Quality and Application Capability Analysis of Production Data ' Source ' Based on Ubiquitous Power Internet of Things

Hua Xun[12 a*], Zheng Kou[12b*], Jianchun Han[1c*]

[1] Inner Mongolia Electric Power (Group) Co., Ltd. Inner Mongolia Electric Power Research Institute Branch, Hohhot, 010020
[2] Inner Mongolia Autonomous Region High Voltage and Insulation Technology Key Laboratory, Hohhot, 010020

Abstract—Under the background of digital power grid, in order to deepen the digital application ability of Inner Mongolia power grid, the quality and application ability analysis of production data ' source ' based on pan is carried out to control risk, reduce crisis and realize the credible availability of data. In order to fully grasp the quality of the ' source ' of power grid production data used for interaction, the data of equipment account, technical parameters, maintenance record, test record and defect information in the production management information system are selected, and the quality analysis based on the integrity, correctness and consistency of the data used for interaction is carried out. It is found that there are problems of data ' source ' information error, missing and garbage data. In view of these problems, the decision of data management is put forward, and the suggestions for intelligent operation and inspection management and control of data are put forward, so as to improve the application ability of data and realize the value of data.

1. Preface

Under the background of digital power grid, in order to deepen the digital application ability of Inner Mongolia power grid production, it is necessary to analyze the quality and application ability of the production data ' source ', so as to prevent and control risks and reduce the crisis. In the process of data analysis, the data directly used and analyzed without testing the data quality, and the final result will be invalid or wrong. Therefore, it is necessary to fully grasp the quality of the data ' source ' based on the ubiquitous power Internet of Things, as well as the causes and manifestations of the root causes of poor data generation, and as an important support for data correction and transformation. At the same time, it has become the most basic link in the process of data analysis. The production management information system of Inner Mongolia Electric Power Company is a power big data platform with data integration and application as the core, which can realize the interaction and application of the whole data of production equipment. Under the pace of power grid information construction, the production management information system is growing, and the operation and maintenance data such as production equipment ledger, inspection, defect, test and maintenance are also increasing day by day. These data contain many rules and characteristics closely related to production safety and equipment management. The standardization, correctness, timeliness and integrity of data are becoming more and more important to the support and decision-making of production work. However, because the system production data belongs to second-hand data, which is re-entered, this will lead to many attributes of some historical data are not complete, and the lack of key production data. For example, the lack of manufacturers, technical parameters, equipment models, and operation time will affect the family defect analysis of equipment and the in-depth research and life-cycle analysis and evaluation of equipment failures. Therefore, the error and lack of data ' source ' have a very important impact on equipment evaluation and production operation decision-making, and data quality is ' determined by a series of parameters such as data consistency, correctness, and correlation '.In order to improve the utilization value of production data, systematic mining of data, scientific combing of data, and responsible application data, it is necessary to analyze the quality of production data. Put forward effective measures to achieve data reliability.

2. Production data ' source ' quality analysis

The quality analysis of the production ' source ' data based on the ubiquitous power Internet of things is a huge project. In order to achieve the effect of seeing the big from the small, the data of the production management information system since the company 's condition-based maintenance work is selected as an example. By checking the

[a]* 1169555952@qq.com, [b]* kouzhengdky@qq.com [c]* hang0507@163.com

equipment account, technical parameters, test reports, maintenance records, defect records and other information, the equipment data is analyzed from the perspective of integrity, correctness and consistency.

2.1 Data integrity analysis

According to the requirements of DL / T 1624-2016 ' Power System Plant and Main Equipment Naming Specification 'and the existing equipment management specifications of Inner Mongolia Electric Power Company , the integrity of primary equipment basic ledger information, technical parameter information, maintenance records, test reports and defect records was checked and analyzed.

2.1.1 Data integrity analysis of basic ledger

According to the requirements of the company 's " Grid Transmission and Transformation Primary Equipment Account Entry Specification ", more than 3 million key data of substation primary equipment are analyzed according to the unified rules of commonness of all equipment according to the nature of equipment parameters. It is found that there are common problems in the six technical parameters of specification model, manufacturer, manufacturer nature, factory date, commissioning date and pollution area level. The overall integrity rate of primary equipment account information is 86.0 %.

2.1.2 Technical parameter integrity analysis

According to the requirements of the company 's ' Condition-based maintenance substation primary equipment classification and parameter specification ' , about 1.18 million technical parameters of transformers and circuit breakers are selected as the sample data of this analysis, and the parameter items are non-empty as the inspection standard to analyze the integrity of the parameters. The average integrity rate of the technical parameters of these two types of equipment is 63.27 %, and all types of technical parameters of the equipment are more or less missing. The lack of technical parameters of these equipment will affect the results of accurate evaluation of the whole life cycle of the equipment, especially in the process of lean evaluation of the equipment, the analysis of the change rate of operation data is the key factor.

2.1.3 Defects, test, overhaul data integrity analysis

According to the requirements of the company 's " Guidelines for Condition-based Maintenance of Power Transmission and Transformation Equipment " , " Test Procedures for Condition-based Maintenance of Power Transmission and Transformation Equipment ", " Defect Management Standards for Power Transmission and Transformation Equipment " and " Requirements for Collection and Management of Equipment Status Data Information ", the integrity of defects, tests and maintenance data in production MIS is checked. The lack of one or more information in the statistical table is defined as incomplete record entry. From 2011 to 2018, a total of 208,000 maintenance records were recorded in the production management information system, with an average integrity rate of 66.0 %. The defects recorded 30,000 pieces of information, with an average integrity rate of 57.6 %. The test recorded 286,000 pieces of information, and the average integrity rate was 70.6 %. With the popularization of live detection technology, the amount of equipment test and maintenance data is increasing. Therefore, improving the integrity of data plays an important role in data analysis, safety production, decision-making and early warning.

2.2 Data correctness analysis

There are many contents for checking the correctness of equipment parameters. According to the key requirements of condition-based maintenance work, the key data in condition-based maintenance work is checked for correctness. The inspection scheme is divided into two parts : the basic account data of primary equipment and the technical parameters of equipment for correctness inspection, and the inspection rules are formulated.

2.2.1 Basic ledger data correctness

According to the requirements of the company 's ' production management information system power transmission and transformation equipment account entry specification ', more than 3 million key parameters of more than 210,000 substation primary equipment in the network are selected according to the data combing method to select the routine basic data for correctness analysis. The eight basic data of equipment voltage level, factory date, operation date, equipment name, equipment type, equipment model and equipment interval are mainly selected for data correctness inspection. These eight parameters are very important parameters in the process of condition evaluation and condition maintenance of equipment. Define simple inspection rules : the expression of the voltage level, the relationship between the date of delivery and the date of commissioning, the name of the device does not match the device type, interval attribution, device model description. Although only five simple rules, but according to the statistics found that the correct basic ledger information is as follows : the date of delivery and commissioning date is correct accounted for 63.8 %, equipment name and equipment type is consistent with the situation accounted for 67.2 %, and the correct interval accounted for 69.7 %.

2.2.2 Correctness of technical parameters

According to the ' condition maintenance substation primary equipment classification and parameter specification ' on the network equipment parameters fill in the case of a ' thick line ' check, the main transformer

157,000 technical parameters average correct rate of 79.4 %, of which 23 technical parameters correct rate of less than 60 %, mainly voltage regulation, circulation, wiring groups. The correct rate of technical parameters of 1020000 circuit breakers in the network is 82.6 %, and the correct rate of some technical parameters is less than 50 %, mainly including circuit breaker classification, mechanism model, closing resistance value and so on.

2.3 Consistency analysis of data

In the process of data interaction, data consistency can be basically achieved. However, in the process of data ' source ' control, there will be a certain degree of deviation in the collection and entry of ' source ' data, resulting in inconsistent data results. For example, consistency analysis can be performed by comparing ledger parameters with nameplate parameter information. This collection of 110kV and above transformer cooling system data and nameplate data, the consistency of the two were analyzed to reflect the quality of production management information system big data. Through the consistency analysis of the cooling system ledger information and nameplate data of 1254 transformers, it is found that only 72.6 % of the transformer cooling system is consistent with the nameplate data. Equipment ledger data inconsistent with the nameplate data, data information errors lead to low data quality, statistical data deviation.

3. Data quality status analysis

The production management information system is the main body of the production management of Inner Mongolia Electric Power Company. It takes the condition-based maintenance as the core, relies on the operation and maintenance management information system, and is based on the equipment ledger. The whole life tracking management of the equipment accurately controlled by the production process has become the core system that runs through and guides the production work. Through data analysis, it is found that the data quality in the current production management information system has the following problems :

3.1 Equipment information missing or error, poor data integrity

Through combing the equipment account data, it is found that the technical parameters of some equipment are not entered in accordance with the regulations, and there are missing parameter information and filling errors. In particular, the lack of some key data has a certain impact on the accurate evaluation of equipment and the decision-making evaluation of safe production. For example, the error or lack of transformer short-circuit resistance parameters has a significant impact on the accurate analysis of the actual situation of transformers ; the lack of equipment manufacturers and specifications in accordance with the ' power grid equipment family defect management approach ' to the same manufacturer with the same type of equipment family defect analysis, operation rules, maintenance recommendations and other related general information can not be shared, the network equipment can not be integrated early warning research, can not develop a strong guidance of the maintenance program ; the absence of the date of delivery and the date of commissioning will lead to the failure to correctly analyze the operating life of the equipment, while affecting the operation of the equipment, the formulation of maintenance plans, and the supervision of old equipment. Lack of pollution level will result in the inability to correctly analyze the frequent problems of equipment running in a substation, can not get similar equipment on the operation, maintenance and other aspects of technical assistance ; due to the lack of equipment base ledger information such as equipment specification model, location name and so on will lead to maintenance, defects, incomplete test records ; errors in data have a great impact on the statistical analysis of production data, which will lead to deviations in decision-making. Consolidating basic data plays a vital role in safety production.

3.2 Errors and duplicate garbage data cause system redundancy

The data in the production management information system shows a trend of power function growth. In practical applications, there is a lack of effective means to check the original data. Due to the unfamiliarity of the input personnel to the system, a large amount of garbage data is caused. For example, when checking the data information in the production system, it is found that there are a large number of errors and duplicate garbage data in various records. These data problems will lead to the system running more and more slowly. In the process of data reading, these garbage data will be read every time because they are not processed, which also causes great confusion to the data analysis and affects the decision analysis.

4. Solution strategy

Based on the rapid development of ubiquitous power Internet of Things demand for credible data spiral upward trend, high-quality data should be complete and accurate data, which can effectively improve the insight of decision-making and accurate prediction ability, and high-quality data is the most valuable data, in order to solve the above problems put forward the following strategy :

4.1 Establish data quality management system

The lack and error of data is actually the lack of standardized management of clear data quality, such as the lack of maintenance content, elimination content, test conclusion and other related content. The lack of equipment maintenance, defects, test records closed-loop management is not standardized, resulting in incomplete input records. Through the establishment of data quality

management and control system, the requirements of data management in the whole life cycle of data generation, storage and application are clarified, the responsibilities of data quality in different stages are clarified, and a unified management and control system is formed. For example, the equipment is returned to the factory for maintenance, the cooling system is reformed, the nameplate information is replaced, and the ledger should be updated in the production management information system in time ; when the equipment parameters are entered, they are compared with the nameplate information. Effective control measures can achieve regular evaluation and dynamic management of data quality. The data quality control model is shown in Figure 1.

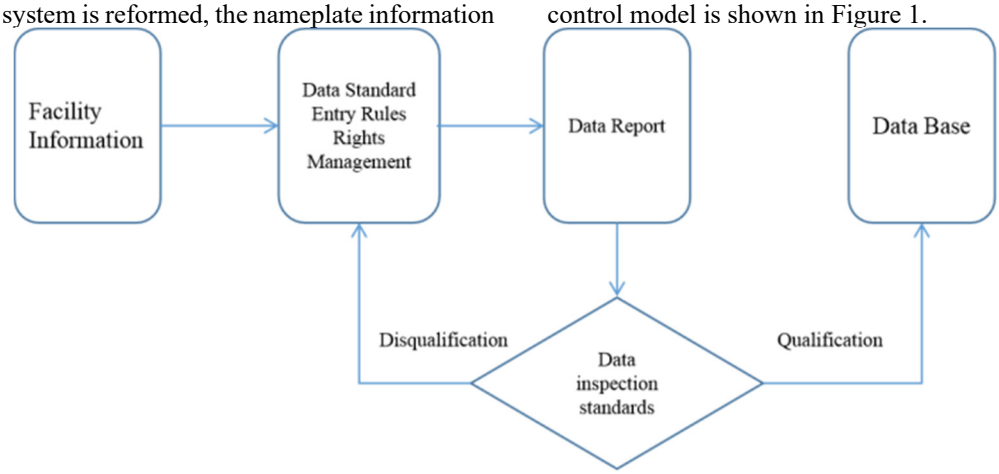

Fig.1Data quality control model

4.2 Conduct governance based on data 'source 'problem

For the missing and wrong information that already exists in the system, find the root cause of the problem through data governance, improve the missing data, correct the wrong information, delete the garbage data, continuously improve the accuracy and integrity of the data, eliminate the inconsistency between the information, reduce the number of redundant data, improve the quality of the source data, establish a standardized data application standard, improve the value of the data, maintain a high degree of consistency between the production data and the work target task, enhance the credibility of the data, and realize the effectiveness of the data life cycle. The data governance decision domain model is shown in Figure 2.

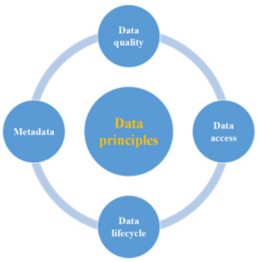

Fig.2 Data governance decision domain model

4.3 Enhance the application of data interaction

Through the data quality management system, the data can be interoperable and interoperable, and the data can be widely shared. The evaluation technology based on power grid operation, maintenance, test and basic data mapping is established to improve the data quality objectively and accurately. The data is used as an industry to deeply explore the value of data assets, and the data can be used as a valuable asset of the company in business, management and strategic decision-making to give full play to the value of data assets.

4.4 Application of rules to analyze data status

To carry out the analysis of the status of data quality, the rule weight configuration method is used to comprehensively evaluate the data inspection task, realize the quality inspection of business data, mark the quality problems of data, generate data analysis and analysis report, and use the measurement rules and inspection methods to carry out data quality inspection and evaluation of system tables and fields, find the problems of data, correct the problem data, and realize the comprehensive quality control of key business data. Using the clustering and association analysis method in statistics, the data automatic creation research and data association mining analysis based on Inner Mongolia power grid production management system are carried out. The built-in data quality check rules include specification check, duplicate data check, record missing check, reference integrity check, result set comparison, data data abnormal check, fluctuation check, timeliness check, business logic check, null value check, range check.

5. Summary

With the application of ubiquitous power Internet of Things in production, the coverage of production service business is expanding, which brings a series of accumulated data quality problems, such as missing

parameters, incomplete data, inaccurate data, data duplication, inconsistent data attributes, thus affecting the reliability of data information, resulting in unsatisfactory information construction and deviation of decision-making. The use and analysis of the production of big data must be based on high-quality data, which is a strong guarantee for the value of data generation. In the big data environment, the quality of data will directly affect the level of data value, thus affecting people 's analysis and decision-making. Therefore, in view of the current problems of data accumulation in Inner Mongolia Power Grid, the establishment of data quality management system and data governance scheme based on Inner Mongolia Power Grid decision-making can effectively standardize the management of data quality and realize the ' source ' data reliability of equipment life cycle.

Acknowledgments

This work was supported by the science and technology project of Inner Mongolia Electric Power Company : Research and Application of Test Capability and Detection Standard of Internet of Things Perception Layer Equipment Based on Digital Inner Mongolia Meng Electric Power Intelligent Operation and Inspection System (2022-15).

References

1. Meng Xiaofeng, kind. Big Data Management : Concepts, Technologies and Challenges [J].Computer Research and Development, 2013,50 (1) : 146-169 .
2. Cai Li, Zhu Yangyong.Big data quality [M].Shanghai Science and Technology Press, 2018.
3. Zhou Dong. Successful guarantee of data quality application system [J].China Information, 2006 (12) : 39-40.
4. National Energy Board. Power System Plant and Main Equipment Naming Specification DL / T 1624-2016 [Z]. China Electric Power Press, 2016.
5. Inner Mongolia Electric Power (Group) Co., Ltd. Production management information system transmission and transformation equipment account information entry specification Q / ND10501 02-2016 [Z].Hohhot , 2016.
6. Inner Mongolia Electric Power (Group) Co., Ltd. State maintenance substation primary equipment classification and parameter specification Q / ND10503 01-2016 [Z].Hohhot, 2016.
7. Inner Mongolia Electric Power (Group) Co., Ltd. Transmission and transformation equipment condition maintenance guidelines Q / ND 10503 03-01-20-2016 [Z].Hohhot , 2016.
8. Inner Mongolia Electric Power (Group) Co., Ltd. Test Procedure for Condition-based Maintenance of Power Transmission and Transformation Equipment Q / ND 10501 06-2018. China Electric Power Publishing House, 2018.
9. Inner Mongolia Electric Power (Group) Co., Ltd. Substation equipment defect management standard Q / ND 20502 04-2016 [Z].Hohhot , 2016.
10. Inner Mongolia Electric Power (Group) Co., Ltd. Management measures for familial defects of power grid equipment Q / ND 20502 0401-2016 [Z].Hohhot,2016.