Smart Meter Evaluation System Based on Big Data

Yu Zhou1,*, Xuesong Shao1, Zhuowen Mu1, Qixin Cai1, Yue Li1, fan Gao1

1STATE GRID JIANGSU MARKETING SERVICE CENTER (METROLOGY CENTER), Jiangsu, 210019, Nanjing, China
*Corresponding author e-mail: zhouyu@jmsc.sgcc.com.cn

Abstract. Aiming at the problem of unstable operation of smart meters, this paper proposes a smart meter evaluation system based on big data. First introduce big data related technologies, such as data storage, analysis, mining, etc. Secondly, design the big data smart meter evaluation system model. Finally, use big data technology and clustering algorithm to realize the design of the big data smart meter operation evaluation system. The system can convert massive data from multiple systems into operational evaluation reports, which helps to reduce the waste of human and material resources.

Key words: Big data; Smart meter; Data mining; Distributed storage

Introduction
With the continuous development of the State Grid, electric meters have become more intelligent, and smart meters have been widely promoted and applied. On this basis, the coverage of its electrical information collection system is continuously expanding, which has increased the massive data of smart meters, making the operation of smart meters appear more complex, sudden, and multi-faceted. Therefore, in this case, how to ensure the stable operation of the electric meter is a problem worthy of discussion, and it is related to the interests of the people and social harmony. Chen Guang et al. conducted research on the big data analysis technology for smart meter quality in modern supply chains, which is helpful to provide reference for the stable operation of electricity meters; Qian Bin et al. proposed a meter measurement based on edge computing. In recent years, the country has increasingly higher requirements for the technical management of electric meters, and it is stipulated that the operation of electric meters can be stabilized by means of on-site inspection and rotation of electric meters. At present, electricity meter operation information is mainly distributed in the metering production scheduling platform, information collection system, and marketing business system. However, there is no metering business system that comprehensively detects and evaluates the operation status of smart meters in real time. Therefore, this paper designs a smart meter operation based on big data. The state evaluation system improves the timeliness and accuracy of the evaluation of the running state of the smart meter.

1. Basic method
1.1 Big data storage technology
Big data storage technology is diversified, with a variety of databases, and for different data structures, the most used databases are also different. When the data is clearly structured, you should choose...
relational database clusters and distributed databases; when the data is not structured and the data is large, you should choose NoSQL databases. And the distributed file system (HDFS) is suitable for large-scale, unstructured Archived data.

1.2 Big data analysis and processing technology
Big data analysis and processing technology is a method of real-time analysis, mining and batch calculation of massive data through specific analysis algorithms, combined with distributed computing programming models and parallel execution engines, and its purpose is to meet various performance analysis requirements. The big data analysis logic is: the Agent component can receive the static and dynamic resource information on the corresponding node at any time, and can also monitor the software module designated by the node. When the system receives one or more analysis tasks that conform to the distributed computing model, the distributed task scheduler cluster can schedule and allocate resources according to node information and resource organization strategies, and then use the parallel execution engine to make the task execution and results Converged reports are reliable and efficient.

1.3 Data mining technology
Data mining is a process of discovering valuable and novel data from massive databases. Under normal circumstances, data mining includes nine steps: one is to prepare relevant data, the other is to select data. The third is to perform data preprocessing. The fourth is to remove invalid data. The fifth is to determine the data mining goal. The sixth is to use corresponding mining algorithms. It is to mine data. The eight is to end data mining. The nine is to obtain evaluation reports. Data mining methods mainly include clustering, classification, correlation rules and regression, etc. Different methods have different algorithms and different application scenarios.

2. Smart meter evaluation system model based on big data
The model architecture of this paper is composed of five parts: data acquisition layer, data storage layer, data analysis and processing layer, evaluation display layer and user layer. The function of the data collection layer is to collect relevant data, then verify the data and clean up invalid data, etc. The main function of the data storage layer is to store various relevant data in this layer to provide data support for subsequent data processing and evaluation. The main function of the data calculation and analysis layer is to analyze and calculate the data. After the data is analyzed and processed, the data results are visualized through the knowledge display and sharing layer, and the results are sent to external systems to effectively integrate knowledge mining and business innovation.
3. Implementation of smart meter evaluation system based on big data

3.1 Implementation method

The system in this paper transmits data through a distributed system, and stores the data collected by the data layer in HDFS, then uses the MapReduce parallel computing method in the analysis system to preprocess and transform the data, and finally uses data mining tools to realize data mining and analysis. The specific implementation method includes three steps:

1. Use Bayesian algorithm to obtain on-site calibration, inspection records, operating years, etc.;
2. Obtain the evaluation result of the operation status of the smart meter through the clustering algorithm and regression algorithm, and after displaying and tuning the result, obtain the operation diagnosis report of the smart meter;
3. Finally, make a rotation plan and on-site calibration based on the diagnosis report. The realization process is shown in Figure 5.

Figure 1 structure of smart meter evaluation system based on big data
3.2 Big data mining platform and clustering algorithm
Finding the accumulation phenomenon in big data and describing it quantitatively is the focus of data mining. Cluster analysis is a method of classifying samples based on similarity, thereby reducing the difference of the same attribute in the same category, and increasing the difference of different attributes in different categories. In the actual application of smart meter evaluation, clustering algorithm, regression algorithm and Bayesian classification algorithm are used. Common algorithms in cluster analysis include Euclidean distance, cosine similarity and relative entropy. The formulas are as shown in formula (1), formula (2), and formula (3) respectively, and their purpose is to calculate similarity.
The k-means algorithm in the clustering algorithm is the most classic algorithm. The algorithm is simple and easy to understand, faster, and the calculation process does not require global variables or formulas, which makes MapReduce simpler, so this article chooses this algorithm to perform data mining. The principle of the k-means algorithm is to divide the data set into multiple samples N, divide its objects into k segmentation groups Si again, and then use an iterative method to optimize the exploration method, which can find the cluster center xj according to the attributes, and make the sum of the mean square errors of the segmentation group reach the minimum. The data center, the number of reducers, and the Euclidean distance are used as the data input of the k-means algorithm, and the files and center points of the data set machine classification are used as the data output.

The k-means algorithm is divided into three steps: One is to partition the data set, and send the center point file of the data block to the Map node. Different Map nodes have different tasks. After each task is executed, according to the principle of proximity allocates the data block to the central point, combines the central point and the corresponding data to form an intermediate result, and transmits the result to the Reduce node. The second is to classify the intermediate result, and then calculate the average of the intermediate results of the same cluster. Use the average value as the new center point. When the new center point remains stable, the operation has been successfully completed. The third is to observe the new center point to see if it is the same as the original center point. If it is inconsistent, it means The model does not converge and needs to be iterated again. Otherwise, the iteration ends.

3.3 System implementation
The metering online detection module in the electricity collection system is mainly responsible for detecting and analyzing the abnormality of the smart meter. The marketing business system provides business information such as on-site meter rotation and meter calibration, and the metering production scheduling platform is responsible for detecting the meter information. Combining the data information of these three external systems, the evaluation results of the operation status of the smart meter are obtained through the system of this paper, which provides a certain reference for the formulation of a rotation plan and on-site calibration, which is beneficial to ensure the normal operation of the smart meter.

4.Conclusion
This paper proposes and implements a smart meter evaluation system based on big data. Through the big data distributed storage and analysis calculation method, the data calculation is made more real-time, and the problems of large data volume and wide processor distribution are avoided. This article builds a big data mining platform to realize massive data mining and analysis, collect data from different external systems, and collect them in the same system, which is the big data intelligent evaluation system of this article, so as to realize the operation of smart meters. The analysis of the status and the more timely and accurate evaluation results are conducive to solving the problem of waste of manpower and material resources in the management of electric meters. Due to my limited knowledge and technology, there are deficiencies in the research, and I need to continuously improve the system and improve the application performance of the system.
References

[1] Chen Guang, Chen Shaobing, Li Zekun, etc.(2021). Research on big data analysis technology for smart meter quality in modern supply chain. Automation and Instrumentation, vol.5, pp: 180-183.

[2] Qian Bin, Cai Ziwen, Xiao Yong, et al.(2019). Data collaborative detection scheme of electric meter measurement system based on edge computing. China Electric Power, vol.52, no.11, pp:145-152.

[3] Wang Hao, Wei Tongjia, Zhang Wei, et al.(2019). Design of automatic test system for smart meters. Information Technology, vol.43, no.11, pp:145-150.

[4] Xue Bin, Zhang Xiangdong, Duan Li, et al.(2019). Real-time assessment method of universal smart meter status based on ubiquitous power Internet of Things. Electric Power Big Data, vol.22, no.11, pp: 38-43.

[5] Hao Fangzhou, Sun Qizhen, Shen Chao, et al.(2019). Analysis of user-side power consumption behavior of distribution network based on high-dimensional random matrix. Guangdong Electric Power, vol.32, no.11, pp: 111-119.

[6] Cui Shengsheng, Sun Jianfeng, Ma Bin, etc.(2020). Research on Smart Meter Data and Supervised Learning to Detect Non-technical Loss. Industrial Instrumentation and Automation, vol.01, pp: 122-126.

[7] Pan Lei, Yang Yan, Lian Hao, etc.(2020). Abnormal electricity consumption detection based on high-dimensional random matrix. Guangdong Electric Power, vol.32, no.11, pp: 111-119.

[8] Cheng Yingying, Du Jie, Zhou Quan, et al.(2020). Evaluation method of electric energy meter operating status based on random matrix theory and clustering algorithm. China Electric Power, vol.53, no.11, pp:116-125.

[9] Chen Liang, Lv Binbin, Fang Qinbin, et al.(2019). Analysis and application of station operating situation based on massive data of smart meters. Electric Power Big Data, vol.22, no.2, pp: 74-80.

[10] Jiao Zhen, Wang Zhaojun, Guo Hongxia, et al.(2020). Using measurement data and clustering algorithms to detect non-technical losses. Computing Technology and Automation, vol.39, no.4, pp: 61-67.