses the contribution of each network node in the data network to the data sharing routing activity to evaluate whether the network node is a malicious node.

The relevant characteristics of different nodes are different during evaluation. For data requesting nodes and data providing nodes, evaluate whether the nodes provide and use data as required to analyze credibility. The coordinating node in the blockchain where the data user is located evaluates the data availability, completeness, and accuracy of the data providing node, and evaluates the reliability and availability of the network node. The coordinating node in the blockchain where the data providing node is located, evaluates whether the data requesting node leaks data, legally uses the data, and evaluates the reliability and availability of the network node. For the coordination node, evaluating whether the task is completed is the evaluation basis. For network nodes, the trust of the nodes that contribute to the route is increased. This paper takes the trustworthiness evaluation of network nodes as the research object.

In order to facilitate the evaluation of the contribution of each network node, the initial contribution of each network node  $v_i \in V$  is set to the same value, which is represented by  $k_i$ . If the network node  $v_i \in V$  belongs to the node on the data sharing route  $G_j^{DSR}$ , use  $T_{G_j}^{t,u_i}$  to represent the contribution of the network node  $v_i \in V$  to the data sharing route  $G_j^{DSR}$  at time  $t$ , and use formula (7) to calculate.

$$T_{G_j}^{t,u_i} = F(T_{G_j}^{t-1,u_i}, G_j^{DSR}) \quad (7)$$

$F(T_{G_j}^{t-1,u_i}, G_j^{DSR})$  represents the update function of the contribution degree of the network node, which is calculated

using formula (8). Among them,  $RW_{G_j^{t-1,DSR}} = yes$  represents network node  $v_i \in V$  provides useful work for data sharing route  $G_j^{DSR}$  at time t-1, and the reward at this time is  $\lambda_{yes}$ .  $Map_{f_j^{t-1,SFC}} = no$  indicates that at t-1, network node  $v_i \in V$  provides harmful work for data sharing route  $G_j^{DSR}$ , and the penalty at this time is  $\lambda_{no}$ . When coordinating nodes to construct a data sharing route, if the contribution of a certain network node is less than the threshold  $k^{Th}$ , the network node is regarded as a potential malicious node, and the network node cannot participate in the work of the data sharing route.

$$F(T_{G_j}^{t-1,u_i}, G_j^{DSR}) = \begin{cases} T_{G_j}^{t-1,u_i} + \lambda_{yes}, & RW_{G_j^{t-1,DSR}} = yes \\ T_{G_j}^{t-1,u_i} - \lambda_{no}, & RM_{G_j^{t-1,DSR}} = no \end{cases} \quad (8)$$

## VI. PERFORMANCE ANALYSIS

### A. NETWORK ENVIRONMENT

Because the allocation of data sharing routing resources is the key task of data sharing. The pros and cons of data sharing routing resource allocation play a key role in the success rate, security, and reliability of data sharing. Since the success rate of data sharing routing resource allocation is related to the utilization rate of the underlying network resources, this paper verifies the algorithm performance from two dimensions: the success rate of data sharing routing resource allocation and the utilization rate of underlying network resources. In order to verify the effect of the algorithm in this paper on the allocation of data sharing routing resources, the two dimensions of data transmission routing mechanism and malicious node identification mechanism are used to compare this paper with traditional algorithms. Through the analysis of traditional algorithms, the comparison algorithm selected in this paper is based on the data sharing routing resource allocation algorithm based on shortest path (DSRRAAoSP). The path allocated by algorithm DSRRAAoSP for data sharing routing is the shortest bottom path.

In order to realize the algorithm performance analysis, the GT-ITM [26] tool is used to generate the network topology environment and simulate the process of data sharing routing and resource allocation. The network topology environment includes five domain networks, each domain network including the number of bottom-layer nodes is uniformly distributed between (30,70), and the number of bottom-layer links is uniformly distributed between (60,100). The bandwidth resource of each underlying link is 1000 Mbps. The start node and end node of each data sharing route are randomly selected from five domains. The link bandwidth resources of data sharing routing obey the uniform distribution between (10,30). The life cycle of each data sharing route is 3 time units. In terms of performance analysis indicators, the two dimensions of the success rate of data sharing routing resource allocation and the availability of underlying network resources are used for analysis. The resource allocation success rate is calculated by dividing the number of data sharing routes that successfully obtain resources by the number of data sharing routes that apply for resources. The availability rate of the underlying

network resources is calculated by dividing the amount of available bandwidth of the underlying network resources by the total bandwidth data of the underlying network resources.

### B. ANALYSIS OF DATA TRANSMISSION ROUTE ARRANGEMENT MECHANISM

In order to analyze the performance of the data transmission routing arrangement mechanism, the data transmission routing arrangement algorithm of this article is added on the basis of the algorithm DSRRAAoSP to obtain the shortest and most reliable routing. The modified algorithm is represented by DSRRAAoSPRC (data sharing routing resources allocation algorithm based on shortest path and relative cost).

The comparison results of the resource allocation success rate of the algorithm before and after optimization is shown in Fig. 2. In the figure, the X axis indicates that the number of data sharing routes has increased from 100 to 3000. The Y-axis represents the success rate of the underlying network in allocating resources for data sharing routing. It can be seen from the figure that as the number of data sharing routes increases, the success rate of resource allocation gradually decreases. This is because the increase in the number of data sharing routes requires more underlying network resources. When the number of data sharing routes gradually increases, the number of resources of the underlying links becomes less and less, which cannot meet the bandwidth resource requirements of data sharing routes. In terms of performance analysis of the two algorithms, the algorithm DSRRAAoSPRC is higher than the algorithm DSRRAAoSP. This is because the algorithm DSRRAAoSPRC takes the relative cost of the link as an element of link selection, thus realizing the load balancing of the underlying link resources and avoiding the scarcity of some key link resources.

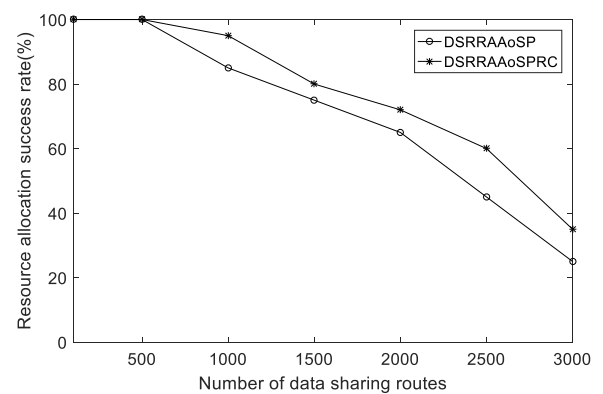


FIGURE 2. Comparison of the success rate of resource allocation.

The comparison of the availability of underlying link resources is shown in Fig. 3. The X axis indicates that the amount of available bandwidth of the underlying link has increased from 300 Mbps to 1500 Mbps. The Y-axis represents the availability of underlying link resources. When the number of data sharing routes is 1000, the comparison of the availability of underlying link resources under the

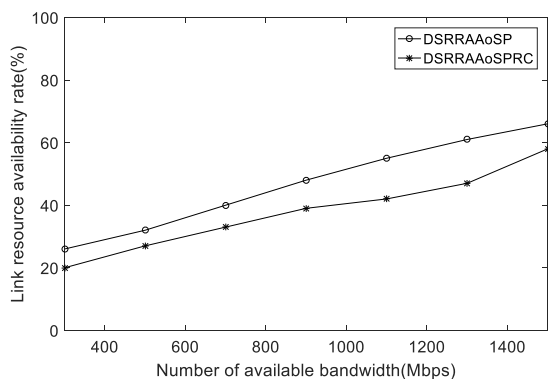


FIGURE 3. Comparison of the availability of underlying link resources.

two algorithms is shown in Fig. 3. It can be seen from the figure that the number of available bandwidths of the underlying link increases, the availability of the underlying link resources under the two algorithms is increasing. This is because the increase in the number of underlying link bandwidths can allocate more optimized underlying link resources for data sharing routing. In terms of comparison of the two algorithms, the algorithm DSRRAAoSPRC allocates underlying link resources for more data sharing routes, so the availability of underlying link resources is low. In contrast, the traditional algorithm DSRRAAoSP allocates underlying link resources for fewer data sharing routes, so the availability of underlying link resources is higher.

C. ANALYSIS OF MALICIOUS NODE IDENTIFICATION MECHANISM

On the basis of DSRRAAoSPRC, the malicious node identification mechanism of this article is added to obtain a safe and reliable route, which is represented by DSRRAAoSPRCMNI (data sharing routing resources allocation algorithm based on shortest path and relative cost and malicious node identification). The malicious node is simulated by exaggerating the bandwidth of all underlying links connected to it. When a malicious node exaggerates its link resources, the underlying link cannot complete data transmission.

Fig. 4. compares the success rate of resource allocation when malicious nodes interfere. The X axis indicates that the number of data sharing routes has increased from 100 to 3000. The Y axis represents the success rate of resource allocation when malicious nodes interfere. Fig. 4. shows the comparison result of the resource allocation success rate when the bandwidth resource amount of each underlying link is 1000 Mbps. It can be seen from the figure that the number of data sharing routes increases, the success rate of data sharing route resource allocation under the two algorithms is decreasing. This is because the increase in the number of data sharing routes requires more underlying link resources for resource allocation. When the utilization rate of the underlying link resources is too high, more data sharing routes cannot obtain the required underlying link resources, resulting in a gradual

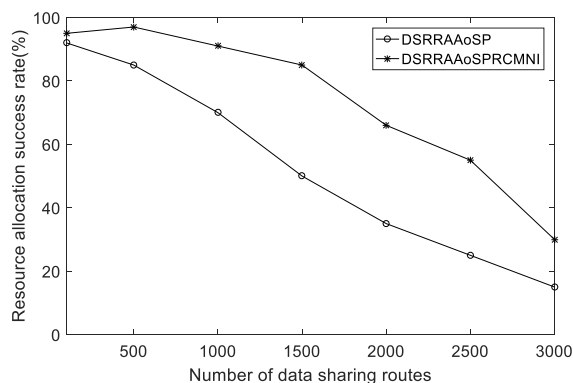


FIGURE 4. Comparison of the success rate of resource allocation when malicious nodes interfere.

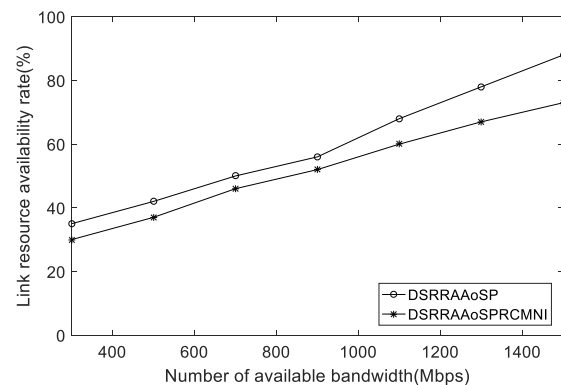


FIGURE 5. Comparison of the availability of underlying link resources when malicious nodes interfere.

decrease in the success rate of resource allocation. In comparison of the two algorithms, the resource allocation success rate of the optimized algorithm DSRRAAoSPRCMNI is higher than that of the algorithm DSRRAAoSP. This is because the algorithm DSRRAAoSPRCMNI deletes malicious nodes before resource allocation, thereby improving the success rate of resource allocation. The algorithm DSRRAAoSP allocates links to malicious nodes in the resources allocated to the data sharing route, which leads to the failure of resource allocation.

Fig. 5. shows the comparison result of the availability of underlying link resources when malicious nodes interfere. The X-axis indicates that the amount of available bandwidth of the underlying link increases from 300Mbps to 1500Mbps. The Y-axis represents the availability of underlying link resources when malicious nodes interfere. Fig. 5. shows the availability of underlying link resources under the two algorithms when the number of data sharing routes is 1000. It can be seen from the figure that the number of available bandwidth of the underlying link increases, the availability of underlying link resources is increasing. This is because the number of underlying link resources increases, and the number of idle underlying link resources increases. In terms of comparison between the two algorithms, the optimized algorithm DSRRAAoSPRCMNI has a higher underlying link



resource availability rate than the algorithm DSRRAoSP. This is because the algorithm DSRRAoSPRCMNI allocates underlying link resources for more data sharing routes. More data sharing routes occupy more underlying link resources.

## VII. CONCLUSION

With the rapid increase in the amount of power Internet of Things data, data sharing between power companies has become an important task. The main problems in data sharing of power Internet of Things in a multi-domain environment are low data security and low network reliability. To solve this problem, this paper proposes a data sharing network model and mechanism for the Internet of Things in a virtualized environment. In the performance analysis section, the algorithm in this paper is compared with the traditional algorithm. From the comparison results of the data transmission routing arrangement mechanism, it can be seen that the algorithm in this paper allocates resources based on the relative cost of network resources. Network resources have better load balancing characteristics, which improves the reliability of data sharing networks. From the comparison results of the malicious node identification mechanism, it can be seen that the algorithm in this paper can identify malicious nodes, avoid the unavailability of the data sharing network caused by malicious nodes, and improve the reliability of the data sharing network. This article improves the security and reliability of the data sharing network from the dimensions of data security and network reliability. Because the enthusiasm of data sharing participants is a necessary condition for data sharing work. In the next step, based on the research results of this article, game theory and auction mechanism are used to study data sharing issues from the perspective of economics, and to promote the healthy development of data sharing work.

## REFERENCES

- [1] Q. I. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani, "MeDShare: Trust-less medical data sharing among cloud service providers via blockchain," *IEEE Access*, vol. 5, pp. 14757–14767, 2017.
- [2] M. Shen, J. Duan, L. Zhu, J. Zhang, X. Du, and M. Guizani, "Blockchain-based incentives for secure and collaborative data sharing in multiple clouds," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 6, pp. 1229–1241, Jun. 2020.
- [3] C. H. Liu, Q. Lin, and S. Wen, "Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3516–3526, Jun. 2019.
- [4] M. Shen, X. Tang, L. Zhu, X. Du, and M. Guizani, "Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7702–7712, Oct. 2019.
- [5] Y. Chen, X. Sun, and Y. Jin, "Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 4229–4238, Oct. 2020.
- [6] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, Feb. 2019.
- [7] B. Hu, Y. Gao, L. Liu, and H. Ma, "Federated region-learning: An edge computing based framework for urban environment sensing," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.
- [8] W. Yang, Y. Zhang, K. Ye, L. Li, and C.-Z. Xu, "FFD: A federated learning based method for credit card fraud detection," in *Proc. Int. Conf. Big Data*. Cham, Switzerland: Springer, 2019, pp. 18–32.
- [9] Y. Hu, X. Sun, Y. Chen, and Z. Lu, "Model and feature aggregation based federated learning for multi-sensor time series trend following," in *Proc. Int. Work-Confer. Artif. Neural Netw.* Cham, Switzerland: Springer, 2019, pp. 233–246.
- [10] N. Yoshida, T. Nishio, M. Morikura, K. Yamamoto, and R. Yonetani, "Hybrid-FL for wireless networks: Cooperative learning mechanism using non-IID data," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.
- [11] Y. Sarikaya and O. Ercetin, "Motivating workers in federated learning: A Stackelberg game perspective," *IEEE Netw. Lett.*, vol. 2, no. 1, pp. 23–27, Mar. 2020.
- [12] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "Adaptive federated learning in resource constrained edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 3, pp. 1205–1221, Jun. 2019.
- [13] X. Wang, W. Wu, and D. Qi, "Mobility-aware participant recruitment for vehicle-based mobile crowdsensing," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4415–4426, May 2018.
- [14] J. Ni, A. Zhang, X. Lin, and X. S. Shen, "Security, privacy, and fairness in fog-based vehicular crowdsensing," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 146–152, Jun. 2017.
- [15] L. Xiao, T. Chen, C. Xie, H. Dai, and V. Poor, "Mobile crowdsensing games in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1535–1545, Feb. 2017.
- [16] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1495–1505, Apr. 2019.
- [17] N. More and D. Motwani, "A blockchain-based decentralized framework for crowdsourcing," in *Proc. Int. Conf. Image Process. Capsule Netw.* Cham, Switzerland: Springer, 2020, pp. 448–460.
- [18] Y. He, H. Li, X. Cheng, Y. Liu, C. Yang, and L. Sun, "A blockchain based truthful incentive mechanism for distributed P2P applications," *IEEE Access*, vol. 6, pp. 27324–27335, 2018.
- [19] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang, and E. Hossain, "Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3154–3164, Dec. 2017.
- [20] T. Vogels, S. P. Karimireddy, and M. Jaggi, "PowerSGD: Practical low-rank gradient compression for distributed optimization," in *Proc. Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2019, pp. 14259–14268.
- [21] M. Mohri, G. Sivek, and A. T. Suresh, "Agnostic federated learning," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 4615–4625.
- [22] S. Savazzi, M. Nicoli, and V. Rampa, "Federated learning with cooperating devices: A consensus approach for massive IoT networks," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4641–4654, May 2020.
- [23] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-i.i.d. data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3400–3413, Sep. 2020.
- [24] N. Yoshida, T. Nishio, M. Morikura, K. Yamamoto, and R. Yonetani, "Hybrid-FL for wireless networks: Cooperative learning mechanism using non-IID data," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.
- [25] J. Kang, Z. Xiong, D. Niyato, H. Yu, Y.-C. Liang, and D. I. Kim, "Incentive design for efficient federated learning in mobile networks: A contract theory approach," in *Proc. IEEE VTS Asia Pacific Wireless Commun. Symp. (APWCS)*, Aug. 2019, pp. 1–5.
- [26] E. W. Zegura, K. L. Calvert, and S. Bhattacharjee, "How to model an internet network," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, vol. 2, Mar. 1996, pp. 594–602.



**SHUNLI ZHANG** received the M.S. degree in computer science from the Taiyuan University of Technology, China, and the Ph.D. degree in computer science from the Beijing University of Posts and Telecommunications, in 2012. He is currently a Postdoctoral Researcher and a Senior Engineer of Jinzhong University, Shanxi, China. He has authored journal and conference papers about the resource management and fault diagnosis of network. His research interests include management for virtual networks and next generation networks.

• • •