



# Design and Implementation of Power Data Traceability System Based on Blockchain

Fangjian Shang<sup>1</sup>, Xinyi Li<sup>1</sup>, Xin Wen<sup>1</sup>, Shaoyong Guo<sup>2</sup>, Chang Liu<sup>2</sup>,  
and Haiyang Hu<sup>2</sup>(✉)

<sup>1</sup> State Grid Jibei Information and Telecommunication Company, Beijing, China

<sup>2</sup> Beijing University of Posts and Telecommunications, Beijing 100876, China  
oceanhu@bupt.edu.cn

**Abstract.** With the development of new power system, the construction of green energy production and consumption market system has been recognized by everyone. And the key step is to trace the source of green power transactions and provide the society with credible green power certificates. At present, the power data traceability system monitors the key processes of green power from production to consumption, but there are some problems, such as difficult to identify the authenticity of power data, unreliable transaction records, imperfect supervision and so on. To solve the above problems, this paper provides a trusted traceability scheme based on blockchain for power data traceability. Firstly, this paper analyzes the power data traceability scenario and proposes a power data traceability model. Then, according to the characteristics of hyperledger fabric, this paper proposes a data access control method based on attribute encryption to realize the data access rights management of different roles. Then, this paper puts forward corresponding identification schemes for different data identification, and constructs a multi-dimensional data model of traceable data. Finally, this paper implements and tests the power data traceability system, which proves that the solution has good performance and high security.

**Keywords:** Power data traceability · Blockchain · Hyperledger fabric · Data access control

## 1 Introduction

The 13th five year plan for power development (2016–2020) put forward in 2016 clearly stated that we should deepen the reform of power system, improve the power market system, establish relatively independent and standardized power trading institutions, establish fair and orderly power market rules, and initially form a fully functional power market. In the information age, power data can be regarded as an important enterprise asset of power enterprises, and the effective management of massive data is an essential task. In the process of data capitalization, the value of data assets can be developed and improved by creating data products, carrying out data services and data circulation, in which data traceability is an important link.

Since the 11th five year plan, China has gradually incorporated energy conservation and carbon reduction into the national economic and social development plan. In September 2020, China put forward the “dual carbon” goal, striving to achieve carbon peak by 2030 and carbon neutralization by 2060. The power sector is one of the key areas to achieve the goals of carbon peak and carbon neutralization. Green power trading can promote clean and low-carbon production of energy, ensure low-carbon operation of power systems, and help enterprises accelerate the completion of the overall goal of carbon peak and carbon neutralization, so as to achieve a high degree of unity of economic, social and environmental benefits, so as to form a new power system dominated by clean energy. It is an important means to realize the reform and efficient development of the power industry, an inevitable need to adapt to the development trend of technological progress in the energy field and promote the transformation and upgrading of the system, and an urgent need to promote the transformation of energy to clean and low-carbon. The traceability of green power product information can fully ensure that the user can purchase pure “green power”, and ensure the consistency of green power products and consumption from the perspective of business relationship and power generation and supply relationship, so as to avoid production waste and achieve real “zero carbon emission”.

In the current green power transaction process, there are many problems, such as difficult to identify the authenticity of power data, unreliable transaction records, imperfect supervision and so on. Therefore, it is necessary to put forward a reliable traceability solution for power data. The challenge of power data traceability makes the use of new technologies an inevitable trend. Blockchain integrates multiple core technologies, such as distribution, cryptography, consensus mechanism and smart contract, and has been widely used in all walks of life, including finance, food, logistics, etc. The characteristics of blockchain, such as transparency, multi-party consensus, tamper proof and traceability, have brought a new cracking idea to the reliable traceability of power data.

## 2 Related Work

The development of blockchain technology has brought a new development direction to the energy industry and can promote the intelligent development of power grid system. Reference [1] proposed the concept of double chain blockchain, which separates data blocks and transaction blocks, only adds power data asset summary to the consensus, and uses the interactive process of biological gene expression for reference to take the blockchain storing operation data as DNA blockchain, and the blockchain storing transaction information and data summary as RNA blockchain. Reference [2] designed a power data asset trading platform, in which a random authentication model for identity management was designed through Ethereum’s smart contract, POA consensus mechanism was adopted, periodic monitoring synchronization mechanism and data access tracking mechanism based on B+tree were proposed. Reference [3] studies the power data management architecture, analyzes the power data recording process, and ensures that the data is not tampered with by means of asymmetric encryption between block nodes and power data recording nodes; At the same time, a process of power data sharing is proposed by using asymmetric encryption technology; The security management

guarantee system of power data is proposed from two aspects of storage security and transmission security.

At present, there is still a lack of recognized concept of data traceability. The concepts vary due to different application fields. Here, refer to the definitions of some scholars. In reference [4], it is defined as the origin of data and its movement between databases in the field of database system. In reference [5], data traceability is defined as a kind of metadata, which is used to record the evolution process of workflow, annotation information, experimental process changes and other information; Reference [6] puts forward that all kinds of data have a life cycle, and data traceability is to record the changes and processed information of data in the whole life cycle from generation to extinction or transformation. The concept of W3C is that “the traceability of each resource is a record that describes the entity and the contained production and delivery processes or otherwise affects the resource.” At the same time, there are many forms of data traceability models. Reference [7] studies the OPM model, which has three core concepts: Artifact, Process and Agent. Artifact can be a physical object or a digital representation in a computer system. Process represents a single or a series of actions generated by Artifact. Agent is the main influencing factor of process, which can promote, control and affect the execution of Process. Reference [8] studies the PROV-DM model. Its core structure also includes three common aspects of Traceability Technology: Entity, Activity and Agent. Entity can be concrete objects such as buildings, or abstract concepts such as ideas; Activity is the action on entity; Agency, as an object, plays a certain role for the existence of Entity and Activity, which is essentially a special Entity or Activity.

At present, there are two solutions to the attribute based data access control mechanism, namely CP-ABE (ciphertext policy attribute based encryption) and KP-ABE (key policy attribute based encryption). In Reference [9], the CP-ABE mechanism is used and the original Fabric CA module in hyperledger fabric is combined to realize user level fine-grained secure access control of blockchain data.

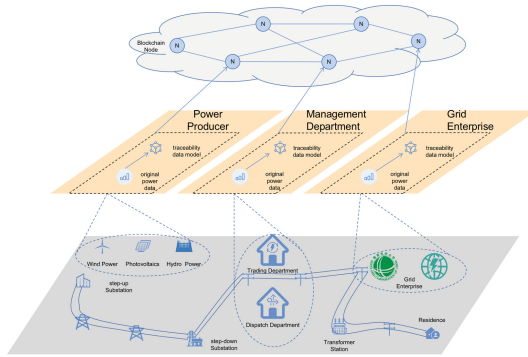
### 3 Power Data Traceability Based on Blockchain

In this paper, a three-tier architecture is established to realize the reliable traceability of power data based on blockchain. As shown in Fig. 1, the structure is composed of a data generation layer, a data conversion layer and a blockchain network layer. The functions of each layer are described as follows:

**Data generation layer:** it is composed of various sensors to obtain the power data generated in the whole life cycle of power data, including the whole process of power generation, transmission, transformation, distribution and power consumption.

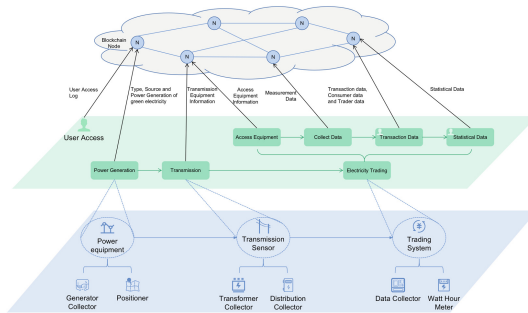
**Data conversion layer:** it consists of local nodes. The power data of the data generation layer is converted into data meeting the storage requirements through the traceability data model, and the original power data is stored locally.

**Blockchain network layer:** it consists of blockchain nodes. The identification and hash of the original power data in the data conversion layer are stored in the blockchain to realize the lightweight management of power data. The blockchain uses the raft consensus mechanism to ensure the credibility of the data on the chain. At the same time, it uses the smart contract to query the data records, and prevents the data on the chain from being tampered by comparing with the original power data stored locally.



**Fig. 1.** Power data traceability architecture based on blockchain

The data storage process is shown in Fig. 2. In the power generation stage, the type, source and power generation of green electricity are collected through the power generation terminal; In the transmission stage, the information of substation, distribution, data collector, watt hour meter and other equipment passing through the transmission process is collected through edge equipment; In the transaction stage, the transaction data, consumer data and trader data are recorded through the power grid system; In the three stages, there are user access links, which collect access records while users access data. Finally, all data is stored through blockchain nodes.



**Fig. 2.** Data storage process

### 3.1 Power Data Traceability Model

This paper constructs and describes the application scenarios based on the prov-dm model. The nodes and edges in the power data traceability model are defined according to three types and seven relationships in the PROV-DM model. Power generation companies trade electricity with power grid companies through market operation agencies, and power grid companies sell electricity to small and medium-sized end users through acquisition equipment and marketing systems. In this process, there are also operations such as data query, modification and statistics.

This paper designs the following three types of nodes:

- (1) Entity node: used to represent data. It is represented by  $E_n$ . E is the abbreviation of Entity;
- (2) Agent node: used to represent an individual or enterprise. It is represented by  $A_n$ . A is the brief form of Agent;
- (3) Activity node: used to represent operations on data. It is represented by  $AC_n$ . AC is the abbreviation of Activity.

This paper designs the following five types of connecting edges:

- (1) Used: used to connect the Active node and the Entity node used in the activity process. The Active node points to the Entity node;
- (2) WasGeneratedBy: used to connect the Active node and the new Entity node generated by the activity. The Entity node points to the Active node;
- (3) WasAttributedTo: used to connect the Agent node and the Entity node. The Entity node points to the Agent node;
- (4) WasAssociatedWith: used to connect the Agent node and the Active node. The Active node points to the Agent node;
- (5) WasDerivedFrom: used to connect two Entity nodes. The new Entity node points to the old Entity node.

According to the above description, this paper constructs the power data PROV-DM model as shown in the Fig. 3. The measurement data, marketing data and other data belong to Entity nodes, which are shown as oval nodes in the figure; Power trading, data collection and other operations are Active nodes, which are shown as rectangular nodes in the figure; Power generation enterprises, power grid enterprises and other enterprises are Agent nodes, which are shown as pentagon nodes in the figure. Connect the nodes according to the five edges defined above, and finally get a directed acyclic graph.

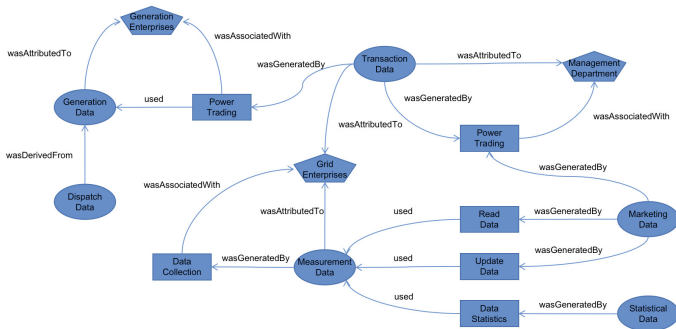
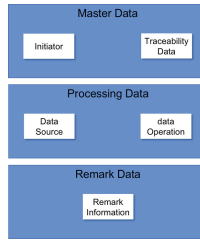


Fig. 3. Power data PROV-DM model

### 3.2 Power Data Model

#### Data Model

As shown in Fig. 4, the traceability data model includes master data, processing data and remark data.



**Fig. 4.** Data model

Master data includes two dimensions: initiator and traceability data. The data structure of initiator and traceability data is shown in Table 1 and Table 2.

**Table 1.** Data structure of initiator

Initiator Attribute	Explanation
elec_prov_actor:id	Unique ID of initiator
elec_prov_actor:name	Name of initiator
elec_prov_actor:desc	Related Description of initiator
elec_prov_actor:type	Type of initiator(person/company/other)

**Table 2.** Data structure of traceability data

Traceability Attribute	Explanation
elec_prov_data:id	Unique ID of traceability data
elec_prov_data:name	Name of traceability data
elec_prov_data:company	Company of traceability data
elec_prov_data:type	Type of traceability data
elec_prov_data:region	Region of traceability data

Processing data includes two dimensions: data source and data operation. The data structure of data source and data operation is shown in Table 3 and Table 4.

**Table 3.** Data structure of data source

Traceability Attribute	Explanation
elec_prov_source:dataId	Source data of traceability data

**Table 4.** Data structure of data operation

Traceability Attribute	Explanation
elec_prov_op:type	Operation on source data(read/update/process)

For the remark information in the remark data, it's able to add some user-defined information.

### Data Identification Scheme

This paper constructs a set of identification rules to make the traceability data of power supply chain unique in the world. This section mainly describes in detail the identification scheme of the Agent and Entity nodes mentioned above.

- (1) Agent: different Agents must be distinguished in the traceability system. Specific examples of Agent may include power generation enterprises, market operation institutions, power grid companies or individuals. When enterprises are involved, they can be identified by using Global Location Number (GLN); When an individual is involved, the Global Service Relationship Number (GSRN) can be used for identification. See Table 5 for details.

**Table 5.** GS1 standard

Code	Full Name	Explanation
GLN	Global Location Number	A code that uniquely identifies legal entities, functional entities and physical entities involved in supply chain activities. The global location code is composed of manufacturer identification code, location reference code and verification code, and is represented by 13 digits
GSRN	Global Service Relationship Number	It provides a globally unique identification code for service providers to access data related to services provided to service recipients. The global service relationship code consists of three parts: the manufacturer identification code, the service object code and the verification code, which is an 18 digit code

- (2) Entity: identify power data according to physical dimension, application dimension and time dimension, as shown in Fig. 5.

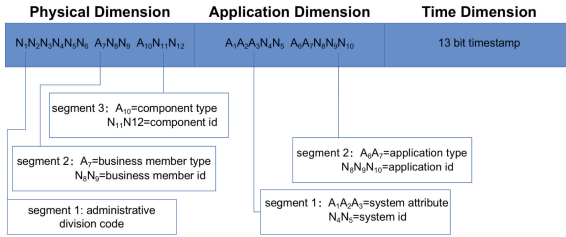


Fig. 5. Three dimensions scheme

**Data Access Control**

CP-ABE (Ciphertext Policy Attribute Based Encryption) means that the ciphertext corresponds to an access structure and the key corresponds to the attribute set. When decrypting, only the attribute values in the attribute set can meet the access structure in the ciphertext. All users obtain the key from the attribute organization through their own conditions and attributes, and the encryptor formulates the access control of the message.

According to the characteristics of Hyperledger Fabric with channel isolation, this paper designs three attributes: channel ID, organization ID and user ID. The specific access structure example is shown in Fig. 6. This access structure indicates that the specified users  $use^i$  and  $use^j$  in  $org^i$  of the specified organization in the specified channel  $channel^i$  have access rights. In this paper, the Fabric CA module is used to realize the initialization, key generation and distribution of CP-ABE, and the ciphertext is stored on the chain to achieve more secure data access control.

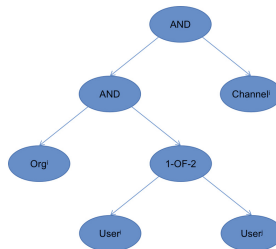


Fig. 6. Access structure



## 4 Experiments and Analysis

### 4.1 Storage Module

Storage modules are divided into local storage module and blockchain storage module, both of which are run in docker. The local storage module selects the Neo4j database, and Fig. 7 shows the Neo4j container. Hyperledger Fabric is selected for the blockchain storage module. Figure 8 shows the Hyperledger Fabric container.

CONTAINER ID	NAMES	PORTS
ce4b6ee7d4cd	Neo4j	0.0.0.0:7473-7474->7473-7474/tcp, 0.0.0.0:7687->7687/tcp

Fig. 7. Local storage module

CONTAINER ID	NAMES	PORTS
4721ca839962b	dev-peer0.org2.example.com-electric_1-02488c8f928e88a8a10783a9f67b3803aa16938c022381a10511779510833	
1922ca83226cc5	dev-peer0.org1.example.com-electric_1-92488c8f928e88a8a10783a9f67b3803aa16938c022381a10511779510833	
3e3b3099f227	dev-peer0.org1.example.com-fabcar_1-114604a491871bf18b23d6676d8cc08865b3acc0e2274f165e086079a8f276	
081f162212797	dev-peer0.org2.example.com-fabcar_1-114604a491871bf18b23d6676d8cc08865b3acc0e2274f165e086079a8f276	
02477c64c26	cli	
4ea97cac803c	peer0.org2.example.com	0.0.0.0:9051->9051/tcp, 7051/tcp, 0.0.0.0:9445->9445/tcp
509c26991316	peer0.org1.example.com	0.0.0.0:7052->7052/tcp, 0.0.0.0:9144->9144/tcp
38ce8f21901	orderer.example.com	0.0.0.0:7050->7050/tcp, 0.0.0.0:7053->7053/tcp, 0.0.0.0:9443->9443/tcp
13bca9f97b08	couchdb0	4269/tcp, 9180/tcp, 0.0.0.0:1098->1098/tcp
f092f6407b0b	couchdb1	4269/tcp, 9180/tcp, 0.0.0.0:7090->7090/tcp
933581e37439	ca_org1	0.0.0.0:7854->7854/tcp, 0.0.0.0:17854->17854/tcp
a7280c84e127	ca_orderer	0.0.0.0:9804->9804/tcp, 7804/tcp, 0.0.0.0:10804->10804/tcp
a41098ea3363	ca_org2	0.0.0.0:9804->9804/tcp, 7804/tcp, 0.0.0.0:10804->10804/tcp

Fig. 8. Blockchain storage module

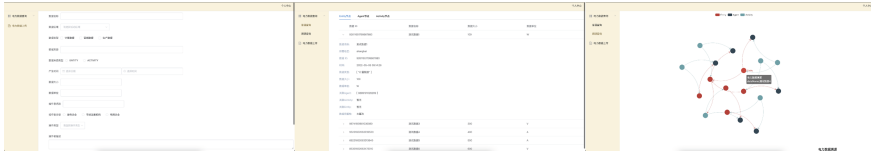
### 4.2 Data Operation Module

The project includes entity layer, mapper layer, service layer and controller layer. Entity layer corresponds to entities in the database; The mapper layer mainly does the work of the data persistence layer. Some tasks responsible for contacting the database are encapsulated here, specifically encapsulating the basic operations of adding, deleting, modifying and querying; The service layer is mainly responsible for the application logic design of the business module. First, the interface is designed, then its implementation class is designed, and one or more mappers are encapsulated again to form a service; The controller layer is responsible for the request forwarding, accepts the parameters passed from the front end, calls the interface of the service layer to control the business process, and gets the return value before passing it to the page.

### 4.3 Data Display Module

The data display module selects the Vue framework and divides the front end into login module, registration module and display module. The display module is divided into three sub modules: top navigation bar, side navigation bar and main module.

In view of the lack of power data sets, the power data upload function is designed, which is mainly used to simulate the generation of power data. Power data query is divided into ordinary query and traceability query. Ordinary query displays the data in the form of list, including detailed data attributes. The data integrity ensures that the data is not tampered by comparing the local data with the hash data on the chain; The traceability query uses the open source visualization library Echarts to display the data in the form of a graph, including only the identification of the data, as shown in Fig. 9.



**Fig. 9.** System function

According to the statistics, the time-consuming of the relevant functions of the system is shown in. Ordinary query needs to compare the data on the chain and the locally stored data to prevent data from being tampered with, so it takes a long time ; Traceability query can quickly display local data in the form of graph and takes a short time; Data uploading requires data to be stored locally and on the blockchain, so it takes a long time (Table 6).

**Table 6.** System function time consuming

Function	Time Consuming
Ordinary Query	2513 ms
Traceability Query	376 ms
Data Upload	10018 ms

## 5 Conclusion

This paper analyzes the power data traceability scenarios and entities involved in the power life cycle. With the help of PROV-DM model, the power data traceability model is discussed. A multi-dimensional data model is proposed and two data identification schemes are proposed for different types of data. The data access control method based on attribute encryption is discussed. The power data traceability system architecture is designed. The blockchain achieves high data consistency, tamper proof and traceability, and one of the costs is performance. Therefore, combined with the local data storage Neo4j, the lightweight management of data on the chain is realized. At last, the power data traceability system is realized.

**Acknowledgment.** This paper is supported by the science and technology project of State Grid Jibei Information & Telecommunication Company “Trusted Sharing Technology and Application of Winter Olympic Power Data Based on Blockchain” (52018E20008J).

## References

1. Lin, J., Cui, W.: Power-data-asset trading platform architecture based on double-chained Blockchain. *Electric Power* **54**(11), 164–170+180 (2021)
2. Guo, Y., Yu, L., Zhang, H., Yan, Y., Sun, Y.: Trading platform of power data asset based on consortium Blockchain. *Frontiers of Data Comput.* **3**(01), 48–59 (2021)
3. Long, Y., Zhong, Y., Gan, R.: Research on power data management architecture based on blockchain. *Power Systems and Big Data* **24**(08), 19–24 (2021)
4. Buneman, P., Khanna, S., Tan, W.: *Data Provenance: Some Basic Issues*. Springer, Heidelberg (2000)
5. Greenwood, M., Goble, C., Stevens, R.: Provenance of e-Science Experiments-Experience from Bioinformatics. 223–226 (2003)
6. Dai, C., Wang, T., Zhang, P.: Survey of data provenance technique. *Appl. Res. Computers* **27**(09), 3215–3221 (2010)
7. Shen, H., Zhang, X.: Data provenance model in semantic web environment: an overview. *New Technol. Library and Inf. Service* **04**, 1–8 (2011)
8. Zhao, Y., Dai, C., Ni, J., Kong, L.: Data provenance on semantic web based on PROV. *Computer Science Appl.* **5**, 141 (2015)
9. Qiu, Y., Zhang, Y., Cao, Q., Zhang, J., Chen, X., Jin, H.: Blockchain data access control scheme based on CP-ABE algorithm. *Chinese J. Network Inf. Security* **6**(03), 88–98 (2020)