LEO IoT based big data management and analysis platform design for intermodal containers

Jieyin LYU1,2,* , Shouqin Zhou2, Hao Liu2 and Xiangmo Zhao1

1Chang’an University, School of Information Engineering, Xi’an, P.R.C.
2Shenzhen CIMC Intelligent Technology Co.,Ltd., Shenzhen, P.R.C.

*Email: jieyinlv@chd.edu.cn

Abstract. In view of the global information management requirements of intermodal containers, combined with the development status, this paper proposes an IoT model based on LEO satellites, and a big data management and analysis platform based on the model. This paper designs and analyses the scheme of the platform, and proposes a middleware approach to solve the problem of multi-source heterogeneous data access of containers intelligent terminals. Then, this paper designs business processes such as user and device registration, middleware registration, historical data access, and timely data access for the platform. The big data analysis and management platform has been prototyped and verified; the platform has the advantages of compatibility, openness and convenience. This study has certain value.

1. Introduction

Intermodal transport is usually a continuous and comprehensive form of cargo transport organization, using containers as the transport unit and organically combines different modes of transport. It is suitable for a variety of transportation modes such as waterways, roads, railways and aviation. In international trade, around 90% of goods are completed through intermodal transport. As the main carrier of intermodal transport, the location, status, security and other data in the container transportation process are often the content of the intermodal transport participants, usually using the means of the Internet of Things for collection, transmission and management, so called intelligent container [1-3].

At present, there are at least three problems in the intermodal transport container Internet of Things. This paper designs a big data[4] management and analysis platform based LEO satellites IoT [5] for intermodal containers to solve the below problems.

1) Platform data access method: Container transportation has the characteristics of global circulation and operation. It requires the transmission of IoT communication data with global coverage, low power consumption and low cost. However, the usual RFID and cellular communication cannot meet the above requirements, and the combination of LEO satellites and LPWA can fill the shortcomings of the technology. In recent years, the global LEO satellite industry has developed rapidly, such as Orbcomm Satellite constellation, Tianqi Satellite constellation, etc. At the same time, NB-IoT, LORA and other LPWA technologies have also appeared.

2) Platform data access protocol: The custom communication protocol used for data transmission between the container intelligent terminal and the container data operation management
platform. Due to the lack of a unified communication protocol and data format, the generated data is multi-sourced. The data of the structure makes it impossible for the container operation and management platform to collect the data generated by the intelligent terminals produced by various manufacturers quickly and conveniently. How to solve the problem of global container multi-source heterogeneous data access is a problem in the current container management platform.

3) Data interaction between platforms: In the entire business chain of container intermodal transport, including cargo owners, freight forwarding companies, car companies, carriers, ports, customs, factories, highways and management departments, etc., participate in division of labor and cooperation. However, there are many platforms in the industry chain at present, there is no unified communication protocol between the platforms, and the data formats are different, leading to information barriers, poor communication, and etc.

2. LEO satellites based IoT
The LEO satellites based IoT model is shown in figure 1. It consists of intermodal container intelligent terminal [6], transmission channel, big data management and analysis platform and third-party service platform.

In this model, the intelligent terminal is manufactured by various terminal manufacturers. Data transmission channels can be mobile cellular networks or LEO satellites. The big data management and analysis platform can be used to access, manage and publish data. The third party service platform can obtain data from the big data management and analysis platform, and carry out corresponding business interaction and trade. Considering the commercial requirements of low power, low cost and full coverage, Orbcomm Satellite Constellation is adopted for LEO satellite data transmission channel and NB-IoT for LPWA low power WAN for cellular mobile network data transmission channel.

3. The platform design
The big data platform is a public service platform and a data storage center. It does not care about the specific data content transmitted. Therefore, the commercial value of data transmitted and the specific data content are not visible to the big data management and analysis platform. The big data platform gives unified data and interface standards. Terminal manufacturers develop their own middleware to encrypt, decrypt and access big data management and analysis platform.

The big data management and analysis platform is an open data platform that includes four main modules: data center, data collection, data release, data distribution, and operation management. The design of the big data platform needs to include these four modules as well as middleware communication protocols and typical business processes. The scheme of the platform is shown in figure 2.
The main technical difficulties of the big data platform are twofold. On the one hand, it is the demand for the expansion of heterogeneous services, that is, facing different link providers need to establish corresponding data transmission protocols, facing different devices need to accommodate the corresponding middleware. On the other hand, it is multi-user demand for scalability in data distribution.

Therefore, the big data platform adopts the architecture of the cloud platform. On the one hand, a vertical extension mechanism is adopted in the access module for data collection to support different application layer transmission protocols of multiple link providers in the future. On the other hand, in the data distribution module, the middleware layer adopts a horizontal expansion mechanism to improve the system's concurrency and load capacity, and to support unlimited expansion of middleware.

### 3.1. Data center module design

Data center module includes data access layer and big data center. Data access layer mainly defines the interface of writing and reading information to big data center, which can be used by data collection module, data publishing module and management platform, such as connecting database interface, closing database interface, inserting intelligent terminal information interface, inserting user information interface, querying user information interface, querying user equipment interface, etc. The scheme of the data center module is shown in figure 3.

Data stored and analyzed in big data centers come from data access layers, including data sent by intelligent terminals, user data, user-device binding relationship data and so on. Big data center
database uses distributed storage architecture. Redis is used for real-time data, and MongoDB is used for business and other data with low real-time requirements. The two databases are selected according to different business needs. At the same time, Hadoop extracts the data from both of them regularly for big data analysis.

Real-time data of big data platform, which accesses through distributed message middleware, is real-time processed, saved to data warehouse, and pushed to management platform or third-party platform for display.

Through regular batch processing, data are extracted from distributed database for statistical analysis, mining and modeling processing (for real-time processing, real-time prediction based on the model established by batch processing), and the statistical results are stored in the database for platform query display.

3.2. Data collection module design
Because the data generated by different intermodal containers and their intelligent terminals are multi-source and heterogeneous, so the data collection module adopts the open design idea, in order to collect the data stream transmitted by different transmission protocols between different link providers, and convert it into a unified data format and store it in the data center module.

Data collection module mainly includes four levels: terminal authentication and identification, content forwarding layer, middleware layer and data access layer. The specific design of each layer is as follows.

- Terminal authentication and identification: When the data of the intelligent terminal is forwarded from the terminal access layer to the authentication layer, the terminal ID is verified, that is to say, its validity is verified; invalid ID information is eliminated directly, and a valid ID terminal further identifies the middleware adapted by the terminal manufacturer.
- Content forwarding layer: After terminal authentication and identification, the access load layer will deal with data, then it is sent to the relevant middleware for parsing and standardization, and responds to the communication link public server.
- Middleware layer: After receiving the data package from the upper layer, the middleware encapsulates the data in the specific format required by the data through the analysis function module defined by the device manufacturer and the data re-encapsulation module, and transmits the data to the next layer through TCP/IP protocol.
- Data access layer: After receiving the data from the middleware, the data access layer calls the related interface functions of the data access layer and stores the data in the big data center.

3.3. Data release module design
Data distribution module is designed and developed for the third party. It consists of different business modules. Each business module provides a logical way to realize business. Business module receives requests from access layer, calls corresponding business logic methods, updates database content through database access layer interface, and generates business response required by users. In big data platform, an application server host only carries one business module, and each module runs independently. When new business needs arise, only the corresponding business module and data access layer interface need to be added, in order to have high scalability.

This paper provides four business modules, including user authentication, user equipment management, device information query, data distribution. The data in the big data center is sent to the third party by HTTP protocol through PUSH or PULL.

- User authentication: The big data center accepts and processes HTTP requests distributed by the access layer, extracts and encapsulates relevant parameters, and verifies the user's identity. When the user's authentication is successful, the system will allocate a unique Session Id, and store the Session Id in the database user table to generate a successful login response package. After that, every user's request needs to carry Session Id, and the system will authenticate user
requests according to Session Id; when user authentication fails, the login failure package will be generated directly.

- User equipment management: Users need to manage their own devices, such as add, change, delete, they can login to a web page we provide third-party developers for device management operations.
- Device information query: After identification, the user establishes a connection with the platform. At this time, if the user needs to query the device owned by him, he sends a device information request to the background through HTTP. After business analysis of data access layer, the request is sent to the device information query service cluster, which retrieves the big data center according to the content of the request, and then returns the data to the third party platform.
- Data distribution: While the data access layer writes the data from the intelligent terminals into the big data center. According to the specific business, it can transmit a signal to the data distribution module. The data distribution module directly pushes the intelligent terminals data to the third party platform.

3.4. Operation management module design

The operation management module is mainly responsible for the functions at the operation level of the big data platform. On the one hand, it provides system management such as state control between the system and the server hosts; on the other hand, it provides business-level management functions such as middleware deployment and management, traffic statistics, user management and registration.

4. Data access middleware design

Middleware can shield complex operating systems and network protocols in heterogeneous systems, and provide standard program interfaces and protocols. In order to be compatible with different data protocols of terminal manufacturers, big data platforms need to customize uniform data writing specifications. The terminal manufacturer needs to develop the corresponding middleware to access to the big data platform. The data protocol of the terminal manufacturer needs to be converted into the standard data writing protocol of the open platform through the middleware. Data reading of big data platform requires terminal manufacturers to develop corresponding reading middleware. Reading middleware converts the standard data format of big data platform into the data format customized by terminal manufacturers for customers to use. The middleware controller assigns an independent port to each terminal access parsing middleware, and packages the middleware into dynamic link library (DLL) files, Jar or independent programs for platform programs to read and parse. If it is a stand-alone program, the data can be transmitted to the platform through TCP/UDP, HTTP, MQTT, CoAP and other communication protocols after the data analysis is completed. Middleware data flow for dynamic link library (DLL) and independent programs are shown in figure 4.

![Middleware data flow](image)

**Figure 4.** Middleware data flow.
By formulating the standard of intermodal container terminals accessing data middleware, the intelligent terminal can access big data management and analysis platform according to the unified standard protocol, realize the fast communication and data sharing between the middleware and the platform, which is conducive to the unified centralized management of intermodal container terminal data. The big data platform provides unified interface to the outside world through API, and carry out timely and effective data analysis, and it can improve the efficiency and security of intermodal transport logistics.

5. Typical business processes design

The main business of big data management and analysis platform includes user and terminal device registration, middleware registration, historical data access and instant data access. Because the big data platform designed in this paper is a standardized public open platform for the whole intermodal container industry, it is necessary to standardize the design of its key business processes.

5.1. User and terminal device registration process

User and device registration business is carried out between operation management platform and user interaction platform. User interface shown in figure 2 is their information exchange interface. User interaction platform carries out new user registration, user information filing, platform device number application, platform device number and hardware binding application, etc. Operations management platform audits user information filing and platform equipment number filing applications. User and terminal device registration process is shown in figure 5.

![Figure 5. User and terminal device registration process.](image)

5.2. Middleware registration process

The Middleware is developed by terminal equipment manufacturer according to the standard of big data management and analysis platform. After the development of middleware, middleware...
registration is needed in big data management and analysis platform. Then, middleware can be used in big data platform. The registration of middleware is also accomplished by information exchange between operation management platform and user interface platform through user interface. The middleware registration can only be carried out after the its terminal manufacturer registers. Middleware registration process is shown in figure 6.

5.3. Data access process
Data access includes real-time data access and historical data access, it refers to data access of big data management and analysis platform by operation management platform or third party platform. This part mainly shows the call and analysis process of middleware. It is shown in the figure 7.

6. Platform prototype development and verification

Figure 7. Historical data access process (Calling middleware).

Figure 8. Prototype and verification.
According to the above design, this paper develops the prototype of big data management and analysis platform. The platform prototype application is carried out in combination with “The Belt and Road Initiative” container intermodal transport. The platform prototype and its validation applications are shown in figure 8.

7. Conclusion

As the main means of intermodal transport, Containers connect intermodal transport in series all over the world. With the development of globalization, international intermodal freight logistics trade is more frequent, and the iterative progress of science and technology is gradually meeting the increasingly strong demand of human beings for information acquisition such as location, status and safety of containers worldwide. However, there are still many areas to be studied in the unified data access, processing and distribution platform for data operation management and analysis.

The big data management and analysis platform based on LEO satellite Internet of Things proposed in this paper has the characteristics of openness, simplicity and low cost. No matter who is the terminal manufacturer and what technology the terminal adopts, as long as the terminal manufacturer provides the middleware of the middleware protocol standard, it can access the big data management and analysis platform proposed in this paper easily. As a public service platform for data acquisition, storage and distribution, the platform serves the customers of the whole supply chain of intermodal container transportation. The solution proposed in this paper adopts the integration of LEO satellite and LPWA, which has the characteristics of low cost and low power consumption.

This study provides a better design idea for the global operation and management of container intermodal transport and the cost reduction and efficiency enhancement of the global freight logistics supply chain.

Acknowledgments

Supported by the National Key Research and Development Program of China (2017YFB0503702, 2018YFB0505201 and 2017YFC0804809), Shenzhen Municipal Technology Program(JSGG20170414114406355), Introducing Innovative Scientific Research Team Plan of Dongguan City (201636000200039), and Shenzhen Satellites Based Equipment IoT Technology Engineering Laboratory ([2016]724). Thanks to my research team.

References

[1] Lang W , Jedermann R , Mrugala D , et al. The “Intelligent Container”—A Cognitive Sensor Network for Transport Management[J]. IEEE Sensors Journal, 2012, 11(3):688-698.
[2] Dittmer P , Veigt M , Scholz-Reiter B , et al. The Intelligent Container as a Part of the Internet of Things[C]. IEEE CYBER 2012. IEEE, 2012.
[3] Hidalgo Fort E , Garcia Oya J R , Munoz Chavero F , et al. Intelligent Containers Based on a Low-Power Sensor Network and a Non-Invasive Acquisition System for Management and Tracking of Goods[J]. IEEE Transactions on Intelligent Transportation Systems, 2017:1-6.
[4] Xuejiao Chen, Xiaofeng Hu, Yang Liu. Summary of Research on Big Data [J]. Journal of System Simulation, 2013(S1):142-146.
[5] Qu Z , Zhang G , Xie J . LEO Satellite Constellation for Internet of Things[J]. IEEE Access, 2017:1-1.
[6] Colombaroni C , Fusco G , Isaenko N , et al. Optimization of container operations at inland intermodal terminals[C].IEEE International Conference on Models & Technologies for Intelligent Transportation Systems. IEEE, 2017.