Research on Data Quality Evaluation and Comprehensive Analysis Decision-Making under the Integration of Regulation and Control

ke Li, Yuxing Lin

Zaozhuang Power Supply Company of State Grid Shandong Electric Power Company, Zaozhuang, 277000, China

Abstract. With the advent of the era of big data, our society has become a "data society". The development and utilization of data resources has become an important force to promote social progress, and data quality is the basis for the validity and accuracy of data analysis conclusions. In view of the current situation that it is difficult for the power grid dispatch center to conduct a comprehensive analysis of the monitoring and dispatching big data, a method for evaluating the quality of operation data in substations based on electrical physics knowledge and manual data discrimination is proposed, and an online data mining and risk early warning technology system under the integration of regulation and control is developed. Early warning indicators are identified to realize early warning of power grid security risks through correlation analysis. This system has been applied in Shandong Dispatching Center.

Keywords: Integration of dispatching and supervisory control; Data mining; Risk assessment; Risk early-warning.

1. Introduction

The traditional power grid operation management mode is no longer suitable for the current development trend of the power grid with the continuous deepening of the reform of the power system. In order to promote the development of smart grid, the power grid has launched an integrated operation and management model [1]. The integration of regulation and control means the integration of power dispatching and monitoring. The high integration of the centralized monitoring business of power transmission and transformation equipment operation and the power grid dispatching business are promoted by optimizing and integrating power grid dispatching and equipment operation resources [2]. However, the monitoring and dispatching information mastered by dispatchers has grown rapidly, and power grid data has the characteristics of big data with the advancement of State Grid Corporation's informatization construction and the development of power grid equipment operating status detection technology. At the same time, the complexity and expansion of the power grid scale also increases the difficulty for dispatchers who implement the integrated management of regulation and control to quickly and comprehensively grasp the security risk situation of the power grid. Although the PSASP dynamic safety assessment and early warning system has been researched and developed, operators cannot use various analytical and research software to analyze and study the operation status of the power grid due to the change of operation mode in the continuous development of the accident [3] [4]. Therefore, it is of great significance to study the big data mining and grid risk identification and early warning system based on grid monitoring and dispatching information under the integrated management mode of regulation and control to improve the awareness level of dispatching operators on the safe operation status of the grid and maintain the reliable and economical operation of the system.

At present, Shandong Electric Power Group will adopt a centralized monitoring method for all 500 kV substations, and the main station side of the monitoring system will monitor the operation mode and equipment status of each station in real time. Therefore, it is necessary to conduct in-depth mining of these different types of "massive" information after the power grid control center has mastered more abundant power grid and substation equipment operation information.

A method for evaluating the quality of operating data in substations based on electrical physics knowledge and manual data discrimination is proposed, and a data mining and risk early warning system under the integration of regulation and control according to this method is developed based on the above background. The correlation between different types of monitoring data sets such as SCADA data, substation monitoring data, meteorological data, etc. is established, and then conduct comprehensive analysis to find the hidden logical relationship behind many clues. The early warning of accident risks can be realized through the identification of early warning indicators, so that low-value density data can fully have decision-making value, and the ability to control the risk of power grid dispatching operation and the ability to respond to emergencies in the "integration of regulation and control" mode can be improved. Finally, the optimal integration of power grid dispatching and equipment operation resources is used to promote the high integration of centralized monitoring of power transmission and transformation equipment operation and power grid dispatching business.

2. Theoretical Research

2.1 Data Mining

Data mining is the discovery of patterns from a large amount of data, and the so-called model is a form of information used to describe the data. The analysis methods of data mining mainly include: association analysis, clustering mode, concept description, deviation analysis, classification mode and time series evolution mode, etc. Power grid data has become increasingly rich, and data mining based on analysis, artificial intelligence and simulation has also been widely used with the development of condition monitoring technology. In the past, data mining analysis methods could only rely on limited experimental statistics or expert experience for modeling, resulting in great limitations in the application of these models due to the lack of data. Nowadays, multi-source heterogeneous data such as equipment and environment are continuously accumulated and accumulated with the development of equipment monitoring and detection technology.

Starting from the thinking of finding an effective data mining and analysis method, this paper looks for the common characteristics of various correlations in the system under the background of increasingly abundant data related to the operation of power grid equipment. On the basis of geometric principle analysis and mathematical concept derivation, a non-parametric identification algorithm for complex correlation based on maximum information and a dynamic correlation sorting early warning algorithm are designed. the identification and measurement of implicit or unknown correlations without relying on model parameters are realized through online analysis and regulation of large data sets, so as to mine data.

2.2 Security risk warning

Risk early warning is to analyze the risk level on the basis of identifying and evaluating the uncertain factors and safety hazards faced by the power grid, and give risk warnings according to the risk status. Risk early warning is a means of risk management. Based on risk management theory, risk early warning should follow data monitoring - risk identification - risk assessment - risk forecasting.

Risk identification refers to obtaining information such as abnormal operation of the power grid through correlation analysis, and identifying potential security risks in the power grid based on the operation of the power grid and equipment information.

Risk assessment refers to the evaluation of the safety level of the power grid to obtain a variety of operational risk indicators (normal, abnormal, severe). To conduct a comprehensive assessment of power system risks, it is necessary to study and set risk indicators. At present, there have been studies on the risk index system at home and abroad, including discussing the basic concepts and importance of power system risk assessment, summarizing risk assessment models and calculation methods, probabilistic risk indicators, formulating the principle of risk index selection, and using AHP to evaluate risk indicators, etc. [5] [6] [7]. However, the above indicators can only be evaluated for

the time scale of the power grid, which is not general and cannot cover all application scenarios. Due to the lack of consideration of various influencing factors of the power grid, the risk indicators of the above research are still not perfect. This paper adopts a risk index system combining cross-line-driven and event-driven to measure the operation status of the power grid from the dimensions of time and space [7] [8]. This paper adopts the risk classification standard of the British Health and Safety Commission - the ALARP principle to formulate a framework system for grid security risk classification [9]. The risk degree is defined as normal, abnormal and severe according to the size of the severity. The quantitative value of different risk indicators needs to be analyzed and determined by combining a large number of simulation analysis calculations and expert knowledge.

Risk warning is to give an alarm according to the risk indicators obtained from the risk assessment results. When the monitoring index reaches or exceeds the risk value set in the risk assessment classification, before the risk expands, it will give an early warning, and green (normal), yellow (abnormal), and red (severe) are used to indicate different levels of alarms.

3. System Design

3.1 The Process of Data Quality Evaluation

The process of data quality evaluation is shown in Fig.1.

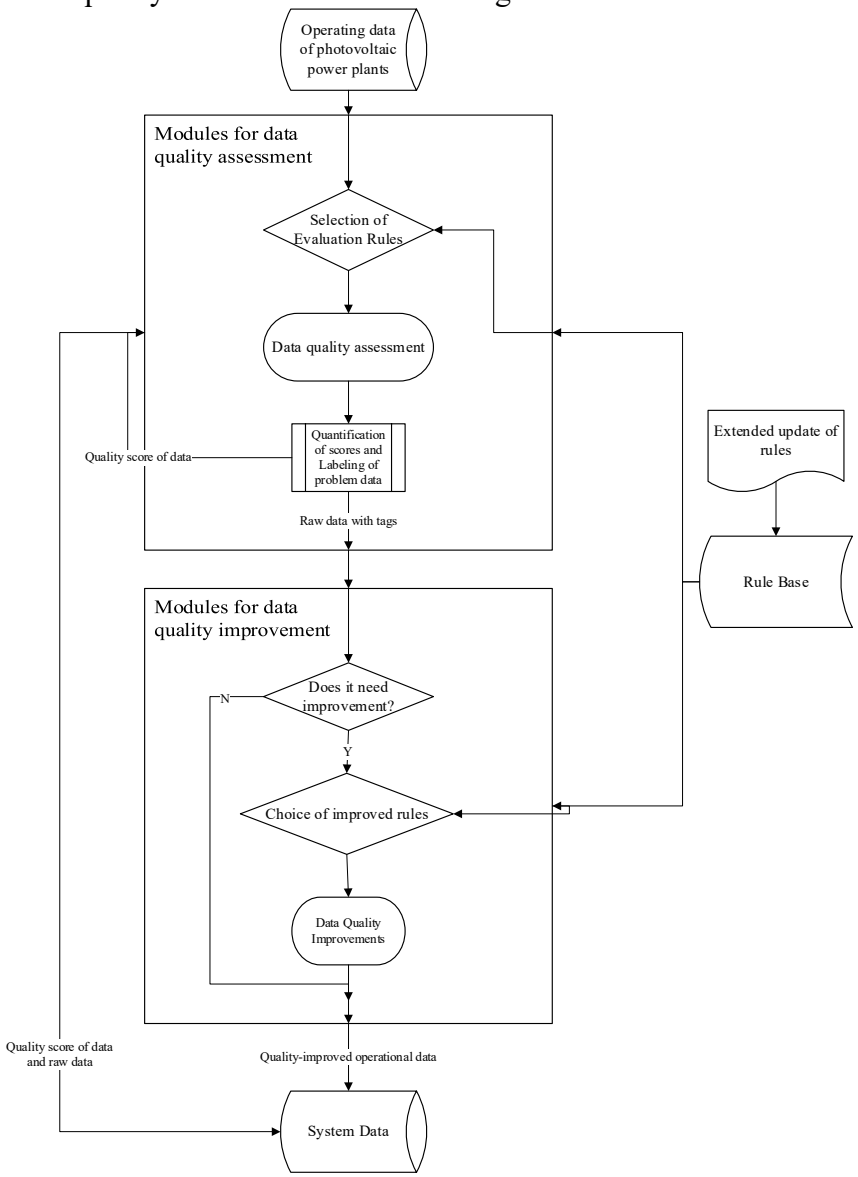


Fig.1 The process of data quality evaluation

After the substation operation data enters the data quality evaluation system through the acquisition channel, it is received by the data quality evaluation module. The data quality evaluation module identifies the type and quantity of access data, selects data quality evaluation rules, performs data quality evaluation, and outputs evaluation results. After receiving the data quality evaluation results and data, the data quality improvement module first determines whether improvement is required. If yes, the data quality improvement module selects the appropriate data quality improvement rules, execute the improvement function, and then output the improved data and the original evaluation results, otherwise output the data and results directly.

According to the actual needs of the substation connected to the power grid, the data evaluation rule base and each module are configured. Data evaluation rules include electrical physical laws and manual data discrimination. The electrical physical law is the physical law that the electrical physical quantities contained in the substation operation data obey, such as Ohm's law; Manual data discrimination is the transformation of the idea of manual data screening into data quality evaluation rules applicable to the system. This kind of rule is different from the physical law, which uses the changing characteristics of the data itself, such as historical data, changing trends, and the changing characteristics of related data to analyze and improve the quality of the data. The rule base is extensible and supports the update, deletion and addition of rules.

Evaluation indicators include accuracy, completeness, consistency and timeliness. Accuracy is used to characterize the deviation of the data value from the true value. Integrity is an indicator used to characterize whether the data is complete. Consistency mainly reflects whether the data conforms to the system requirements in terms of format and type. Timeliness reflects the degree of delay between data reception time and data time. he problem of data quality is mainly manifested in the aspects of accuracy and integrity due to the particularity of the power system.

3.2 System Architecture Diagram

Under the above data quality evaluation methods, corresponding systems are developed. The system is supported by a distributed analysis and computing platform, effectively integrates grid monitoring and detection of online and off-grid data, detects abnormal grid operation, and quantifies its risks. The risk of safe operation of the power grid is assessed, and the operating status of the power grid is effectively perceived, which provides a basis for dispatchers to make decisions. The overall structure of the system includes data source layer, data interface layer, data platform layer and system application display layer.

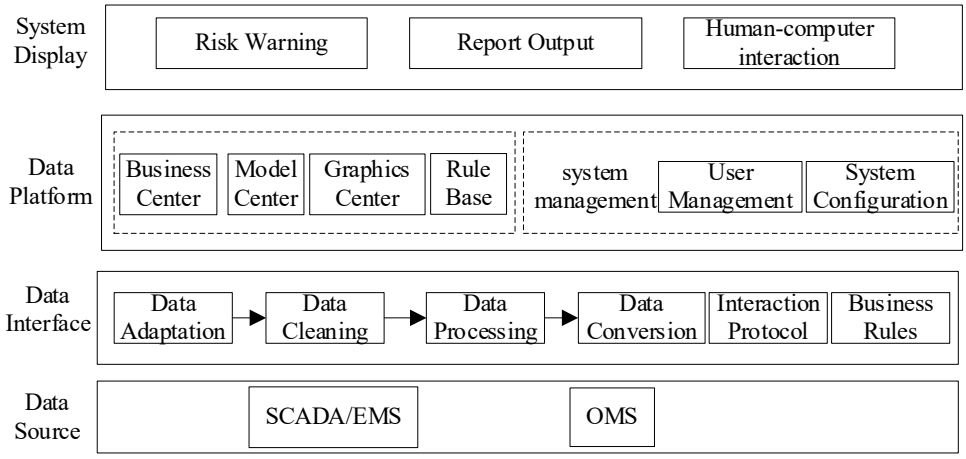


Fig.2 The system architecture diagram

The data mainly comes from the energy management system and the production management system OMS. The data of each system in the data source layer is collected through the design and development of different interface adapters, which provides the required equipment monitoring and

power grid parameters for the risk early warning system, and realizes the data exchange between the regulation early warning system data and different application systems.

The data interface layer is mainly based on the Java platform, CentOS x64 system and MySQL database. The interface for obtaining data from the data source system is studied, and the data is processed based on the interactive data protocol and business standard rules, and finally the required data is stored in the database of the risk warning system.

The data platform is mainly used to analyze and mine control data, which includes data correlation analysis, risk identification, risk assessment quantification and risk early warning.

The data display is mainly the visualization of the system, which provides intuitive and clear system analysis results and risk situations through the visual interface display of human-computer interaction.

4. Demonstration of System Innovation Functions

4.1 Difficulty Assessment of Frequency Modulation

Analyze the adjustment characteristics of the power plant's ACE in time and space, and understand the recent ACE adjustment performance, which is convenient for dispatchers to pay attention in advance. At the same time, the ACE adjustment difficulty of each plant is quantitatively evaluated, and the ACE adjustment difficulty and time period are given.

4.2 Difficulty Assessment of Pressure Regulation.

Some plants with poor voltage adjustment can be found through the statistics of the bus voltage whose voltage fluctuation is greater than the specified value, which is conducive to taking pre-control measures in advance, and also provides a basis for evaluating the adjustment level of each plant. Through the analysis of the coordination of voltage adjustment, it can be concluded whether the power plant and the power station can coordinately adjust the voltage level of the station. At the same time, the difficulty of voltage adjustment can be quantified through the statistical analysis of the limiting range of the voltage near the upper and lower limits, and the final adjustment difficulty index can be given.

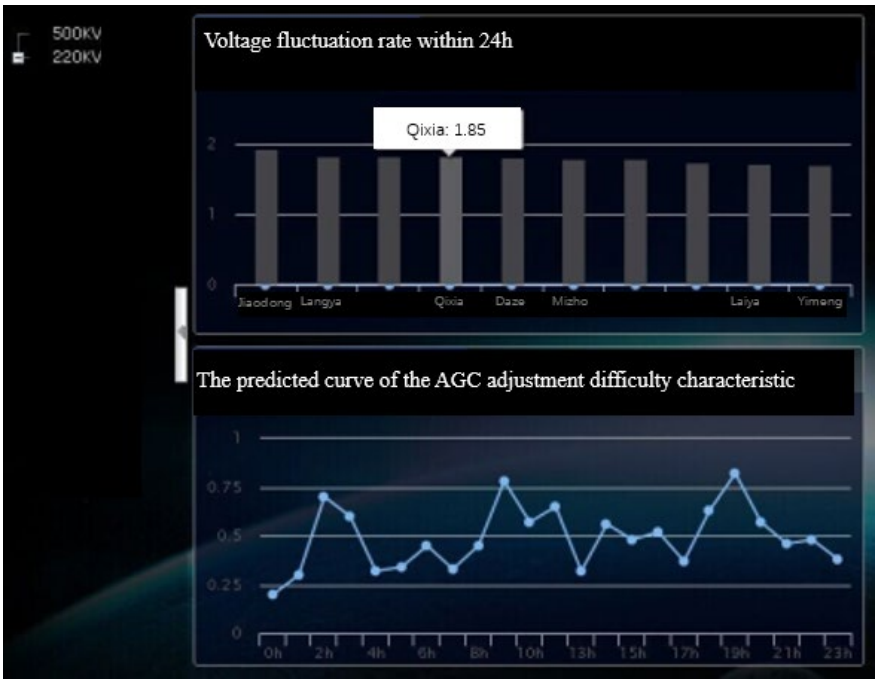


Fig.3 The interface effect of frequency modulation and voltage regulation

4.3 Analysis of Load Characteristics

In addition to the traditional statistical indicators of load characteristics, this system analyzes the degree of fit between regional and network-wide loads. The judgment of whether the regional adjustment characteristics are consistent with the whole network is obtained, and the analysis of the simultaneous rate of the power grid load can more effectively guide the orderly allocation of electricity quotas.

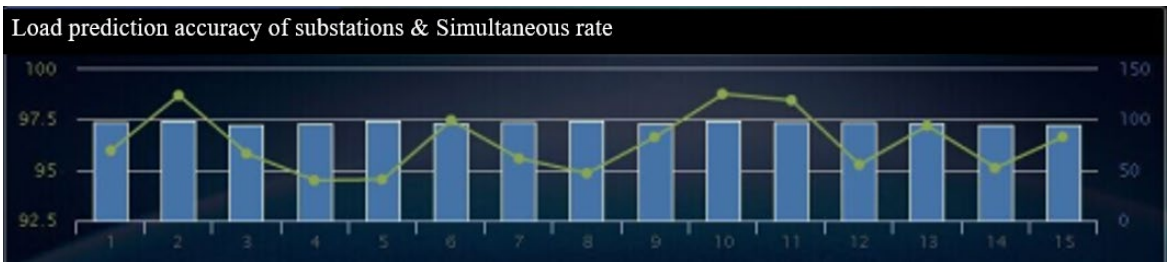


Fig.4 The Interface Effects of Analysis of Load Characteristics

4.4 Assessment of Security Risks

Risk assessment starts from the perspectives of power grid, equipment, environment, etc. The uncertain factors faced by the system are comprehensively considered, and the operating status of the system measured by formulating security risk indicators is to discover hidden security risks in the power grid in time.

4.5 Safe and Stable Online Warning

The risk assessment result is displayed visually after the risk level is obtained according to the power grid security risk index. Different levels of alarms on the entire network map are represented by green (normal), yellow (abnormal), and red (severe), so that the dispatcher can clearly view the security risk status of the entire network on the home page, so as to issue targeted instructions, and take safety precautions.

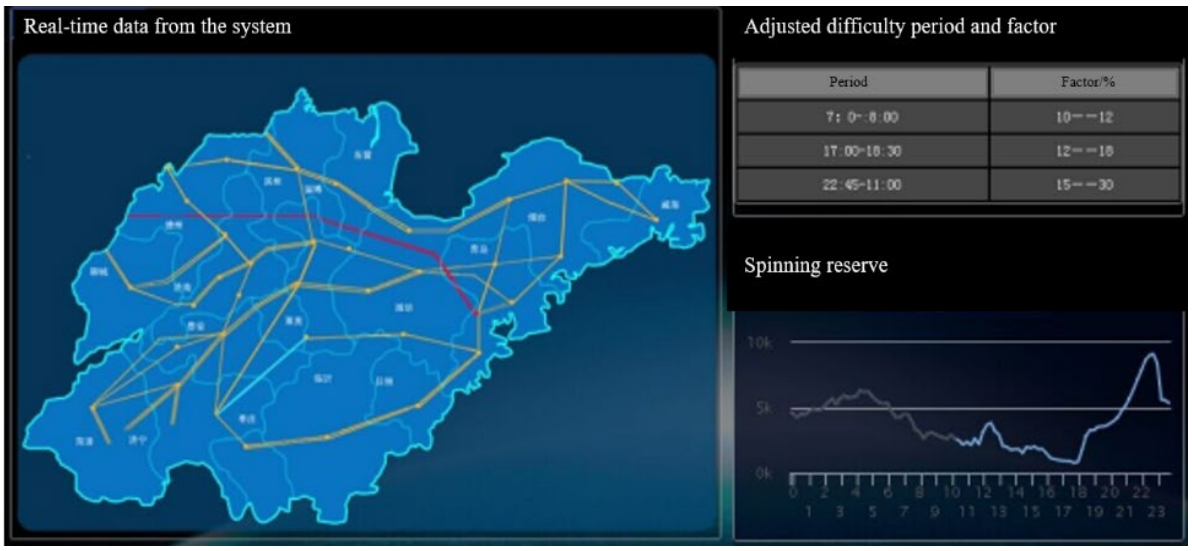


Fig.5 The interface effect of data mining and risk early warning

4.6 Report Output

This system not only has a good display and human-computer interaction interface, but also has the function of automatically generating scheduling, monitoring and power grid logs and outputting them.

5. Conclusion

The online data mining and risk early warning system under the integration of regulation and control has been put into use in Shandong Electric Power Dispatching Center, and will also be applied to the municipal dispatching center. Engineering practice shows the early warning system integrates dispatching and monitoring information. An integrated data platform for regulation and control are provided, and the mining, evaluation and early warning of grid regulation data are realized. The system adopts distributed storage, parallel processing and analysis, which has high computing efficiency and greatly improves the work efficiency of dispatchers. At the same time, the system is scalable. the requirements of system integration are satisfied because of the standardized design. The visual interface shows the risks and problems existing in the current operation of the power grid, which provides online information reference for the work management of dispatchers.

The regulation integrated online data mining and risk early warning system proposed in this paper improves the current situation that the dispatch center is difficult to control various risk factors for the safe operation of the power grid. The accident risk can be early warning by identifying the early warning indicators, which provides an effective analysis tool for the power grid dispatching operation. In the next step, relevant research can be carried out on the auxiliary decision-making of risk control. The system will automatically provide a reasonable dispatching strategy and plan for the staff to refer to when the power system operates abnormally, which could improve the ability of the power grid to deal with risks and avoid accidents.

References

[1] Lv Hongbo, Interation of power grid dispatching and supervisory control in operation and management mode, CN: North China Electric Power University, 2011.

[2] ZHANG Caiyou, ,DING Yimin, FENG Hua, “Reflection on integration of operation and maintance for transformation equipments,” Zhejiang Electric Power. vol. 30, pp. 50-53, March 2011.

[3] CAI Bin, WU Sunong, WANG Shiming, WU Weuchuan, ZHANG Boming, ZHANG Changlin, ZHOU Dongliang, “Power Grid on-line security and stability analysis and forewarning system,” Power System Technology. vol. 02, pp. 36-41, January 2007.

[4] YAN Jianfeng, YU Zhihong, TIAN Fang, ZHOU Xiaoxin, “Dynamic security assessment & early warning system of power system,” Proceedings of the CSEE. vol. 28, pp. 87-93, December 2008.

[5] ZHANG Jianhua, ZHANG Guohua, DUAN Manyin. “Indices system and methons for urban power grid security assessment,” 2009 Asia-Pacific Power and Energy Engineering Conference, March 28-30,2009, Wuhan, China.

[6] LU Bo, TANG Guoqing, “Application of risk-based security assessment in power systems,” Automation of Electric Power Systems. vol. 22, pp. 61-64, November 2000.

[7] WEN Yunfeng, CUI Jianlei, ZHANG Jinjiang, GUO Chuangxin, HUANG Lianggang, SHEN Kun, “Design of a security risk management system for power system dispatching and operation Part one Concepts and design of architecture and function,” Automation of Electric Power Systems. vol. 37, pp. 66-73, May 2013.

[8] CUI Jianlei, WEN Yunfeng, ZHANG Jinjiang, GUO Chuangxin, HUANG Lianggang, “Design of a security risk management system for power system dispatching and operation Part two, Risk index, assessment methodologies and application strategies,” Automation of Electric Power Systems. vol. 37, pp. 92-97, May 2013.

[9] CALLAGHAN B, WALKER T. Reducing risks, protecting people: HSE ‘s decision-making process. England: HSEBOOKS,2001.