Control on Abnormal Data Overflow of Distribution Network Management Platform

Li Bo, Zhao Ruifeng, Lu Jiangang, Guo Wenxin, Liu Yang

Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangdong, Guangzhou 510600, China
38899707@qq.com

Abstract. A large amount of system operation and equipment status data in current power systems are recorded and stored in text form, and text analysis has become an important basis for power system operation analysis. However, the current research still lacks relevant experience in applying it to the disposal of abnormal data spills in power grids. Therefore, the text-oriented data mining is first introduced in the paper to explain its basic concepts and implementation processes. Then, combining text grid-related data abnormal spillover requirements, text analysis and decision-making methods are proposed. Finally, considering actual development and application situation, the implementation benefits are analyzed. The experiment proves that the data abnormal overflow control technology studied in the paper can effectively guarantee the normal operation of the system.

1. Introduction

The exception handling of power grid is an important content of power system dispatching operation. Based on historical data mining analysis of abnormal handling in power grid, improving the handling capacity is a key issue in current research on abnormal handling of power grids. Moreover, considering that the historical data of grid data abnormal spill disposal are all recorded and stored in text form, the text analysis and decision-making method of grid data abnormal spill information has become an important issue in the research content.[1-2]

Based on the research of text analysis and decision-making methods in power systems, technical system design in terms of business correlation analysis and equipment defect statistics has been carried out in the paper. What’s more, the scenario matching and data mining for data exception spillover control are based on the structured text records. By matching with the elements of the problem to be solved, the corresponding strategies and implementation suggestions for the historical problems with the highest matching degree are provided to users for their decision analysis, so that efficient mining and analysis of huge historical data can be achieved[3].

2. Implementation Process of Text Library Type Data Mining

The biggest difference in the implementation technology between text library type data mining and database type data mining problems is that the basic data of database type data mining problems is actually structured data, while a large number of text records which are not structured are difficult to extract and mine key information. In addition, based on the current technical data, the implementation process of the text library type data mining problem is shown in Figure 1, which can be roughly divided into structured framework design, text record structured processing, and scene matching and data mining analysis.
Structured framework design is the implementation basis of text library type data mining problems, which is also an important prerequisite for efficient mining. In other words, the structured framework design is to extract the key elements from the general problems reflected in the text record, so that the specific content of the text record can be reorganized and stored according to the matching relationship with each key element. Moreover, the structured processing of text record refers to the process reorganizing text records based on a structured framework. For historical data, manual compilation-based methods are needed to re-compile it, while for newly added data, the design of the entry stage in the automation system can be used to standardize the format of text records to achieve automatic integration and storage[4-5].

3. Text Analysis and Decision-making Method for Abnormal Overflow Information of Power Grid Data

The recording and storage of grid data abnormal spill information are mainly text. Therefore, based on the implementation process of the text library type data mining problem in the paper, the structured framework design requirements and scenario matching-based data mining methods are mainly introduced to support the text analysis and data mining problems of abnormal overflow information in power grid data.

3.1. Text Structured Framework for Abnormal Overflow Information of Power Grid Data

During handling grid data abnormal spills, not only the type of data abnormal spills is needed, but also the importance of the equipment such as substations or lines that have an abnormal data spill in the power grid needs to be considered. Then, in line with relevant regulations or standards, targeted disposal strategies are formulated by assessing the impact of the data abnormal overflow on grid operation. Besides, according to the ideas mentioned above, the text structured framework of abnormal overflow information in power grid data can be divided into 3 levels and 7 aspects, as shown in Figure 2[6-7].
The first layer is the basic information layer including the data exception spill type and the device type, which is used to store the basic information for data exception spill handling.

The second level is the decision analysis layer including the accident event level prediction and the procedure specification, which is used to store the key information of the dispatching operation control personnel in the decision-making process based on the basic information.

The third level is the data abnormal overflow disposal layer including the adjustment of power plant output, adjustment of grid operation mode, and user load control, which is used to store the disposal measures specified after evaluation.

The structured design of abnormal data overflow information in the power grid is only a framework design, which needs to be further refined and improved in combination with actual conditions and management requirements in the actual system development and application process[8-10].

3.2. Scene Matching Index Definition and Mining Method

According to the structured text record results of abnormal overflow information in the grid data, in actual applications, it is necessary to further match the scene of the problem to be processed with the historical record, so that the most similar data abnormal overflow information and its disposal data can be filtered out for scheduling operation control personnel’s reference decision analysis. In addition, the scenario matching index is defined as the similarity degree of anomaly type, equipment type and accident event level prediction.

In one aspect type, the similarity index of the anomaly to be processed and the anomaly in the historical data is defined as follows.

$$\mu_{i,j}^{n} = \begin{cases} 0 & Q_i^n \neq Q_j^n \\ 1 & Q_i^n = Q_j^n \end{cases}$$ (1)

In formula (1), $\mu_{i,j}^{n}$ is the similarity index recorded in the $n^{th}$ aspect of the anomaly $i$ to be processed and the anomaly overflow record $j$ of some data in the historical data, and $Q_i^n$ and $Q_j^n$ respectively indicate the recorded contents of the $n^{th}$ aspect of the anomaly $i$ and the historical anomaly $j$ to be treated. When the two records are the same, the similarity index value will be 1. Otherwise, it will be 0.

The scenario matching is defined as the average of similarity in all aspects of data overflow and historical abnormal records in the three aspects of analysis to be matched, which can be expressed as follows.

$$R_{i,j} = \frac{\sum_{a=1}^{3} \mu_{i,j}^{n}}{3}$$ (2)

In Equation (2), $R_{i,j}$ is the scene matching degree of abnormal overflow $i$ of the data to be processed and abnormal record $j$ in historical data, whose limit is manually given. If the scene matching degree exceeds given limit, it indicates that the two scenes have a high degree of similarity, and their disposal methods and reference procedures can be used for the data overflow treatment. Otherwise, the similarity is low and has no reference value. The criteria mentioned above can be expressed as follows.

$$R_{i,j} \geq R_{\text{set}}$$ (3)

In equation (3), $R_{\text{set}}$ is the artificially given limit of scene matching.

4. Power Grid Data Simulation Experiment

In order to prove the effectiveness of the designed data exception overflow control technology, a virtual environment required for the information communication data management experiment under big data is built on the Windows 10 operating system. The experiment uses Ligence database server and the development language is C#. Moreover, during the simulation experiment, the data-less management
action is used as the comparison object of the experiment. Through comparison, the control validity of abnormal overflow in grid information and communication data is verified.

Through the design of the above simulation environment, the data management technology proposed in the paper is used to simulate the abnormal overflow of information and communication data, where two technologies are used to test the encryption strength of the information communication data, and the comparison results of the tests are shown in Figure 3.

![Figure 3](image)

**Figure 3** Comparison of encryption strength of information data abnormal overflow control

It can be seen from Figure 3 that the encryption strength with the data abnormal overflow control technology is stronger. During abnormal overflow control, as signal-to-noise ratio gradually decreases, the strength of data abnormal overflow control technology can be better encrypted and the control performance is better.

In addition, the data exception overflow control technology and data-less overflow exception technology are used to test the data processing efficiency. The test comparison results are shown in Figure 4.

![Figure 4](image)

**Figure 4** Comparison of data processing efficiency

It can be seen from Figure 4 that compared with the data-less control technology, the data exception overflow control technology has a significant difference in the efficiency of data processing, which indicates that the data exception overflow control technology can be used to process data more efficiently, so that the security of information communication will be greatly improved.
5. Conclusion
Focusing on the characteristics of the widely used text-type data recording method in the grid data abnormal overflow information, the implementation process of text library-type data mining is introduced in the paper. Then based on the actual data overflow information of the grid, the research focuses on its structured framework design, scenario matching index definition, and data mining methods, which has been practically applied in the power grid of certain province in China. With significant results, the method proposed in the paper has a prominent effect on improving the efficiency of dispatch exception handling.

Reference
[1] Li Li, Hua Kui, Jiang Yan, et al. Review of research on key technologies for multi-source heterogeneous data processing of transmission lines [J]. Guangdong Electric Power. 2018 (08)
[2] Ma Qiang, Guan Yan, Bai Yingwei, et al. Research on equipment correlation technology of power grid business system based on big data [J]. Automation and Instrumentation. 2018 (04)
[3] Du Xiuming, Qin Jiafeng, Guo Shiyao, et al. Text mining of typical fault cases of power equipment [J]. High Voltage Technology. 2018 (04)
[4] Yan Danfeng, Zhu Qiao, Tang Yi, et al. Visualization of the status data of power transmission and transformation equipment [J]. Computer Applications. 2017 (S1)
[5] Yu Zhenhua, Yang Shiyong, Shu Zhengyu, et al. Research on identification and classification method of power network operation monitoring information based on clustering [J]. Hubei Electric Power. 2016 (04)
[6] Qiu Jian, Wang Huifang, Ying Gaoliang, et al. Text information mining technology and its application in circuit breaker full life state evaluation [J]. Automation of Electric Power Systems. 2016 (06)
[7] Huang Tianen, Guo Qinglai, Sun Hongbin, et al. Model-data hybrid-driven power grid security feature selection and knowledge discovery key technologies and engineering applications [J]. Automation of Electric Power Systems. 2019 (01)
[8] H.W. Hui, C.C. Zhou, S.G. Xu, F.H. Lin, A Novel Secure Data Transmission Scheme in Industrial Internet of Things, China Communications, vol. 17, no. 1, 2020, pp: 73-88.
[9] F.H. Lin, Y.T. Zhou, X.S. An, I.You, K.R. Choo, Fair Resource Allocation in an Intrusion-Detection System for Edge Computing: Ensuring the Security of Internet of Things Devices, in IEEE Consumer Electronics Magazine, vol. 7, no. 6, 2018, pp: 45-50. doi: 10.1109/MCE.2018.2851723.
[10] J.T. Su, F.H. Lin, X.W. Zhou, X. Lv, Steiner tree based optimal resource caching scheme in fog computing, China Communications, vol. 12, no.8, 2015, pp: 161-168.