

Research on Data Mining Technology of Company Operation Monitoring

Guang Chen^{1, a}, Ju Qiu^{2, b}, Tingzi Pan^{3, c}, Dongxiao Niu^{4, d}, Di Pu^{5, e *},
Ting Zhao^{6, f}, Haini Qu^{7, g} and Lihan Qin^{8, h}

¹State Grid Energy Research Institute Co. Ltd., Beijing, China;

²Beijing China Power Information Technology Co., Ltd., Beijing, China;

³Beijing China Power Information Technology Co., Ltd., Beijing, China;

⁴School of Economics and Management, North China Electric Power University, Beijing, China;

⁵School of Economics and Management, North China Electric Power University, Beijing, China;

⁶Global Energy Internet Research Institute Co. Ltd., Beijing, China;

⁷State Grid Shanghai Municipal Electric Power Company, Shanghai, China;

⁸State Grid Jibei Electric Power Company Limited, Beijing, China;

^achenguang@sgeri.sgcc.com.cn, ^bqiuju1@sgitg.sgcc.com.cn, ^cpantingzi1@sgitg.sgcc.com.cn,
^dniudx@ncepu.edu.cn, ^epd@ncepu.edu.cn, ^fzhaoting_00@qq.com, ^gqin.lihan@nc.sgcc.com.cn,
^hhattiequ@126.com

* The Corresponding Author

Keywords: Power grid corp operation monitoring; Data mining technology; Index system construction; Database building

Abstract. This paper mainly studies the operation monitoring data mining technology of Power Grid Company. In order to meet the need of operation monitoring business in the era of big data, this paper deeply studies the data mining technology suitable for company operation monitoring, builds a database of data mining technology and forms a data mining technology subject database according to the different application of company operation monitoring data mining technology to solve the problem of data mining for the operation and monitoring business. The research on the company's operation monitoring data mining technology helps company analyze the laws and trends of operation monitoring business, promote the company's quantitative decision-making and enhance the company's decision-making efficiency. At the same time, this study can reduce duplication of resources, increase the value of data assets, and ultimately improve the lean operation and management of power grid.

Introduction

The emergence of large data is triggering a profound global change in technology and Commerce [1]. At the same time, the continuous development of large data also has an important impact on the future development of electric power enterprises [2, 3]. In recent years, the application of data mining in power system has been deeply analyzed at home and abroad. The research fields mainly focus on, such as dynamic security assessment, load economic dispatch, load forecasting and fault diagnosis [4, 5].

Paper [6] takes the operation and monitoring center of a city power grid as an application object, relying on operation management and control system, and researches data mining to explore potential correlation, build operation monitoring index system, set up company risk early warning and prevention and control system, and introduce data mining technology to carry out meticulous operation analysis, quickly and accurately locate loopholes in management, so as to provide guarantee for improving economic efficiency and management efficiency. Paper [7] design and implement the power marketing analysis system based on Data Mining Technology, including data mining modeling function and analysis of scene function, expounds how to apply data mining technology to the operation and management of power supply and power supply business in power grid enterprises through the construction of the analysis of the scene model. At the same time,

according to the business logic of different topics, different data mining algorithms are selected and data models are established, which strengthened the enterprise level analysis ability of data for power supply and power supply business, and made the company's resource allocation better, internal control power stronger, management efficiency and overall efficiency higher. Based on the development requirements of the national Power Grid Corp operation monitoring (control) business, paper [8] analyzes the current situation and existing problems of the operation monitoring (control) work, and put forward the construction model of operation monitoring (control) system based on big data mining technology, business theme database and data asset correlation library. It has a good reference for further improvement of operation monitoring (control) in power grid enterprises. Paper [9] mainly studies the business content, index system, information collection and processing flow of the visual management system of power operation monitoring and control, expounds the analysis method of electric power operation, puts forward the early warning of electric power security and the emergency handling mechanism based on case reasoning, and realizes the visualization of the power management process. Based on the research of the equipment state monitoring of large data technology, paper [10] discusses the application of the key technology of large data mining in state monitoring. Paper [11] establishes a unified and complete production and management, monitoring index system of comparable weight, production and operation of the company, through the construction of monitoring index system, timely, comprehensive and accurate grasp of the company's operating performance, management performance and core resource discovery, coordinate and solve company incompatible problems and carry out multidimensional analysis and depth mining business cross, effectively ensuring the stable development of the company.

Based on the above analysis, it can be found that the application of data mining technology in power system is mainly focused on the part of operation monitoring of electric power company, and there is a lack of comprehensive and systematic data mining technology library that can be applied to all operations of power companies. Therefore, based on the requirements of the company operation monitoring data mining, this paper has carried out the research of the Power Grid Corp operation monitoring data mining technology. Through the construction of data mining technology library, the problem of data mining of the company's transportation supervision service will be solved.

The Construction of Company Operation Monitoring Index System

In order to support the overall development of the operation monitoring business, it is particularly important to build a scientific and reasonable index system which covers the operation management and core business of corporations.

The index system of the operation monitoring is derived from all levels of corporation strategy and operation management. The indicators at all levels are interrelated and mutually supported to ensure the ultimate goal is the realization of the corporation's strategic objectives. According to the idea of index system construction, the overall structure to form the index system is illustrated in the table 1.

According to Table 1, we can figure out that the corporation operation monitoring index system consists of five aspects: external environment monitoring, key indicators monitoring, operation status monitoring, key work completion monitoring, operation index unusual alteration and early warning monitoring. The construction of corporation operation monitoring index system is helpful to sort out the data mining demand for the corporation operation monitoring business, laying a foundation for the construction of the data mining technology library of corporation operation monitoring.

During the period of designing corporation operation monitoring index system, the scalability is fully considered. With the deep development of business, the index system is becoming more and more perfected, gradually covering all aspects of the corporation's daily operation management. In the meantime, with the expansion of the corporation's strategic focus, the performance evaluation

indicators of the corporation will be further refined and the index system will be expanded and improved.

Table 1 Corporation operation monitoring index system

The first class indexes	The second class indexes	The third class indexes
External environment monitoring	index monitoring	Official PMI; Choi new PMI; CSPI
	Electricity power analysis	Total Electricity Consumption
	Macroeconomic analysis	Gross Domestic Product(GDP); Consumer Price Index(CPI); Producer Price Index(PPI); The proportion of the primary, secondary and tertiary industries in the GDP; Total Retail Sales of Consumer Goods(SRG); Money stock; Currency reserves; Industrial added value
	Policy analysis	Electric Power Reform-Price Reform; 13rd Five-Year Development Plan and so forth
Key indicators monitoring	Performance evaluation index	Operating performance; Execution efficiency of operation plan(Line loss rate); Operation and maintenance efficiency of power grid(The system reliability of the power grid); The completion rate of power marketing tasks(Electricity sales); The quality of power supply and customer service; Human capital efficiency
	Completion of the same industry' benchmarking	The benchmarking of provincial companies; The benchmarking of municipal companies; The benchmarking of professional institutions
Operation status monitoring	Planned budget	Comprehensive plan implementation; Financial budget implementation
	Investment construction	Fixed assets investment; Capital construction; Production technical reform; Non-production technical reform; Information construction; Marketing special; Sporadic procurement projects; Construction scale of project; Production scale; Under construction scale
	Production operation	Number of accidents in production safety(Power grid accident; Equipment accident); Network load
	Purchase and sale of electric power	The quantity of electricity purchase; The quantity of electricity sales; Direct purchase of electricity; Comprehensive line loss rate; Average price for electricity purchase; Average price for electricity sale; Price difference between purchase and sales
	Operation Benefits	Total profit; controllable expense; ROE; Net operating income; Total operating costs
	Quality service	Telephone traffic; Completion rate of quality service; The fulfillment rate of fault repair arrival time; The completion rate of fault repair arrived on-site time; the number of work orders with customer complaint reports
	Resource situation	Manpower; Financial resources; Material resources; Electric network facilities; Installed capacity
	“Twenty-four solar terms’ work completion rate	—
Key work completion monitoring	Comparison over the same historical period	—
Operation index unusual alteration and early warning monitoring	Unusual work orders co-processing situation	Co-processing work orders; The number of completion of work orders; the completion rate of worker orders
	Early warning of operational risks	Implementation of integrated planning and budget projects procurement; Line blackout situation; Quality of power supply; The area with high Complaints incidence; Fault repair concentrated area

Requirement Analysis of Data Mining on Company Operation Monitoring Business

From the perspective of business attributes, during the functional requirement analysis, the operation monitoring (control) workbench system is mainly used to support the operation and management of power grid enterprises about the company daily operations management activities and work, such as the comprehensive monitoring on the company operation and management, operation analysis, coordination control, panorama display, etc. The core business of the center includes mainly four aspects: comprehensive monitoring, operation analysis, coordinated control and panoramic display, which is shown in Fig. 1.



Figure 1. Operations monitoring center core business

24 hour on-line dynamic monitoring of the Company's comprehensive performance, operation status, core resources and key business processes can be carried out through the construction of monitoring models, carding of indicator systems and setting of thresholds for indicators so as to timely find out the changes in the operation of the company and problem and realize automatic security early-warning.

Operation analysis mainly focuses on the daily operation and management activities of SGCC, conducts an in-depth study on the interconnection of different businesses and their impact on the overall operation of the company, which analyzes the changes or problems caused by the company in monitoring and the overall performance, development capability, competition ability, risk management and other aspects, observing the operation of the transaction or the causes of problems, influencing factors, existing risks, and making recommendations for the leadership of the company.

Coordinated control is based on the results of comprehensive monitoring and operational analysis and coordinates with related departments to handle the problems and risks in the operation and management of the company through internal coordination, horizontal coordination and vertical control.

The panorama display mainly focuses on the group operation and intensive management of the company, which makes full use of advanced visualization technology and large-screen display platform to formulate display solutions for different objects and achieve a full range of display.

From a global perspective, the main task of data mining system is to help operations monitoring (control) workbench system to complete the above tasks to meet the above requirements of several functions.

Study on Company Operation Monitoring Data Mining Technology Library

Classification of Company Operating Monitoring Data Mining Technology. With the advancement of science and technology, the development of data mining technology is getting more and more perfect, and according to different classification methods, different classification can be carried out. This article divides the common data mining technologies into four categories according to the operational monitoring business requirements: one is statistical analysis, the other is comprehensive evaluation, the third is prediction and early warning, and the fourth is visual display.

The data mining of company operations monitoring business is shown in table 2.

Table 2 Data mining technology classification table

classification	Classification Standard	Data Mining Technologies	Advantages	Shortcomings
Statistical analysis	Information processing comprehensive monitoring; Statistical analysis of operational analysis	variance analysis	The method is simple and practical	Information analysis accuracy is not high
		Principal component analysis	The method can analyze the correlation between data	Will produce information missing
		Regression analysis	Simple and practical, especially in multi-factor model analysis; correlation coefficient or significance test makes predictive model become more credible method.	The model is simpler and more restrictive.
		Association rules algorithm	Able to find frequent episodes of aggregated knowledge in big data	Need to scan the database multiple times, generating too many candidate sets
		Bayesian classification algorithm	Keep learning, optimizing models, non-linear relationships	Model test is not mature, missing key variables, result will decline in accuracy
		Decision tree (decision tree) algorithm	Good classification speed, high precision, the generation process is simple and the rules generated are easy to understand	It is sensitive to noise and tends to choose more-valued attributes when testing attribute selection
		Clustering algorithm	easy to understand	Sensitive to the initial center point, the number of clusters needs to be determined in advance; easy to fall into local optimum
		Gray correlation analysis	Evaluation results are more objective and accurate	Evaluation steps are not standardized, the calculation process is extremely complicated
comprehensive evaluation	Power Grid Enterprise Integrated Performance Management; Power Grid Enterprise Investment and Asset	Expert evaluation method	It can well integrate lessons learned	Subjectivity is too high
		Analytic hierarchy process	able to rank the importance of an indicator	The method is more subjective
		Entropy	Objective and	Computation is

Management; Power grid business management; Coordinated control Process Management;	method	practical	more complicated
	Rough theory	set Important content of pattern recognition and machine learning	High computational complexity, easy to fall into the local optimal solution
	Fuzzy comprehensive evaluation method	The evaluation result is more accurate	Easily lead to membership information loss
	Data envelopment analysis	Practical and operational	Algorithm results are not ideal, too physical

Continued Table 2 Data mining technology classification table

prediction and early warning	the monitoring alarm, penetrating query of comprehensive monitoring Trend / extrapolation, predictive modeling, outcomes optimization in operations analysis	Time series algorithm	Applies to the existence of time data	Higher requirements for the regularity of data
		Artificial neural networks	Easy to operate, easy to understand for the user; suitable for solving complicated problems of the internal mechanism	Easily fall into the local optimal solution; the training results are unstable; the model is particularly dependent on the training samples; under the premise of high-quality training results, the training time is greatly increased and the convergence speed is difficult to control.
		Gray prediction theory	Simple and convenient; The required sample data less, can be applied to discrete, regular data is not strong	Fitting accuracy is low
		Support vector machines	Supports small sample learning; owns good robustness	Large-scale training samples difficult to implement
		Genetic algorithm	owns good global search ability; optimization results have nothing to do with the initial conditions; scalable, easy to combine with other algorithms	The algorithm is more complex; convergence performance is poor, premature convergence prone to the situation; local search ability

				is poor, resulting in relatively simple genetic algorithm time-consuming
		Particle swarm optimization	Suitable for solving real problems, the algorithm is simple, easy to calculate, solving speed, parallelism	Easy to fall into the local optimal solution
		Co-integration theory	Analysis method is more intelligent	Easy to get the local optimal solution
visual display	the state tracking, data collection of comprehensive monitoring; descriptive analysis in operational analysis; panoramic display	Data visualization technology	Clear, intuitive, clear and easy to understand	Unable to carry out quantitative analysis, and the conclusion is not sufficient for some data processing is not enough.

According to the classification standard, advantages and disadvantages of data mining technology, all kinds of data mining technologies are centralized and unified and can be quickly matched to meet the needs of diverse and full-service data mining technologies.

Construction of Company Operations Monitoring Data Mining Technology Library. In order to improve the performance of data mining technology in the company's operation monitoring business and provide the company with better data support and decision reference, the algorithm can be appropriately improved according to the advantages and disadvantages of the algorithms in table 2, and we propose a more specific and accurate data mining technology to build a comprehensive and diversified library of data mining technologies through combining the data mining technology and the company operation monitoring business (as shown in Fig. 2).

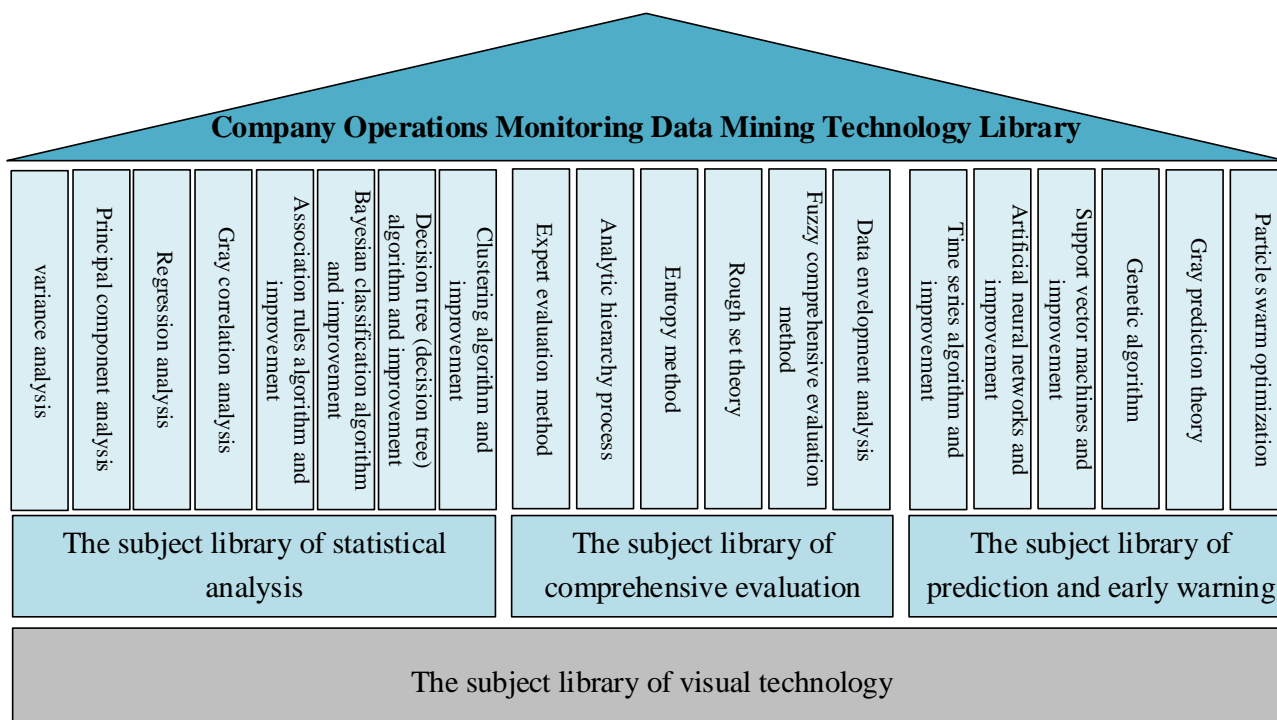


Figure 2. Data mining technology theme library

Among them, the main technical topic library construction is as follows:

The Subject Library of Statistical Analysis

Statistical analysis techniques include variance analysis, principal component analysis, regression analysis, association rules algorithm and improvement, Bayesian classification algorithm and improvement, decision tree algorithm and improvement, clustering algorithm and improvement, gray correlation analysis, etc. They are mainly used in operational analysis of market competition, business development capabilities, risk management and other aspects, to put forward management recommendations and promote the company safe, healthy, orderly and efficient operation. The advantage of the indicator can be used to analyze the status of power grid companies, such as the security status and financial status of power grid, and the disadvantage is that statistical analysis cannot predict the future situation.

The Subject Library of Comprehensive Evaluation

Comprehensive evaluation techniques include expert evaluation method, analytic hierarchy process, entropy method, rough set theory, fuzzy comprehensive evaluation, data envelopment analysis, etc. It is mainly used in comprehensive performance management of power grid enterprises, power grid enterprise investment and asset management, and the main business management, whose advantage is that you can use the comprehensive evaluation to evaluate the performance of the grid companies and other aspects of evaluation, to adjust the current situation based on the evaluation, and the disadvantage is that the comprehensive evaluation is subjective sometimes.

The Subject Library of Prediction and Early Warning

Prediction and early warning includes time series algorithm and improvement, artificial neural network algorithm and improvement, support vector machine algorithm and improvement, genetic algorithm, particle swarm optimization, co-integration theory and gray prediction theory. It is mainly used in the operation forecast of power grid enterprises and to make qualitative and quantitative analysis on the future development trend, applied to the prediction of electric power and electricity, the forecast of construction cost of transmission and transformation projects, and the forecast of conductor icing, etc. Early warning analysis can be carried out by comparing the predicted value with the actual value, such as the mechanism of monitoring and early warning of the risk of electricity recovery, owning advantages that the current development is more perfect and plays an important role in the future planning of the grid company. However, its forecasting technology still exists areas that need to be improved, such as easy to produce local optimal solution and other issues.

The Subject Library of Visual Technology

Visual display technology is mainly applied to the panoramic display business, which realizes the visualization of business information and business process, through the visualization technologies such as large-screen, video and smart devices, so that the visualization technology is presented clearly and intuitively and even in an interactive manner. Its advantage is that the information collected by the power grid industry is displayed in an intuitive way and is clear and easy to understand. However, the results obtained by the visualization technology are mainly used for qualitative analysis, and cannot be used for quantitative analysis, where the data processing is not sufficient and the conclusions are not sufficiently substantiated. At present, the visualization technology can transform the display modes of static, two-dimensional plane and isolated data into dynamic, three-dimensional and continuous graphical display modes to enhance the control over the power grid. The main technologies include Visualization Toolkit (VTK), Java 3D, OpenGL and so on.

Conclusion

This paper mainly studies the operation monitoring data mining technology of Power Grid Company. Through the construction of data mining technology library, we solve the problem of data mining business operations monitoring. First of all, the article organizes the establishment of the company's operation monitoring index system from the five aspects of external environment

monitoring, key indicators monitoring, operation status monitoring, key work completion monitoring, operation indicators change and early warning situation monitoring. Then, from the aspect of business attribute, we analyze the needs of our company's operations monitoring business data mining; finally, to meet the needs of the operation of the monitoring business in the era of big data on the ground, in-depth study of data mining technology for corporate operations monitoring is used to build data mining technology library. The research of company data mining technology lays the foundation for the development of digital business of Power Grid Company, points out the direction for the construction of power grid enterprise operation and management, and improves the digitization level of power grid operation and management.

Acknowledgments

This work was supported by Natural Science Foundation of China (Project No. 71471059), and it was also supported by the science and technology project of State Grid Corporation of China, called "Research and Application of Data Mining Implementation Library on Company Operation Monitoring".

References

- [1] J.Q. Li, J.Z. Liu, C.L. Niu and L.Y. Zhang. *Proc. International Conference on Machine Learning and Cybernetics* (Guangzhou, China, August 18-21, 2005), Vol. 3(2005), p.1642.
- [2] J. Lou, H. Lu, J. Xu, and Z. Qu: A data-mining approach for wind turbine power generation performance monitoring based on power curve, Vol. 10 (2016) No.2, p.137.
- [3] K. Mollazade, M. Omid, F. A. Tab, Y. R. Kalaj and S. S.Mohtasebi: Data mining-based wavelength selection for monitoring quality of tomato fruit by backscattering and multispectral imaging, Vol. 18 (2015) No.4, p.880.
- [4] J. Melvin, T. James and S. Babu: Monitoring the best product from consumers opinion forums based on their budget using data mining technique, Vol. 45 (2017), P.71.
- [5] J. Nakash: Real Time Product Analysis using Data Mining, Vol. 4(2015) No.2.
- [6] Q. Song and H.Z. Chen: Research on the construction of Power Grid Corp operation monitoring (control) system, 2017, No.12, p.60.
- [7] Peng L, Xiaoping Z. Social Stratification and Cooperative Behavior in Spatial Prisoners' Dilemma Games [J]. PLOS ONE, 2015, 10(7):e0131005.
- [8] X.L. Li, X.L. Li, Z.G. Ni, and L. Feng: Analysis and Research on operation monitoring based on large data mining, Vol. 37(2016), No.03, p.38.
- [9] C.M. Sui: Design and implementation of power marketing analysis system based on Data Mining Technology(MS., University of Electronic Science and technology, China 2015).
- [10] M.J. Zhang: Research on visual management of power operation monitoring(MS., Southwestern University of Finance and Economics, China 2013).
- [11] Y. Yan, J. Wu, S. Wu and J. Zhang. *Proc. International Conference on Materials Engineering and Information Technology Applications* (September 2015)
- [12] C. Liu, D. Yang, Z. Zhao, C. Hu: Monitoring Index System Construction Method Research of Production and Operation for Power Group, Vol. 188(2015), p.22.