A multi-source data collection and information fusion method for distribution network based on IOT protocol

Yongxiang Cai¹, Xiaobing Xiao¹, Hao Tian²,³, *, Yu Fu¹, Peng Wu¹, Hongliu He¹

¹ Electric Power Research Institute of Guizhou Power Grid Co., Ltd., Guiyang, China
² State Key Lab of Control and Simulation of Power Systems and Generation Equipments, (Dept. of Electrical Engineering, Tsinghua University), Beijing China
³ Binjiang College of Nanjing University of Information Science and Technology, Wuxi, China

Corresponding author e-mail: tianhao@mail.tsinghua.edu.cn

Abstract. With the construction of Global Energy Internet, Intelligent distribution network has more and more characteristics of multi scene, multi service and changeable measurement and control objects. In order to realize data collection and information exchange among the intelligent terminal and energy, power grid, load, energy storage and other control resources in distribution network, we should research an architecture which can adapt to various data sources Based on dynamic self describing protocol.

1. Introduction
With the development of large power grid control technology, information and communication technology and intelligent sensor technology and the construction of Global Energy Internet, power system has the trend of the scale expansion, the network complexity, the equipment precision and data massification [1]. There are many kinds of distribution network measurement and control terminals. Intelligent distribution network has the characteristics of multi scene, multi service and changeable measurement and control objects [2]. Data collection and information exchange based on interactive communication, which the intelligent terminal and energy, power grid, load, energy storage and other control resources realized can greatly improve the intelligent and universal level of distribution network terminals.

General data exchange methods include shared database, common data exchange file and common format data bus. At present, there is no effective method to collect and fuse all kinds of application system and intelligent terminal data in distribution network. This paper provides a multi-source data collection and information fusion method based on IOT protocol, which based on self describing IOT, use SOA (Service-Oriented Architecture), and realize cloud integrated distribution network information model in IOT.

2. Communication Protocols for the Internet of Things (IOT)
Intelligent terminal networking of energy, power grid, load and energy storage end of distribution network forms a complex electricity IOT. The first step of information fusion is data transmission and communication. In the IOT, communication protocol is the most important to ensure data transmission and communication.
IOT protocols are generally divided into two categories, transport protocol and communication protocol. Transport protocol is responsible for the networking and communication between devices in the subnet. Communication protocol mainly focuses on the equipment communication protocol, running on the traditional Internet TCP/IP protocol. It is responsible for data exchange and communication between devices through the Internet.

2.1. Common IOT protocols

2.1.1. REST/HTTP (Service Invocation of Loose Coupling). Representational State Transfer (REST) is a communication style based on HTTP protocol. REST refers to a collection of architecture constraints and principles. The application or design which meets those constraints and principles is RESTful[3]. REST/HTTP is designed to simplify the system architecture in the Internet, achieve loose coupling between client and server rapidly, and reduce interaction delay between client and server. Actually, REST/HTTP is an package style of service invocation API in the Internet. In the application system of the IOT, after the data is collected, the data service can be opened by opening the REST API, and then, it can be called by other applications in the Internet.

2.1.2. CoAP Protocol (Constrained Application Protocol). CoAP is a limited application protocol, which is applied to protocols in wireless sensor networks. CoAP is a RESTful API that simplifies the HTTP protocol. It is an application-layer protocol in the 6LowPAN protocol stack which is suitable for IP networks with limited communication resources. The goal of CoAP is to solve the problem that the device connects directly to the IP network, in other words, the communication requirements that IP technology is applied between devices or internet and devices. IPv6 technology brings huge addressable space, which not only solves the identification problem of huge amount of equipment and resources in the future, but also enables Internet applications to directly access devices supporting IPv6 without requiring additional gateways. Compared to HTTP, CoAP has the following features:

   (1) Message header compression  
   (2) The transport layer uses UDP protocol  
   (3) Supporting asynchronous communication  
   (4) Supporting resource discovery  
   (5) Supporting caching  

2.1.3. MQTT Protocol (Message Queuing Telemetry Transport). MQTT is an instant messaging protocol developed by IBM, which uses the publish/subscribe mode. All the IOT terminals connect to the cloud through TCP, and then the cloud manages the communication content concerned by device through a thematic way, and is also responsible for forwarding messages between devices[4]. The MQTT protocol provides data transmission and monitoring of remote equipment based on cloud platforms in low-bandwidth, unreliable networks. The MQTT protocol is generally suitable for device data collection to-end (Device->Server, Device->Gateway). Hub-and-spoke network architecture is not suitable for communication between devices, and the control capability of devices is weak. In addition, its real-time performance is poor, generally in second level.

2.1.4. DDS Protocol (Data Distribution Service for Real-Time Systems). DDS is a data distribution service for real-time systems. It is suitable for data communication of distributed, high-reliability, real-time transmission devices. It can well support data distribution between devices, device control, data transmission between device and cloud. At the same time, DDS has very high real-time efficiency of data distribution, which can distribute millions of messages to many devices at the same time in seconds. Since DDS provides many ways to guarantee the quality of service, it is widely used in the fields of high reliability and high security, such as national defense, civil aviation and industrial control. However, its
limitation is that it can only work under the wired network. In other words, there are some difficulties in application under the wireless network, especially in the case of resource constraints.

2.1.5. XMPP Protocol (Extensible Messaging and Presence Protocol). XMPP Protocol based on standard general markup language, XML, which inherits the flexibility of development in the XML environment and therefore has strong scalability, openness and ease of use. Compared with HTTP, XMPP is more suitable for IOT system in communication business process. Developers don't have to pay too much attention to solve the business communication process of device communication, so the development cost is relatively lower.

2.1.6. JMS (Java Message Service). Java Messaging Service is a Java platform API for Message Oriented Middleware (MOM), which is used to send messages between two applications or in a distributed system for asynchronous communication. Java Messaging Service is a platform-independent API, supported by most MOM providers. JMS can send messages from one JMS client to another through a messaging service. Messages are a type of object in JMS that consists of two parts: header and message body. The header consists of routing information and metadata about the message. The message body carries the data or payload of the application.

2.2. Actual Scenario Protocol Combination
When designing the specific system architecture of IOT, the communication requirements of the actual scene should be considered and the appropriate protocol should be selected. The following content takes smart home as an example to illustrate the application direction of these protocols. Smart light control in smart home, DDS protocol can be used to monitor the engine group of power plant. When the power is transferred to millions of households, the inspection and maintenance of power lines can use MQTT protocol. The power consumption of all household appliances can be transmitted to the cloud or home gateway for analysis using AMQP protocol. Finally, users can use REST/HTTP to open API services if they want to publish their energy query services to the Internet.

DDS, MQTT, AMQP and JMS are all based on the publish/subscribe mode. The publish/subscribe architecture has the features of service self-discovery, dynamic expansion and event filtering. It solves the problems of fast data source acquisition, object joining and exiting, interest subscription, reducing bandwidth traffic and so on in the application layer of the IOT system. More than this, it realizes that the connection of objects could be loosely coupled in space (both sides do not need to know the communication address), time and synchronization.

3. General Data Exchange Protocol for Traditional Power System
In order to facilitate the operation of monitoring data migration, transformation and data access generality in distribution network, some applications use the form of general exchange file to access and transmit the operation monitoring data of distribution network. XML format is a common data exchange format in the world, which has the advantages of strong compatibility, flexibility and extensibility, etc. Based on the XML standard and according to its expansibility characteristics, the State Grid Company puts forward the power system data markup language, CIM/E specification, according to the characteristics of power system. It is a new efficient data markup language for power system, which based on the object-oriented abstraction of IEC61970-301 power system common data model CIM (Common Information Model), and repaired the efficiency defects in XML [5]. CIM/E specification is specially used for on-line exchange of data such as power grid model, between application systems of dispatching control center. CIM/E standard format is adopted for operation monitoring data of distribution network as follows:

```xml
<relaysig>
  @id name vl_id  pnt_type  pri_flags  ...
  #122160141620511071 Two-stage pressure plate 0001 1 3
</relaysig>
```
The <relaysig> </relaysig> tag indicates the beginning and end of the data block, with the content of the data block in the middle.

@ represents the header of the behavior data and defines a series of titles separated by spaces or tabs. 
# represents the behavior data row, and the data content and separator are corresponding to the header.

Compared to XML, CIM/E adds content-related symbols, which greatly simplifies the definition of data row and improves the efficiency of data access and transmission.

4. Data fusion based on IOT dynamic CIM/E protocol of SOA architecture

4.1. Dynamic CIM/E Data Model Based on IOT Protocol

With the continuous expansion of distribution network scale, the equipment is becoming more and more complicated, refined and intelligent. The traditional static CIM/E data model can no longer meet the data fusion compatibility and customization requirements of heterogeneous system and heterogeneous equipment.

The IOT Protocol is generally based on publish/subscribe mode. It has the characteristics of service self-discovery, dynamic expansion and event filtering, which can effectively solve the rapid access to data sources, real-time join and exit of heterogeneous systems and heterogeneous devices in the IOT of distribution equipment.

Dynamic CIM/E based on IOT protocol, is to embed CIM/E format data in the form of content or body into IOT protocols such as DDS, MQTT, JMS, REST/HTTP, etc. It uses the characteristics of self-discovery and dynamic expansion of IOT protocol to effectively transmit and fuse the monitoring data of distribution network operation. In the case of REST/HTTP, the protocol includes Header and Body. The header stores the attributes, restrictions and user-defined information of the transmitted data in the form of key-value.

Theoretically, there is no limit to the number of user-defined attributes, so the scalability of the protocol is effectively guaranteed. The body stores the actual data content in an unlimited format. We can encapsulate the operation monitoring data of distribution network into CIM/E, and insert it into the body. Moreover, we can insert multiple data into an array. In theory, the size of the data volume in the body is also unlimited. The structure of the dynamic CIM/E data model based on IOT is shown in Table 1.

Table 1. The structure of the dynamic CIM/E data model based on IOT

| Header | Server address |
|--------|----------------|
| Host   | Connection: Keep-Alive |
| Accept-Charset | Encoding formats that can be sent |
| Accept-Language | Languages that can be sent |
| Type | Message type |
| (Customization) | Custom properties |

| Body |
|---|
| CIM/E | Data 1 |
| CIM/E | Data 2 |
| ... | ... |

4.2. Data fusion based on SOA architecture

SOA (Service-Oriented Architecture), is a distributed service architecture. It links the different functional units (called services) of an application through defined unified interfaces and contracts between these services. The interface is defined in a neutral way, which is independent of the hardware platform, operating system and programming language to implement the service [6]. This enables the services built in such a system to interact in a unified and common way.
The key of distribution network data fusion based on SOA is to realize a general data bus, and provide the standard of data access interface compatible with each acquisition system. Each data acquisition system registers services with the bus registry service center and becomes the data provider, providing data to the bus in the form of an adapter. The data fusion system consumes data from the bus, analyzes all kinds of data according to the protocol, and then uses ETL engine to realize data format summary and store on demand. At the same time, the unified data access service of IOT protocol is provided in a dynamic CIM/E mode. The specific architecture is shown in Figure 1 below.

**Figure 1.** Architecture of data fusion system for distribution network based on SOA.

The unified data access service provides an access interface based on REST/HTTP IOT protocol. The detailed flow chart of data user accessing unified data access service is shown in Figure 2.

**Figure 2.** Data acquisition process of dynamic CIM/E model based on IOT protocol.
Specifically:
(1) Data users send HTTP requests to the unified data access service, requesting a CIM/E model of required data and other relevant conditions.
(2) After receiving the request, the unified data access service calls the data bus unified data access interface to obtain the required distribution network data.
(3) Data bus returns distribution network data to ETL engine.
(4) The ETL engine processes the data, including analysis, extraction, conversion and storage when necessary, and returns the processed distribution network data to Unified Data Access Service.
(5) The Unified Data Access Service assembles the received data according to the CIM/E model in the request, encapsulates it in the IOT protocol, and returns it to the data user in the form of HTTP response.

After the above process, a data user can access the distribution network data in any collection system registered in the bus through The Unified Data Access Service and data ETL engine, which is, using SOA architecture to realize the data fusion of heterogeneous systems in distribution network.

5. Summary
In this paper, IOT protocol is used to encapsulate dynamic CIM/E distribution network data, and SOA architecture is used to solve the data fusion problem of complex heterogeneous devices and systems in the distribution IOT. Firstly, the data heterogeneity and complex IOT of energy, power grid, load and energy storage end of distribution network lead to the difficulty of data fusion. Then it introduces the common communication protocol and IOT protocol which encapsulates CIM/E distribution network data to realize dynamic CIM/E. On this basis, it introduces SOA architecture and designs a data fusion solution for heterogeneous equipment and system of distribution network based on this architecture. The implementation method of multi-source data collection and information fusion of distribution network based on IOT protocol has universal reference value for heterogeneous data fusion of transmission network, distribution network and even the whole energy Internet.

Acknowledgments
This work was financially supported by “Research and Demonstration on Key Technologies of Green and High Reliable Smart Distribution Networks in Huangguoshu Scenic Area (No. 060400KK52190017)” fund.

References
[1] Information on https://baike.baidu.com/item/SOA/2140650?fr=aladdin
[2] Qi Linhai, Ai Minghao. A Voltage Sag Parallel Calculation Method Based on Cloud Computing, Proceedings of the CSEE, 2014, 31:5493-5499.
[3] Ai Minghao, Ge Xianjun, etc. A Big Data analysis based new method for power grid dispatch and control training simulation, CICED 2016, Xi’an.
[4] Mostafa D. Fard, Hessam S. Sarjoughian. A RESTful framework design for componentizing the water evaluation and planning (WEAP) system, Simulation Modelling Practice and Theory, 2021, 106.
[5] Yao Lili. Research on data breakpoint continuation scheme based on MQTT protocol, Internet of things technology, 2020, 10(10): 30-32.
[6] State Grid Corporation of China; "System, Method And Storage Device For Cim/E Model Standard Compliance Test" in Patent Application Approval Process (USPTO 20180307594), Energy Weekly News, 2018.
[7] Information on https://baike.baidu.com/item/SOA/2140650?fr=aladdin