

# Research on power communication data security based on ETL

Xiaoyan Liu\*

Shandong Xiehe University, Yaoqiang town, Licheng district, Jinan city, China

**Abstract.** The use of modern data application technology in information security management programs within power enterprises is a major trend in industry development. Integrate scattered, messy, and non-standard data in electronic communication, extract, clean, and transform business system data, and load it into a data warehouse. ETL mainly relies on cloud computing logic to protect the transfer, storage, confidentiality, and other aspects of data, effectively avoiding the loss, omission, and leakage of information in enterprise data transmission. This study discusses the necessity of ETL based data transmission security protection technology for power enterprises to achieve security, in order to provide theoretical support for subsequent research and practice.

## 1 Introduction

The internal information database of power enterprises contains multidimensional information data, among which power grid program data and user privacy demand data are the most critical. The security of such data information will be directly related to the orderly business development and long-term energy supply of power enterprises. In recent years, due to the complex international situation, the cloud platform of power enterprises, which is related to China's energy security and information security, has repeatedly experienced information and data leakage, seriously hindering the modernization of China's power enterprises' industrial transformation

Progress<sup>[1]</sup>. At the same time, breakthroughs in cloud computing at the technical level have also given energy enterprises a new thinking path for data transmission security protection. If we strengthen the modern application of such technologies, it will inevitably open up a bright development path for power enterprises<sup>[2]</sup>.

## 2 Current situation of data security technology in electric power enterprises

China's power enterprises and their affiliated related enterprises synchronously utilize a unified cloud platform system to achieve secure and efficient transmission of confidential information. At the same time, it can also strengthen the security isolation between

---

\* Corresponding author: [1054878624@qq.com](mailto:1054878624@qq.com)

differentiated security hierarchical networks within the current system, preventing the leakage of confidential information data within the enterprise in collaborative transmission from a physical perspective<sup>[3-4]</sup>.

### **3 Transformation and operation of communication data security with ETL**

As a data source and data consumption layer, the ETL has complex data sources, so the ETL work must be defined using a simple design. There are many ETL tools, this paper introduces Kettle, Kettle in the open source software will ETL abstract into conversion and job two actions<sup>[5]</sup>.

#### **3.1 The conversion of ETL**

The object of conversion processing is the power communication data stream, which consists of three steps: input, processing, and output. There are two types of transformations to convert the extracted data from its previous form to the desired form. One is the unification of data names and formats, such as naming, data format, calculation rules, etc; The second type of data that does not exist in the data source requires the combination, segmentation, and calculation of fields, such as null value processing, data standards (unified metadata, standard fields, field type definitions), data splitting, data validation (such as time rules, business rules, self defined rules), data replacement (such as invalid data, missing data), and data association.

#### **3.2 Job**

Job management mainly includes functions such as general, email, file management, conditions, scripts, and batch loading.

### **4 Factors affecting the data security management of electric power enterprises**

#### **4.1 Network system vulnerability**

In the process of network communication, no matter what information is transmitted, it cannot be separated from the support of network software and systems. Once the support system has vulnerabilities, it will become an important hidden danger to communication network security. Although with the continuous development of relevant software and system iteration, as well as the continuous development of related technologies, China's communication network security has been basically guaranteed, based on the current status of communication network maintenance, system vulnerabilities are still an important factor affecting communication security. The necessity of finding and repairing vulnerabilities is self-evident.

#### **4.2 User access requirements**

From the perspective of network usage functions, its essence will prioritize the requirements of user capture, market demand, and information exchange in the construction and use of network platforms.

### **4.3 Geographical location and regional factors influence**

The establishment and use of communication networks cannot be separated from the technical support of terminal databases and processing centers. Whether it is the construction of communication network base stations or the construction of large-scale information processing systems, it is necessary to meet the requirements of long-distance network transmission.

## **5 ETL based data security for power communication**

### **5.1 Application of updates to firewalls and encryption systems**

The application rate of firewalls and encryption systems is very high in the current field of network defense. As a powerful security barrier between internal and external networks, as well as private and public networks, firewalls can filter out multiple illegal access requirements and control the access permissions of users with poor security through isolation operations and access control functions. This undoubtedly installs a safe door for communication security. In the context of big data, communication technology personnel should utilize the information collection capabilities of big data to enhance their technical limits, conduct research and updates on previous firewalls and encryption systems, establish a sense of technological leadership, and apply systems with higher security and usage performance to network communication.

### **5.2 Establishing mathematical models to analyze data**

Use ETL tools to establish mathematical models and establish mechanisms for abnormal data collection and analysis. One major advantage of big data technology is the universality and rapidity of information collection. When applied to communication network security management, this advantage can be combined with communication technology to establish an abnormal communication data collection and analysis system with excellent development prospects. This system mainly includes two modules: data collection and integrated analysis, among which the data collection module is a key link in the operation of the entire system. From the perspective of data universality, data collection work needs to cover user data, transmission data, traffic monitoring, vulnerability updates, event logs, etc. From the perspective of data capture process, reliable and stable information capture devices are the foundation of data collection. From the perspective of the data analysis module, the application of big data technology is very extensive. The collected data needs to go through the screening and mining of big data technology, and abnormal results need to rely on big data for visualization and data backup, such as generating adversarial networks.

### **5.3 Adopting SaaS cloud security products**

SaaS (Software as a Service) is a cloud computing model that primarily provides owners with a cloud software service, rather than installing the software on a local computer. On the one hand, the advantages of big data technology in communication security management can be reflected through SaaS cloud security products. On the other hand, SaaS cloud security products can utilize big data technology to provide more accurate services and better user experience, use the data warehouse services provided by SaaS to store and manage big data, and carry out specialized security management for communication sensitive information,

such as data encryption, multi-layer key applications Set sensitive information transfer tokens, etc.

## 6 Conclusion

In the context of informatization, the power industry faces opportunities and challenges in business data processing and security due to the existence of various business systems. Dialectically, challenges and opportunities are inevitable development trends, and one thing is to gain experience and achieve significant development through challenges. Therefore, we need to explore the factors that affect data security, analyze the current situation, find security management strategies, and study and discuss how technological innovation can add powerful methods to security.

## References

1. Yu Y, Zong N, Wen A ,et al.Developing an ETL tool for converting the PCORnet CDM into the OMOP CDM to facilitate the COVID-19 data integration[J].Journal of Biomedical Informatics, 2022, 127:104002-.DOI:10.1016/j.jbi.2022.104002.
2. Cai L, Su J .Research on data center construction based on Extract-Transform-Load (ETL)[J]. 2017.
3. Zhang, Qinhe,Li,et al.Design and Implementation of ETL Strategy in Tele-Communication Based on the Tool of Teradata Data Warehouse[C]//0[2023-09-17].
4. Savitri F N, Laksmiwati H .Study of localized data cleansing process for ETL performance improvement in independent datamart[C]//International Conference on Electrical Engineering & Informatics.IEEE, 2011.DOI:10.1109/ICEEI.2011.6021806.
5. Fei W U, Gui-Fen X , Yu-Ping X .Research on design of ontology-based ETL[J].Computer Engineering and Design, 2007.