Innovative Design and Intelligent Manufacturing
L.C. Jain et al. (Eds.)

doi:10.3233/FAIA241140

Big Data Analytics and Applications Based on Metered Asset Full Life Cycle Management Scenarios

Rui LI [a], Ruijia ZHANG [a] and Xiongfei SUN[b1]
[a] *Yinchuan University of Science and Technology, Yinchuan 750001, China*
[b] *Longshine Technology Group Co., LTD, Shanghai200051, China*

Abstract. With the in-depth development of measurement business informatization, how to establish the monitoring, analysis and management mechanism of measurement assets' life-cycle data has become a key problem to be solved. This paper aims to improve the effectiveness of measurement asset life-cycle management through big data analysis technology. Based on the architecture and functions of the measurement production scheduling platform (MDS), this paper uses ETL technology to achieve data extraction and conversion, and analyzes massive data through offline data mining and real-time data mining. The research results show that through clustering analysis, potential correlation analysis and trend prediction, the weak links in the life cycle of measured assets can be effectively identified, the management process can be optimized, and the management efficiency can be improved. The research in this paper is of great significance to the refined management of power metering assets, and provides theoretical and practical support for further improving management efficiency.

Keywords. MDS platform, big data, Full lifecycle management

1. Introduction

With the in-depth development of measurement business informatization, the purchase, arrival and verification of measurement assets are realized in the measurement production scheduling platform, the daily operation management is completed in the marketing business application system, and the electric energy information collection of smart meters is completed in the electric energy information collection system. Under the operation mode of large-scale centralized verification, the automatic production system of measurement, such as the automatic verification line of electric energy meters, intelligent storage system, AGV control system, is rapidly replacing the traditional manual verification method. The overall automation level of verification production facilities, such as "four lines and one warehouse" (including manual verification platform), has significantly improved. Based on the continuous deepening of business development, how to establish a data monitoring, analysis and management mechanism for the whole life cycle of measured assets, effectively monitor and record anomalies through data monitoring, anomaly analysis, rectification feedback, management

[1] Corresponding Author: Xiongfei Sun, Longshine Technology Group Co., LTD, E-mail:13522652309@163.com

evaluation and other means, micro analyze the causes of anomalies, and optimize the business situation and results in the management and control process, Improving the effectiveness of life cycle management of measured assets has become a key problem to be solved at present. With the development of big data analysis technology, it has become a reality to locate the weak links and businesses in the whole life cycle of measurement through data mining and analysis models; It has become practical and feasible to formulate targeted rectification measures and continuously improve the level of lean measurement management.

2. Research Methods

Asset full life cycle management is a kind of green management method that pursues low cost on the basis of management quality and efficiency, which focuses on the relationship between asset safety, efficiency and cost, and in today's reform situation, it is an effective way to improve the technology and equipment level of electric power equipment as well as the operational efficiency of assets, and one of the most important research topics to realize the transformation of China's electric power development mode! [5].

Wang, Z. et al. believed that in a highly competitive global economy, asset management is often one of the last choices to maximize cost savings, especially in many developing countries [6]. When designing and implementing projects, asset management of process industry must consider the commissioning, operation and end of life cycle of fixed assets. However, the current asset management model is inefficient in solving the whole life cycle cost and sustainable development. Then, the asset life cycle management model is proposed for the assets in the process industry. This model combines the concepts of general project management framework and system engineering with operational reliability to solve these inefficiencies. Lim, K. Y. h and others analyzed the application of ALCM in the field of smart grid, proposed a mutual management model, and briefly analyzed the eight indicators included in the complete asset management system [7]. These indicators are composed of basic expressions of predictable income, which helps to optimize the operation of assets in the smart grid. Finally, the management model based on life cycle asset management theory is simply compared with other models, and their advantages and disadvantages are obtained. Achouch, M. and others proposed a MIMS diagnostic model of asset management process based on the whole life cycle in view of the requirements of asset life cycle management for the whole process closed-loop, combined with the characteristics of power grid enterprise asset management, and from the management mode, information sharing, decision-making mechanism and methods, as well as systems and standards [8]. Popescu, i.s. and others tried to apply life-cycle cost management to the repair, failure, inspection and maintenance of overhead line equipment, and calculated various statistical data related to cost and reliability within the life cycle of overhead line equipment. Finally, it is proposed to analyze and evaluate the line asset management strategy in combination with the life-cycle cost [9]. Leng, J. and others proposed that the establishment of effective asset life cycle management is of great significance to improve economic efficiency and long-term development capacity. Life cycle management theory should be applied to large-scale asset management of public management departments, such as transportation infrastructure assets [10].

The error data, batch information, unqualified project data, marketing business system production system fault detection data, scrap fault detection data, etc. in the measurement center production scheduling platform (MDS) platform are extracted by timestamp, and data conversion and processing are realized by ETL data integration technology. The extracted data is analyzed through offline data mining and real-time data mining. Through algorithms such as mathematical statistics and artificial intelligence, the comprehensive quality evaluation based on clustering division, meter verification data, potential association analysis of abnormal scrapping, trend prediction, verification real-time alarm, etc. are studied for massive data.

3. Research methodology

3.1. Measurement production platform architecture

3.1.1. Metering production scheduling platform application architecture

The application architecture of the metering production scheduling platform includes six modules: scheduling and monitoring, production operation management, metering system management, platform auxiliary functions, platform support functions and production quality analysis functions. And through the interface to interact with marketing business and other systems [12].

3.1.2. Technical architecture of the metering production scheduling platform

Measurement production scheduling platform through a unified data interface to achieve data interaction with other business applications such as marketing applications. Measurement center production scheduling platform multi-layer architecture using component technology, interface control, business logic and data processing separation. Technically, it is divided into performance layer, application layer, platform layer (business platform, scheduling platform) and data layer, and interacts with the data center, business application system and four-line library.

3.2. The technical architecture of the analysis platform and the data extraction procedure

3.2.1. Big Data Analytics Platform Technical Architecture

Combined with the current MDS data volume and data characteristics, and taking into account the future development planning of measurement business and data value-added analysis needs, big data technology is used to build a data analysis platform. Its technical architecture is shown in Figure 1 below:

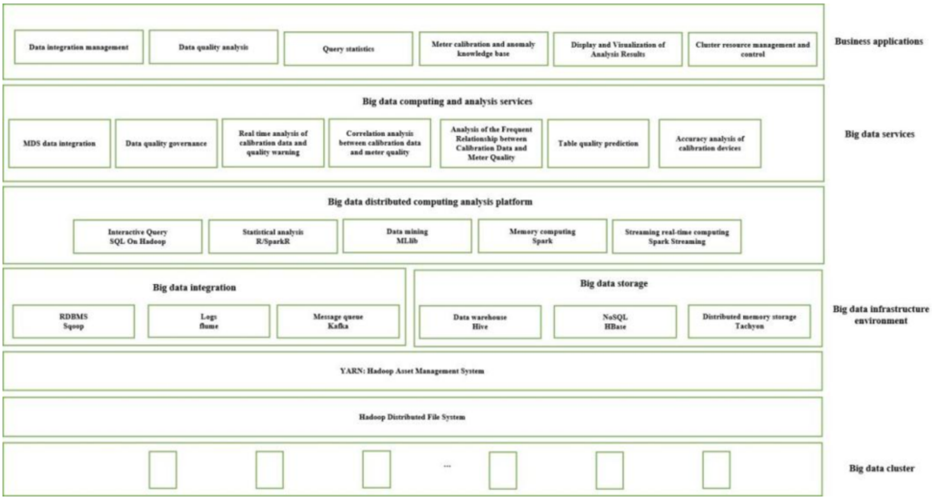

Figure 1. Diagram of the technical architecture of the big data analytics system.

3.2.2. Data extraction procedures

The big data analysis platform is based on the MDS system to analyze relevant quality data, and the extraction of these data is the premise of big data analysis.

The extracted data mainly includes error data from MDS (test conditions, average error, validation pipeline (validation device), test time, etc.), batch information (batch number, manufacturer, equipment specification, arrival time, batch quantity, etc.); Data of nonconforming products (information of verification instruments (asset number, specification, batch number, manufacturer, verification date, verification conclusion, specific verification items and conclusions, verification process); It also includes fault detection data from the marketing business system (commissioning time, disassembly time, fault occurrence time, on-site detection data, indoor verification data, fault description, fault type, fault instrument information (asset number, specification, batch number, manufacturer, etc.); Scrap failure detection data (commissioning time, disassembly time, scrap reason, scrap measurement information (asset number, specification, batch number, manufacturer, etc.), scrap equipment detection information (if any), etc.) [13].

In the process of measuring big data extraction and utilization, the data content, data format and data quality are different due to different data sources. The data source problem needs to be solved. Even data in different formats cannot be converted into data in a unified format, or the converted data cannot be used, and the data is missing [14]. Therefore, ETL data integration technology is used to continuously accumulate and extract data, that is, only the data accumulated in the tables to be extracted since the last extraction in the database is extracted. Through these technologies, data extraction, conversion and processing are realized.

(1) The time byte method is used to extract data. It is a way to capture data changes by snapshot comparison, adding a byte that can mark time on the data source, and when the data source needs to be updated in the system, the value of the time byte will be updated synchronously. When the data needs to be extracted, the time of the system itself is compared with the value of the time byte of the data source to decide which data to extract [15]. Support for the time itself and the time of change, their own changes to the

data do not need to manually synchronize. Some databases do not support the automatic update of the time byte, when the system updates the business data, you need to manually update the time byte.

(2) Data conversion and processing in ETL engine. The main methods are mapping, data filtering, data cleaning, data integration, data encryption and decryption through mapping tables.

By extracting data from different systems, formats and characteristics and realizing data integration through a unified data access platform, comprehensive basic data can be provided, which is an important prerequisite for big data analysis.

3.3. offline data mining and real-time data mining

3.3.1. Offline data mining analysis

Offline data mining and analysis refers to the process of importing the business data of MDS and marketing system into the big data environment, integrating and cleaning them, and then using mathematical statistics or artificial intelligence and other algorithms to analyze the massive data, trying to discover the hidden laws and potential relationships behind the data, so as to provide guidance for the business. In this scheme, the available big data analysis techniques and combinable scenarios include:

(1) Correlation analysis: correlation analysis is to analyze the correlation between two or more variable elements of the observation object to measure the closeness of the correlation between factors [16]. In the MDS system, there are many verification items for instruments, up to more than 100. At present, the verification results of various parameters in the system are managed discretely, and the correlation analysis of various data is not carried out. Based on big data, carry out association analysis on data, identify potential relationships between validation items, and accelerate validation process and quality judgment;

(2) Trend prediction: by collecting and sorting out the past data of observation objects, identify them, analyze the development process and rules of observation samples with time axis, and predict their development trend. In MDS system, measurement error evaluation and standardization analysis of calibration device can be considered;

(3) Potential correlation analysis: potential correlation analysis is mainly used to analyze the correlation between different components in the observation sample. For example, two or more events often occur between different observation samples. Analyzing the current relevance can help to judge the possible problems in the future according to the observation results. For example, in MDS data analysis, we can analyze the potential relationship between validation data, scrap detection data, scrap reasons and other data originally managed in different life cycle stages. According to the set mathematical statistical threshold, filter out the potential relationship with universality, which is helpful to find out the rule between the abnormal scrapping of instruments and the verification data, and then estimate the operation results of instruments.

(4) Comprehensive quality evaluation based on clustering division and feature analysis: for massive analysis samples whose relevant attributes have been quantified and whose number is large but in a chaotic state, clustering analysis method is used. It is not necessary to give classification standards in advance during the mining process, but from the sample data, analyze the similarity between different samples according to scientific data mining algorithms, The observation samples were automatically divided into several groups. The massive verification data in the MDS system are analyzed by

cluster division method, and the instruments with similar quality are divided according to the verification data, so as to realize the comprehensive evaluation of the quality of measuring instruments; On this basis, by analyzing the intra group characteristics of similar quality scales, we can identify the product commonalities between similar quality scales.

(5) Quality prediction: predicting the final category of samples based on existing data is a common scenario for big data mining algorithms. In the MDS system, the existing verification data cannot be effectively associated with the scrapping reason and operation quality of the instrument. Big data technology uses multiple regression and classification methods to mine and ultimately establish an effective quality prediction model based on massive meter verification data and operation quality data, so as to predict the final operation quality of meters according to the verification data. ·

3.3.2. Real-time computing and mining

Real time computing and mining technology is different from traditional OLAP and offline data analysis. It emphasizes the timeliness of data computing, analysis and knowledge mining. It is especially suitable for scenarios where data is continuously generated, real-time calculation and analysis are required, and analysis responses are made in a timely manner. Such as real-time monitoring of the operation status of key operating equipment.

In the MDS system, there may be tens of thousands or even hundreds of thousands of batch meters from one manufacturer. If this batch of instruments has passed the manual sampling inspection, but the error is too large or the hidden danger is too large in the online inspection process, it is likely that there is a problem with the whole batch of instruments. If real-time calculation and mining are adopted, common problems existing in the current batch table can be found through real-time calculation and cumulative analysis during pipeline inspection, and early warning can be given in time.

4. Analysis of results

4.1. Comprehensive quality evaluation based on clustered segmentation

Although all the instruments that passed the test are qualified, there are quality differences between them due to the size of the error and other reasons. The main purpose is to analyze the similarity of instrument quality based on the values of the test items, so as to classify and categorize the instrument quality from a macroscopic point of view [17].

(1) Data processing: the instrument verification data in the MDS system is used as the source of analysis data. For each equipment, the verification data of all items constitute its unique quality behavior data.

(2) Clustering: input all quality behavior data of the instrument into the MDS big data analysis platform; The big data analysis platform calculates the similarity between instruments according to the verification value; For the same validation project, if the values of two equipment are similar, it indicates that the quality of the equipment in the project is similar; Similarly, if the values of two equipment in each verification item are very similar, it means that the quality of the two equipment is very similar [18]. Similarly,

if the values of each item differ greatly, it means that the quality of the equipment differs greatly.

(3) In-depth analysis: on the basis of clustering, in-depth analysis of different groups of meters with good and poor quality, analysis of different production attributes such as manufacturer, production batch, meter type, etc., and in-depth mining of production information affecting the quality analysis of the meter will help to identify the deep-rooted factors affecting the quality of the meter, such as suppliers, brands of parts and components.

The original belongs to different life cycle stages of the management of measuring equipment, completely independent of the data grid together, the establishment of calibration data and instrument quality (such as early retirement) of the associated data matrix, so as to realize the different life cycle data through.

For the newly acquired data matrix, as shown in Figure 2, each row vector is regarded as an independent table behavior data. Use big data association analysis and frequent pattern recognition methods to conduct association pattern recognition on the behavior data of historical instruments. Filter and exclude the data mining results according to the set effective thresholds such as the confidence and support of the association pattern; The final output meets the potential association mode of threshold setting, which is of universal significance in data statistical analysis. For example, through association analysis, the following rules can be output: if a single-phase meter manufacturer is A1, the value of verification item 1 is a2, and the value of verification item 2 is A3, then the probability of failure of verification item A4 of the product may be B%, and the probability of early scrapping due to XX is c%;

Verification data				
Equipment number	Items 1	Items 2	...	Items n
001	0.025	...	...	...
002	-0.32	...	...	...
003	0.053	...	...	...
001	1.501	...	...	...
...	...	...	...	...

+

Qualitative data	
Equipment number	Scrapped in advance
001	yes
002	no
003	yes
001	no
...	...

Figure 2. Matrix of Linked Data.

The association rules based on potential correlation analysis, which are strictly based on mathematical statistics, have a solid mathematical theoretical foundation and are universally applicable. According to the results of the analysis, the values and occurrence probabilities of other meter behavioral data can be inferred from some meter calibration data [19]. It is also possible to analyze potential correlations between hundreds of calibration items individually, as a complete record of meter behavior.

4.2. Trend projections

Evaluation of measurement standards and equipment is carried out on the basis of the trend of measurement errors in the calibration devices. Timely detection and correction of possible measurement errors in calibration devices has improved the level of measurement work.

(1) Data organization: the same instrument for manual calibration data and assembly line measurement instrument calibration data for comparison, the formation of calibration equipment measurement error. Then the measurement error of the calibration equipment belongs to the data associated with the time, the formation of calibration equipment measurement error time series.

(2) Trend analysis: In general, the measurement of calibration equipment is not an instantaneous change in measurement deviation, but requires a relatively long process of deviation evolution. For this reason, the historical calibration performance and error evolution of the calibration device can be analyzed to find its out-of-tune pattern, so as to estimate in advance. After inputting the time-series data generated in the previous step into the big data analysis platform, trend analysis can be used to predict the possible error level of the measurement device in the next cycle or cycles, which can help to detect hidden problems in time.

4.3. Real-time alerts

According to the calibration data of the calibration pipeline, real-time monitoring and quality alarm for the current batch of the device to be tested. The experimental principle is shown in Figure 3.

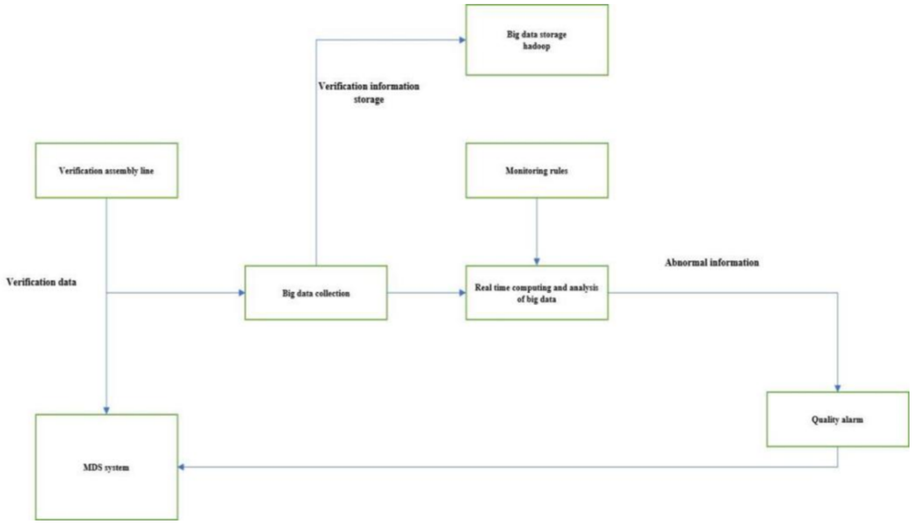

Figure 3. Schematic diagram of quality alarms.

The biggest difference between real-time verification alarm and traditional data query, statistical analysis and alarm after being put into storage is that its data processing is real-time: adopt technical means to make the verification pipeline data of MDS enter the big data system while writing into the database; The data generated by the continuous pipeline is also continuously input to the big data platform; The big data platform processes the input validation data in real time according to the preset business alarm rules, and gives real-time quality alerts to the validation batches that trigger the alarm threshold.

5. Conclusion

This paper first briefly introduces the measurement production scheduling platform (MDS), introduces the data extraction scheme and application architecture of the big data research platform, and the construction scheme of the big data platform of the integrated marketing business application system.

Finally, it focuses on the analysis of asset lifecycle management methods in big data measurement analysis. Combined with multiple application scenarios and demand analysis of big data measurement asset lifecycle management, data analysis methods such as clustering division, potential correlation analysis, trend prediction and other data analysis methods are introduced in detail and future application scenarios.

References

[1] Yu, W., Patros, P., Young, B., Klinac, E., & Walmsley, T. G. (2022). Energy digital twin technology for industrial energy management: Classification, challenges and future. Renewable and Sustainable Energy Reviews, 161, 112407.

[2] Yitmen, I., Alizadehsalehi, S., Akıner, İ., & Akıner, M. E. (2021). An adapted model of cognitive digital twins for building lifecycle management. Applied Sciences, 11(9), 4276.

[3] Succar, B., & Poirier, E. (2020). Lifecycle information transformation and exchange for delivering and managing digital and physical assets. Automation in Construction, 112, 103090.

[4] Kaewunruen, S., Sresakoolchai, J., & Zhou, Z. (2020). Sustainability-based lifecycle management for bridge infrastructure using 6D BIM. Sustainability, 12(6), 2436.

[5] Durana, P., Michalkova, L., Privara, A., Marousek, J., & Tumpach, M. (2021). Does the life cycle affect earnings management and bankruptcy?. Oeconomia Copernicana, 12(2), 425-461.

[6] Wang, Z., Akbar, M., & Akbar, A. (2020). The interplay between working capital management and a firm's financial performance across the corporate life cycle. Sustainability, 12(4), 1661.

[7] Lim, K. Y. H., Zheng, P., & Chen, C. H. (2020). A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives. Journal of Intelligent Manufacturing, 31(6), 1313-1337.

[8] Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On predictive maintenance in industry 4.0: Overview, models, and challenges. Applied Sciences, 12(16), 8081.

[9] Popescu, I. S., Hitaj, C., & Benetto, E. (2021). Measuring the sustainability of investment funds: A critical review of methods and frameworks in sustainable finance. Journal of Cleaner Production, 314, 128016.

[10] Leng, J., Ruan, G., Jiang, P., Xu, K., Liu, Q., Zhou, X., & Liu, C. (2020). Blockchain-empowered sustainable manufacturing and product lifecycle management in industry 4.0: A survey. Renewable and sustainable energy reviews, 132, 110112.

[11] Khzouz, M., Gkanas, E. I., Shao, J., Sher, F., Beherskyi, D., El-Kharouf, A., & Qubeissi, M. A. (2020). Life cycle costing analysis: Tools and applications for determining hydrogen production cost for fuel cell vehicle technology. Energies, 13(15), 3783.

[12] Leng, J., Ruan, G., Jiang, P., Xu, K., Liu, Q., Zhou, X., & Liu, C. (2020). Blockchain-empowered sustainable manufacturing and product lifecycle management in industry 4.0: A survey. Renewable and sustainable energy reviews, 132, 110112.

[13] Rigamonti, L., & Mancini, E. (2021). Life cycle assessment and circularity indicators. The International Journal of Life Cycle Assessment, 26, 1937-1942.

[14] Jbair, M., Ahmad, B., Maple, C., & Harrison, R. (2022). Threat modelling for industrial cyber physical systems in the era of smart manufacturing. Computers in Industry, 137, 103611.

[15] Ozturk, G. B. (2021). Digital twin research in the AECO-FM industry. Journal of Building Engineering, 40, 102730.
[16] Lu, Q., **e, X., Parlikad, A. K., Schooling, J. M., & Konstantinou, E. (2020). Moving from building information models to digital twins for operation and maintenance. Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction, 174(2), 46-56.
[17] Lu, Q., **e, X., Parlikad, A. K., Schooling, J. M., & Konstantinou, E. (2020). Moving from building information models to digital twins for operation and maintenance. Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction, 174(2), 46-56.
[18] Liu, Y., Zhang, Y., Ren, S., Yang, M., Wang, Y., & Huisingh, D. (2020). How can smart technologies contribute to sustainable product lifecycle management?. Journal of Cleaner Production, 249, 119423.
[19] Erkoyuncu, J. A., del Amo, I. F., Ariansyah, D., Bulka, D., & Roy, R. (2020). A design framework for adaptive digital twins. CIRP annals, 69(1), 145-148.