Computer-aided Construction of Big Data Mining Platform in Subway Power Supply Network Operation Monitoring

Lina Guo [1] and Na Suo [2]

[1]Department of Electrical Engineering, Zhengzhou Railway Vocational and Technical University, Zhengzhou, Henan 450000, China, guoln1981@163.com
[2]Department of Electrical Engineering, Zhengzhou Railway Vocational and Technical University, Zhengzhou, Henan 450000, China, Suona06@163.com

Corresponding author: Lina Guo, guoln1981@163.com

Abstract. With the development and progress of the social economy, the subway is playing an increasingly important role in urban development and citizens' daily life. The subway power supply system provides the required electrical energy. Once the power supply is interrupted, it will not only cause the paralysis of urban rail transit, but also it will also endanger the lives of passengers and cause property damage. With the continuous development of science and technology, big data mining platform technology is fully used in value utilization. Compared with traditional data analysis, the data mined is more effective and reasonable, although there are some ways to mine data and data processing. The same, but in the mining process, research and analysis should be based on actual valuable indicators, big data mining should be used as a platform for analysis, the mining technology should be prospected, and the application of big data mining in subway power supply network operation monitoring should be explored. In order to give full play to the role of prior guidance and correction during the monitoring of the subway power supply network operation, the article selects the subway power supply network business as the entry point for monitoring and analysis, and uses big data on the basis of rules and regulations, expert experience, and business common sense. Analysis tools, using trend fitting and building analysis models and other monitoring and analysis methods to explore the operation monitoring and analysis of subway power supply network services. Through business data verification, the goals of discovering business problems, warning business risks, predicting economic benefits, promoting business standardization, and improving efficiency and benefits can be achieved. It has certain reference significance for the operation monitoring and analysis of other subway-related businesses.

Keywords: Big data mining; Subway power supply network; Operation monitoring; Trend fitting
DOI: https://doi.org/10.14733/cadaps.2021.S4.84-94

1 INTRODUCTION

As society progresses, the scale of cities continues to expand, and population density increases rapidly, people's demand for subways continues to increase, and the requirements for subway operation monitoring and safety are also increasing. The power supply system is an important part of the subway. Without the reliable and safe power supply of the power supply system, the normal operation of the subway is impossible. The safe and reliable operation of the power supply system is directly related to the safe and stable operation of the subway. The operation and monitoring of the power supply system directly affects the continuity of power supply for users. Operational monitoring and safety are the characteristics of the system's long-term work, which runs through the entire life cycle of the target system from the conceptual stage to the end of its life. Therefore, it is an arduous system engineering to construct the power supply system into a system with extremely high operational monitoring and safety.

The era of big data has brought a large number of opportunities for enterprises to gain richer, deeper and more accurate insight into market behaviors, but traditional data processing methods mostly stay at the most basic level such as query, statistics and simple analysis. The expert system relies on the "bottleneck" problem of artificial knowledge acquisition, which is helpless in the face of increasing data [1] the huge amount of data and lack of knowledge is a common problem in many enterprises in the era of big data. By using the relevant technology for deep processing of large amounts of data, found that the implicit information and take advantage of, Sun guides the subway supply network enterprise to make the right decision, so that the role of metro power between data can give full play to the out [2]. Is the core of the theory of the large data analysis, data mining algorithms, all kinds of data mining algorithm based on different data types and formats can be more scientifically presented has the characteristics of the data itself. Data mining is excavated from a large number of data (including text) implicit, unknown, on the relationship between the decisions have potential value, patterns and trends, With the knowledge and rules used for decision support model is set up, providing a predictive decision support methods, tools and processes [3]. Luo can take the place of experts excavated from a large amount of data hidden in the knowledge, use of database storage function, and find out the potential relationship between data and the data itself inherent rule [4]. Fully tap the value of the metro power between data, in improving operation management level, improve the level of customer service and support future power grid development, and provide decision support government has made great progress. The basic task of the data mining technology is mainly manifested in the classification and regression, clustering, association rules, sequential patterns, deviation detection [5]. The metro power supply system equipment maintenance cost in the subway accounted for a large part of the operating costs, reduce the maintenance cost can be saving the cost of metro operation. As is known to all, shorten maintenance interval can improve system operation monitoring, at the same time, frequent maintenance activities can also lead to bigger expenses. If maintenance interval is too long, although can save cost for maintenance activities, metro power supply system failure risk, once the system failure, will cause is far greater than the cost of maintenance costs [6]. Therefore, average trouble-free operation time is obtained by monitoring analysis of operation system, make reasonable maintenance plan is one of the most possible cost saving measures [7]. In a metro power supply system, traction substation is the subway to city power grid, the major source of harmonic current, is also equipment for metro power supply system of ac/dc conversion. From the point of view of traction substation rectifier unit, at present has widely used 12 pulse rectifier unit or 24-pulse bridge rectifier, Ben [8] through the 12-pulse rectifier unit 7 working mode analysis, deduction and set up mathematical model of the fundamental wave and harmonic. Zhong [9] for excluding bridge between equilibrium reactor in parallel operation mode of the 24-pulse bridge rectifier for detailed analysis, and deduce the dc

output voltage and line current and bridge between the commutation overlap Angle calculation formula, simulation is conducted by using PSCAD software validation. Gao [10] using Matlabk to build 12 pulse rectifier circuit simulation model, Got the output of the rectifier voltage and current waveform, and the six-pulse rectifier circuit primary side current waveform and FFT (Fourier series) and 12 pulse rectifier circuit primary side phase current waveform and FFT (Fourier series) are compared. Kawagishi [11] using Matlab/Simulink simulation module of supply network ((Power System Blockset) establishment of 24-pulse rectifier unit model is simulated, and the dc rectifier unit side harmonic current and harmonic current of ac are analyzed. Compared with 12 pulse rectifier unit, Dc side current harmonic content is small, the ac current harmonic content is reduced greatly, reduce the harmonic pollution of communication power supply network, at the same time, the output of dc voltage pulse frequency doubled, improve power quality of metro power, improve the overall utilization of the transformer rectifier unit, the only drawback is the cost will be increased accordingly. Using Matlab simulation under the subway run more stray current and the rail potential distribution rule and influence factors, Therefore, the influence of rail potential should not be ignored in the establishment of the dc traction equivalent model. Since the power source modeling of subway and subway will make the node voltage equation nonlinear, the iterative method should be used to solve it [12, 13].

At present, the subway power supply network has generated a large amount of data during the operation period, and the amount of data is PB level. Due to the rapid development and wide application of technologies such as the Internet, mobile computing, and the Internet of Things, a large amount of data will be generated every day for analysis. And extracting valuable information requires big data thinking. Big data thinking is to treat data as a "resource of energy and value, and use data as the core to think about and solve problems, so that massive business data collide, discover its relevance, and realize the value-added of data through professional processing. Marketing business application system and subway power supply network is information collection system data, using big data analysis and mining and model design as extended exploration points to carry out subway power supply network business operation monitoring analysis exploration.

2 METRO POWER SUPPLY NETWORK OPERATION MONITORING ANALYSIS BASED ON BIG DATA MINING PLATFORM

2.1 Thinking of Operation Monitoring Analysis

The big data metro power supply network business operation monitoring analysis and exploration model described in this paper firstly explores the monitoring rules based on the business, and then further explores the monitoring rules and analysis models based on the big data. Figure 1 shows the analysis and exploration model of metro power supply network business operation monitoring based on big data.

Explore business-based monitoring rules by sorting out documented management materials related to systems, standards, processes, organizations and assessments, as well as unwritten expert experience, business common sense, etc., to find business criteria that can help find problems, and carry out after confirmation by business departments Relevant verification work; exploration of monitoring rules based on big data. Based on the exploration of business monitoring rules, starting from business data, through statistical analysis of massive business data, we can find criteria that can help find problems, and use relevant management standards as support , Or judge based on the statistical characteristics of massive business data, and carry out relevant verification work after confirmation by the business department; the analysis model exploration based on big data is based on the exploration of business-based monitoring rules and the exploration of monitoring rules based on big data, from Starting from business data, through the analysis and mining of massive business data, we can find models that can reflect business conditions, reveal regular characteristics, analyze root causes of risks, and evaluate trend impacts, and carry out relevant verification work after confirmation by business departments.

The above-mentioned three kinds of operation monitoring analysis and exploration ideas, their rule verification needs to be carried out based on actual business data, and investigation and tracing of the discovered problems are carried out to find the root cause. Root cause tracing needs to deeply dig and analyze the specific business, and communicate with the business department to confirm. If necessary, new operation monitoring analysis and exploration work can be initiated according to the work idea from the business, to improve the problem criterion and analysis model, and realize the monitoring analysis Integration.

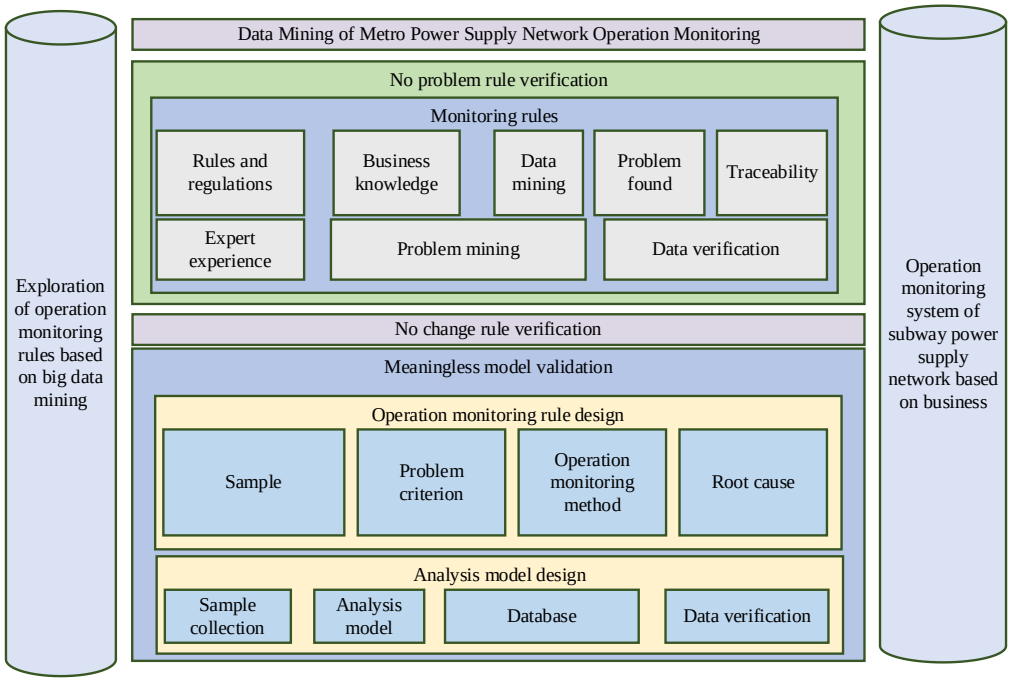

Figure 1: Big data-based metro power supply network business operation monitoring analysis and exploration model.

2.2 Monitoring and Analysis of Subway Power Supply Network Business Operations

Take the metering device change business as an example of the subway power supply network business operation monitoring analysis and exploration: first, sort out the systems, standards, and specifications for the metering device change business, so that the subsequent monitoring and analysis work can be followed; secondly, sort out the metering device changes Division of responsibilities of the business, find business weaknesses, unclear boundaries or undefined responsibilities; finally, sort out the business process of the metering device change, and horizontally connect the system, standards, norms and responsibilities of the metering device change business through the form of a process , Vertical series, and at the same time, through specific work content combing to find out the improvement points of monitoring and analysis business. The business improvement points are shown in Figure 2.

Aiming at the management improvement points in Figure 2, the monitoring rules are explored based on business system standards and specifications, based on big data statistical analysis, and based on big data modeling analysis. On the day of the meter replacement, the business personnel should confirm the indication of the old electric energy meter with the user, perform a special copy of the electric energy indication, and record the old meter to the SG186 system.

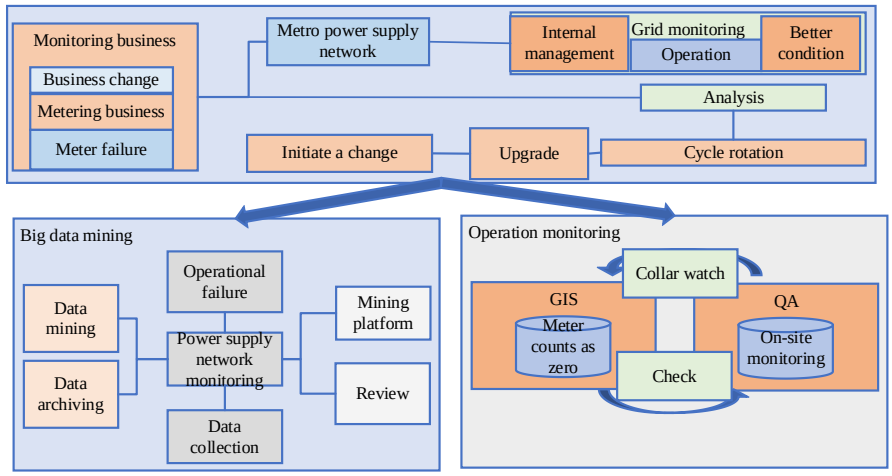

Figure 2: Sorting out business improvement points.

In order to avoid the inconsistency between manual meter reading and on-site readings, Article 28 of the "State Grid Corporation's Electricity Charges Reading and Collection Management Rules" stipulates that "for new electricity customers, customers who change electricity customers, and customers whose energy metering device parameters change, their business The calculation of the electricity bill for the first time after the process is processed shall be reviewed on a household-by-household basis." By comparing the special bargaining degree in the marketing SG186 system with the zero-point collection readings on the day of the meter change in the electricity information collection system, it is found that the special bargaining degree is less than the collected readings and the electricity bill is lost or the special bargaining degree is greater than the collected readings.

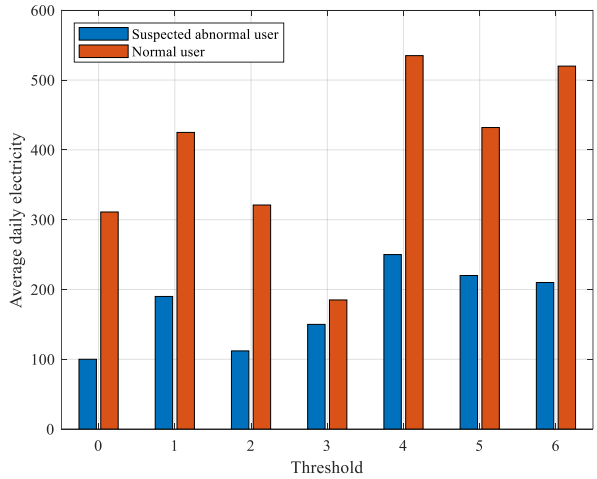

Figure 3: Schematic of dynamic threshold.

When the meter is changed, the meter stop degree-collection stop degree is less than 0, or (the meter stop degree-collection stop degree when the meter is changed) x magnification/daily average electric power full dynamic threshold (the units of the dynamic threshold can be obtained by the agile analysis tool, To suit the actual business status of the unit). The dynamic threshold is shown in Figure 3. Through monitoring, it can be found that the meter replacement personnel did

not confirm the number with the user, or to avoid disputes with the user and reduce the power consumption; there is the failure of the electric energy meter to see the number or the number is abnormal, and the meter replacement personnel did not Check the faulty equipment and estimate the bottom level.

3 DATA MINING AND REFINED OPERATION MONITORING ANALYSIS OF SUBWAY POWER SUPPLY NETWORK

In order to avoid the one-sided analysis conclusion caused by the independent monitoring of indicators, the data mining method is fully used to explore the potential correlation between various data. The index association network is constructed by considering the explicit algorithm association and the implicit service association among the indexes. The change of a certain index will affect other indexes associated with it. For example, the quantity of subway power supply is positively correlated with the income and market share of subway power supply, so the fluctuation of the quantity of subway power supply will lead to the change of other indexes. Therefore, if effective measures are taken to improve the electricity sold, other indicators will also be improved accordingly. Or when conducting abnormal dynamic analysis of indicators, these explicit relationships can help us quickly locate the causes of abnormal dynamic. In fact, in addition to these known explicit associations, there may be some unknown implicit business associations between different indicators. Effective use of this implicit relationship for the company to break professional barriers, improve performance indicators, control risks, will be able to play a multiplier effect. Based on the above analysis, we use big data mining technology to explore the implicit correlation of indicators.

With all the indicators in the hierarchical monitoring system as the analysis object and monthly as the cycle, the index values of the last 5 years were extracted from the business application system. Some missing and inaccurate data were removed through data cleaning, and a total of 346 indicators in 9 majors were finally determined to participate in the calculation. After the demonstration, the regression analysis method is adopted to establish the regression analysis model and calculate the correlation coefficient.

Regression analysis is an analytical method to determine the interdependent quantitative relationship between two or more variables. According to the number of independent variables, it can be divided into unitary regression analysis and multiple regression analysis. According to the relationship between independent variables and dependent variables, it can be divided into linear regression analysis and nonlinear regression analysis. Based on the observation data, it establishes the appropriate dependence relation between variables to analyze the inherent law of data, and can be used for forecasting and controlling. In order to calculate the correlation between two indexes, a unitary linear regression analysis is used. Its task is to find the linear regression equation describing the relationship between the two variables according to several observed values x and Y. In fact, the regression line is used to fit each observation point in the scatter diagram. The common method is the least square method, which minimizes the vertical distance between the line and each point. Even if the sum of squares of the difference between the measured y regression lines reaches the minimum, it is also called the residual sum of squares.

The correlation coefficient r in the linear regression equation is the magnitude of the linear correlation: the closer the absolute value is to 1, the better the linear correlation is, and the fitted line is more similar to the graph line obtained from the tracing point. Its calculation formula is as follows:

$$r = \frac{(x_1 - x)(t_1 - t) + (x_2 - x)(t_2 - t) + \ldots (x_n - x)(t_n - t)}{\sqrt{(x_1 - x)(t_1 - t)} + \sqrt{(x_2 - x)(t_2 - t)} + \ldots + \sqrt{(x_n - x)(t_n - t)}} \quad (1)$$

The regression model was used to carry out correlation analysis and the indexes were compared one by one. It is believed that correlation coefficient: indexes with absolute value of 0.8 and above

have linear correlation. Quantitative analysis and business association analysis are carried out for such indexes, focusing on cross-department implicit correlation indexes. After the index correlation is determined, the early warning and prevention network is established. Once the index changes in the associated network, other related indexes will automatically give early warning according to certain logical judgment, so as to achieve the effect of affecting the whole body, and change the traditional post-monitoring and analysis mode to the pre-risk warning and effective prevention and control mode. On the other hand, by controlling the process data (i.e., the indicator impact factor), the deviation is rectified in time to ensure that the key monitoring indicators reach the expected value.

Metro traction power supply system monitoring index is the traction power supply system operation monitoring on the basis of the history and future of evaluation and the basic starting point, can reflect the subway traction power supply system and equipment of the structure, characteristics, operation condition and its influence to the normal driving electric locomotive, and can be used as a measure of the relevant factors. According to the original data provided by the power supply department, due to the incomplete statistical data, the statistical data of the power supply system failure in 2003 and 2004 are shown in Table 1. Combined with the calculation method of the operation monitoring index of the subway power supply system, the operation monitoring index of the power supply system can be calculated.

Operational monitoring indicators	Total outage time (min)	Total metro delay time (min)	Number of Delayed Trains (columns)	Average subway delay time (min)	Number of failed power outages (times)	Number of power failure caused by external influence (times)
2003	135	13	3	7	13	2
2004	106	1	1	1	10	1

Table 1: Original data of a subway power supply failure.

$$\text{Average subway delay time} = \frac{\text{Subway delay time}}{\text{Number of delayed trains}} = 6.2\text{min} \quad (2)$$

$$\text{External influence blackout rate} = \frac{\text{External influences the number of blackouts}}{\text{Number of failures}} = 8.35\% \quad (3)$$

$$\text{Supply availability} = 1 - \frac{\text{Outage time}}{\text{Total time}} * 100\% = 99.98\% \quad (4)$$

From the above analysis results, it can be seen that the power supply availability in 2003 was 99.98, the average failure time was long, and the external influence was large. In 2004, the total time of power failure was significantly reduced, which indicated that the subway operation department had improved significantly in personnel management and maintenance organization. From the perspective of power supply quality index, the total delay time and the number of delayed subway trains in 2004 were significantly reduced compared with that in 2003, which reflects that the subway power supply operation department has enhanced its ability to organize work and reduce the scope of the fault in case of failure. In addition, the rate of power outage affected by external factors was also reduced, indicating that the main factors causing external failures, namely, local power supply failures, were reduced. This indicated that the coordination between the power supply department and the power department was strengthened, and the maintenance and repair efforts were also increased for the subway.

Therefore, through the analysis and calculation of the traction power supply system operation monitoring index, can be more clearly know that influence of the main causes of power system fault, guiding the subway power supply management personnel to enhance the management of the factors, also thought that maintenance staff to offer reference to the maintenance, planned maintenance, achieve the goal of saving the maintenance cost.

Monitoring based on the business rules for exploration and monitoring rules based on the large data, on the basis of exploration results, starting from the business, to think about different business subject as input, analysis model and converts business subject to demand analysis of data mining, based on a data cleaning and correction of sample data to construct analysis model, and constantly iterative optimization model, after confirmed by the business sector, access to huge amounts of business data for model validation, on the analysis of the model can reflect the characteristics of the business situation, reveal the law, analysis risk returning, evaluating trends, popularization and application. Analysis model construction and verification are shown in Figure 4.

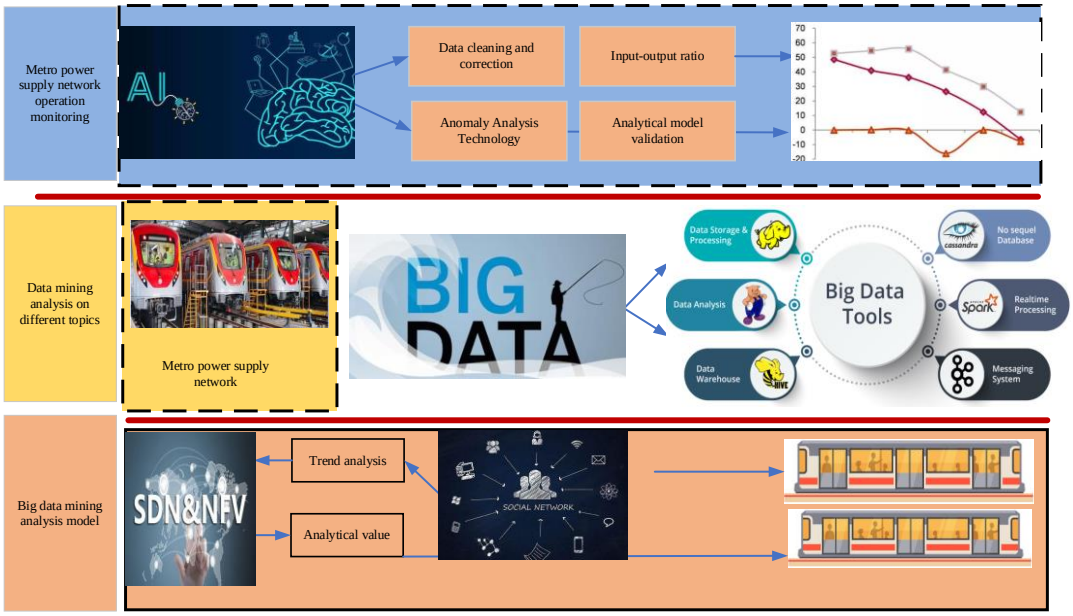

Figure 4: Construction and verification of operation monitoring analysis model.

4 ANALYSIS OF RESULTS

According to the relevant regulations of the State Grid Corporation of metering and connection, after the installation is completed, the person in charge of the installation shall power on to measure the voltage, current and phase sequence, and confirm that the electric energy meter is connected correctly. Through the comparison of the collected power before and after the user changed the meter, it was found that the user's power was not metered due to the wrong wiring or the zero-power problem caused by other reasons. By obtaining the business data of "the power of the electric energy meter before the replacement is greater than zero, and the power of the electric energy meter is equal to zero within a certain period of time after the replacement", the statistical analysis tool is used to fit the optimal value of the monitoring time. The monitoring time fitting diagram is shown in Figure 5.

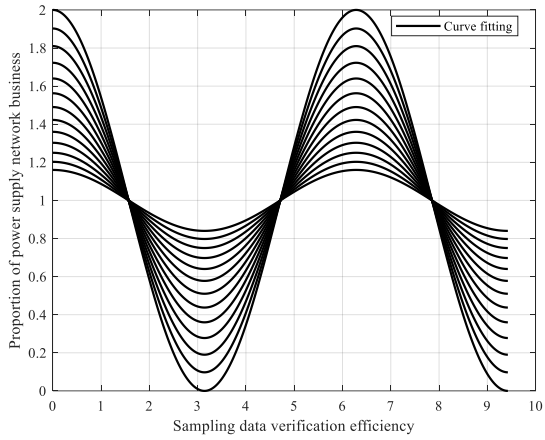

Figure 5: Monitoring duration fitting diagram.

The subway runs on lines with slope and curve length, that is, considering the additional resistance during operation. Based on the data of subway marshalling, subway weight, slope and curve length, based on the traction strategy, programming and simulation calculation on the Matlab platform, the VS (speed one mileage), as (acceleration-mileage) and Pt of the subway operation (Power-time) curve. It can be seen from Figure 6 that after the above-mentioned traction strategy is adopted, the maximum speed of subway operation is 88.03km/h. The average running speed is 86.545km/h, which makes the subway's traction performance fully utilized.

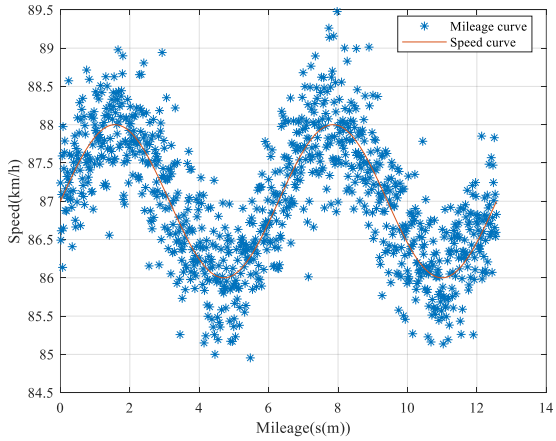

Figure 6: The v-s curve between subway stations considering the line conditions.

It can be seen from Figure 7 that the maximum traction acceleration of the subway is 0.737m/s2 when starting, and the maximum reverse acceleration is -1.252m/s^2 when braking, which is more in line with the requirements of subway safe operation and passenger comfort. It can be seen from Figure 8 that the traction energy consumption is 23.7609kwh, and the braking energy consumption is -19.7111kwh. It shows that even if a more optimized traction strategy is adopted, the energy consumption under the line condition is still higher than that under the straight-line condition. Therefore, it is verified by comparison that the subway traction load modeling considers

the simulation curve obtained by considering the line condition compared with the ideal straight line. The data is more accurate, and it also shows that the energy consumption of subway operations cannot be ignored considering the line conditions.

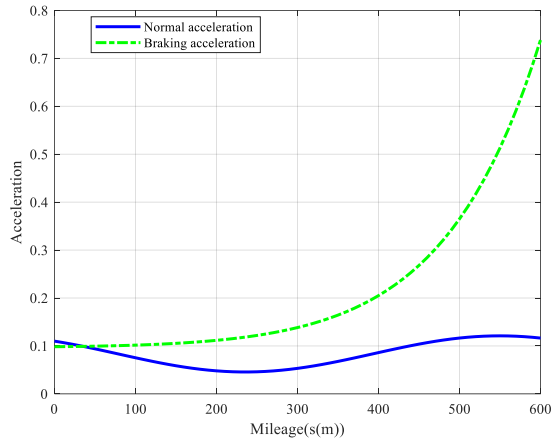

Figure 7: Consider the a-s curve between the subway stations under the line status.

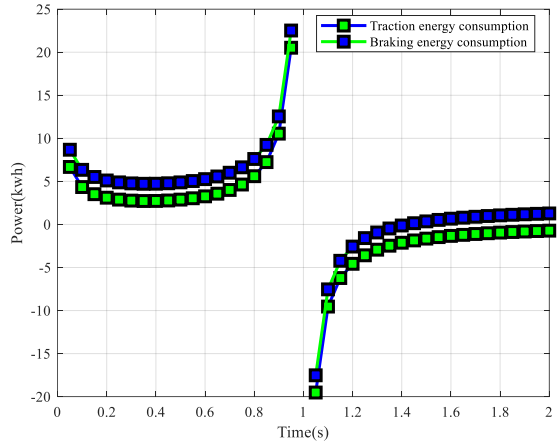

Figure 8: P-t curve between subway stations considering the line conditions.

5 CONCLUSION

The operation monitoring and analysis exploration model of subway power supply network business based on big data can construct practical and effective monitoring rules to carry out operation monitoring and analysis business. Using big data for data mining and modeling analysis can plan and guide future business, facilitate effective risk management and control and issue targeted countermeasures. This article verifies its role in discovering business problems, predicting business management and control risks, promoting business collaboration and efficiency, and improving efficiency and benefits through examples. It is an analysis and exploration program with the characteristics of promotion and replication, and can be widely used in other power-related business operation monitoring analysis in the work, it has become an effective support means for the construction of a comprehensive management and operation monitoring platform.

6 ACKNOWLEDGMENT

The 13th Five-Year Plan of Education Science of Henan Education Department in 2020 project name: Research on evaluation index of online teaching under the background of epidemic disease, project number: 2020YB0547.

Lina Guo, https://orcid.org/0000-0001-9030-0549
Na Suo, https://orcid.org/0000-0002-3561-6468

REFERENCES

[1] Bilal, M.; Oyedele, L.-O.; Qadir, J.: Big Data in the construction industry: A review of present status, opportunities, and future trends, Advanced Engineering Informatics, 30(3), 2016, 500-521. https://doi.org/10.1016/j.aei.2016.07.001

[2] Sun, Z.; Zhang, S.; Liu, N.: Application and analysis of hydrostatic level gauges in deformation monitoring of subway tunnels during operation, Modern Tunnelling Technology, 52(1), 2015, 203-208. https://doi.org/ 10.13807/j.cnki.mtt.2015.01.029

[3] Sutrisnowati, R.-A.; Bae, H.; Song, M.: Bayesian network construction from event log for lateness analysis in port logistics, Computers & Industrial Engineering, 89(11), 2015, 53-66. https://doi.org/10.1016/j.cie.2014.11.003

[4] Luo, Q.; Cao, Q.; Li, W.: Application of collaborative modelling method in Pile-Beam-Arc of Subway station, IOP Conference Series: Earth and Environmental ence, 558(3), 2020, 32049-32055. https://doi.org/ 10.1088/1755-1315/558/3/032049

[5] Hongliang, L.; Zhenhai, Z.: Ontology modeling and case database construction of subway emergency plan knowledge, IOP Conference Series: Earth and Environmental ence, 587(1), 2020, 12076-12083. https://doi.org/ 10.1088/1755-1315/587/1/012076

[6] Wang, W.; Meng, L.; Wu, L.: Research and Application of Data Mining Technology in Library Office Information Construction, Journal of Physics: Conference Series, 1550(3), 2020, 32001-32007. https://doi.org/ 10.1088/1742-6596/1550/3/032001

[7] Maalla, A.; Jia, C.; Ning, X.: The Application of Power Network Data Mining and Optimization processing based on Distribution network, IOP Conference Series: Earth and Environmental ence, 242(2), 2019, 22045-22050. https://doi.org/ 10.1088/1755-1315/242/2/022045

[8] Ben, F.: The Research of Automatic Deformation Monitoring in Subway Tunnel Construction Based on Georobot Technology, Geomatics & Spatial Information Technology, 102(2), 2015, 113-119. https://doi.org/ 10.1016/0889-5406(92)70022-3

[9] Zhong, G.-L.: Application of Slotting and Grading in the Construction of Deep Excavation of Subway, Value Engineering, 5(3), 2015, 3645-3647. https://doi.org/ 10.1109/IGARSS.2005.1526638

[10] Gao, T.; Wang, X.: The Construction of Computer-aided Translation Teaching Platform Based on Corpus, Journal of Physics: Conference Series, 1648(3), 2020, 32059-32064. https://doi.org/ 10.1088/1742-6596/1648/3/032059

[11] Kawagishi, M.; Kubo, T.; Sakamoto, R.: Automatic inference model construction for computer-aided diagnosis of lung nodule: Explanation adequacy, inference accuracy, and experts' knowledge, PLOS ONE, 13(11), 2018, e0207661-e20772. https://doi.org/ 10.1371/journal.pone.0207661

[12] He, X.; Xiang, S.; Luo, H.: Application Research on Computer Aided Simulation of Coastal Wetland Topography and Seawater Inundation Range, Journal of Physics: Conference Series, 1574(1), 2020, 12040-12045. https://doi.org/ 10.1088/1742-6596/1574/1/012040