

Analysis of User Energy Consumption Patterns Based on Data Mining

Weitao Liu¹, Fuqing Wang¹, Hang Shi¹, Yan Zhang¹, Ruobo Chen¹

¹ Chengnan Power Supply Branch of State Grid Tianjin Electric Power Co., Ltd

Abstract. The energy use behavior analysis method can dig out the user's energy use behavior rules from the energy use big data, thereby improving the quality of the grid-side management service in the integrated energy system. Firstly, it summarizes the characteristics of the integrated energy system and constructs the integrated energy system service system; secondly, it summarizes the data-driven electricity consumption behavior analysis research model. Then, it elaborates on the collection and aggregation of electricity consumption information, and refined user classification. Next, the comprehensive application of energy consumption behavior analysis in load forecasting, demand response modeling and other typical scenarios is deeply analyzed. Finally, the challenges that may be encountered in further research are clarified and the follow-up work is prospected.

1 Introduction

Nowadays, with the rapid development of my country's social economy and the maturity of big data and related technologies, promoting the informatization and digital transformation of the power system has become an important task for the construction of the energy Internet. As the main component of the energy Internet, the integrated energy system is an important way for integrated energy service providers to meet the growth of users' energy demand and the diversification of energy use methods. Energy load data also occupies a considerable proportion, which reflects the real energy demand of users. Analysing user energy consumption based on data mining is beneficial to improve the level of user-side service management by integrated energy service providers, and also provides a basis for users to develop personalized energy service services.

At present, most domestic and foreign scholars' research on consumer energy behaviour analysis focuses on the research of consumer electricity behaviour. For example, literature [1] explores the feature extraction method for consumer behaviour analysis in the electricity sales market, using k-means, fuzzy Different clustering algorithms such as clustering and hierarchical clustering realize the pattern extraction of typical users' electricity consumption behaviour, and analyse the basic characteristics of different users' electricity consumption behaviour. Literature [2] puts forward a random matrix correlation algorithm suitable for multi-dimensional big data electricity consumption behaviour analysis based on a multiple power distribution big data platform, and discusses user electricity consumption behaviours for different object scenarios. For example, literature [4]

takes the multi-energy collaborative integrated energy system as the research object, constructs a comprehensive supply-side and demand-side optimization method for demand-side user behaviour analysis, and verifies the effectiveness and feasibility of the proposed method. Literature [5] elaborated on the characteristics of the user-centered integrated energy system, and conducted an in-depth analysis of the current situation of user-side integrated energy, laying a theoretical foundation for user-side management.

As power systems increasingly adopt digital information, decision support based on advanced data analysis is playing an increasingly important role in the operation and management of integrated energy systems. In addition, smart grids and smart meters have also been widely deployed in recent years. Real-time data rapid analysis and decision-making based on machine learning algorithms and big data computing platforms will become an important means for integrated energy service providers to maintain market competitiveness. Therefore, the analysis of user energy consumption patterns based on data mining can dig out user energy behaviour patterns from big data, thereby improving the quality of user-side management services in the integrated energy system, meeting the individual needs of users for energy use, and maximizing benefits.

2 Concepts and related theories

Integrated energy can realize the conversion between different forms of energy. For example, it can convert surplus energy that is not easy to store into other forms of energy that are easy to store, so it can improve the utilization and consumption rate of renewable energy,

* Corresponding author: mayuzebj@163.com

and can meet a wider range of Therefore, it can fundamentally adjust the energy structure and realize the sustainable development of energy. In addition, due to the close coupling relationship between various energy systems, when one of the systems fails or other unexpected situations, the other energy systems obtain corresponding information and use the energy conversion between supplies to make up for the supply gap, making the entire integration The energy system can still operate safely and reliably, ensuring the energy supply to the load demand, and providing a richer and more flexible means for the energy system in emergency coordination and control.

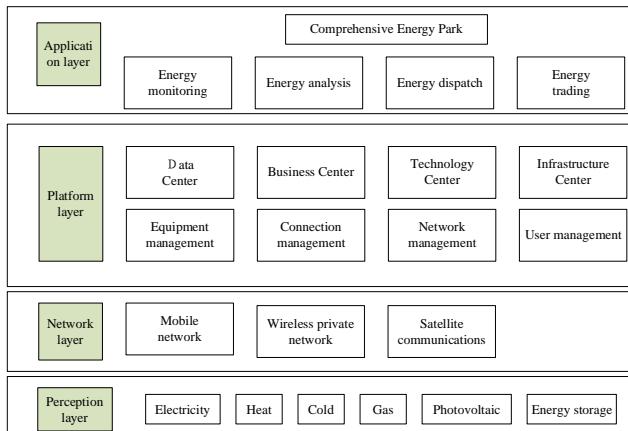


Figure 1 Overall construction of the park's integrated energy service system

3 User energy behaviour analysis

3.1. Analysis of influencing factors of consumer energy behaviour

The influencing factors of users' energy consumption behaviour can be roughly divided into four categories: user, system, environment, and policy. Each category of factors can be further refined, as shown in Figure 2 below.

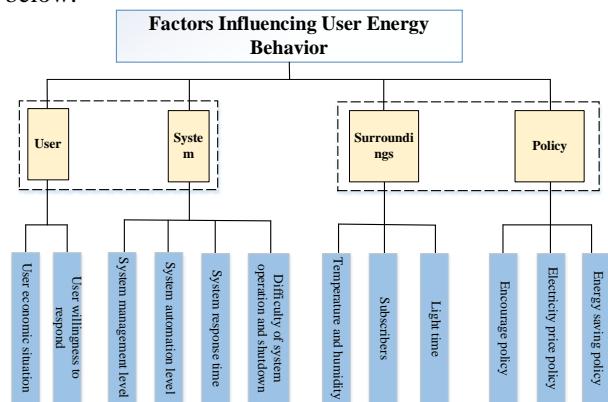


Figure 2 Influencing factors of user energy consumption behavior

3.2. Key technology

3.2.1. Energy consumption information collection and aggregation

Intelligent measurement terminal equipment data is flexibly accessed in real time, and data is interconnected and interoperable, which provides equipment foundation and technical support for user energy use behaviour data collection. At present, the energy consumption information collection system provides the main data source, and can perform data processing, analysis and application at the network edge layer, which is an important link in edge computing [7]. Load energy detailed monitoring can be divided into two types: intrusive and non-intrusive [8], among which: intrusive residential load monitoring (ILM) equips each electrical appliance in the total load with a sensor with digital communication function, and then The cost of collecting and sending energy consumption information on the local area network is relatively high; non-intrusive load monitoring (NILM) is equivalent to configuring a smart meter with additional functions. Compared with ILM, NILM adopts a decomposition algorithm to provide online feedback of sub-item electricity consumption information. It has the advantages of economy and simplicity, and is easy to promote on a large scale. It provides energy usage information with good availability and high value density, which is an analysis of energy usage behaviour. Smart power services with advanced functions such as energy efficiency analysis and demand-side management provide quantitative support, and have the potential to develop into the next-generation AMI core technology. The architecture of NILM-based power consumption information collection and analysis is shown in Figure 3.

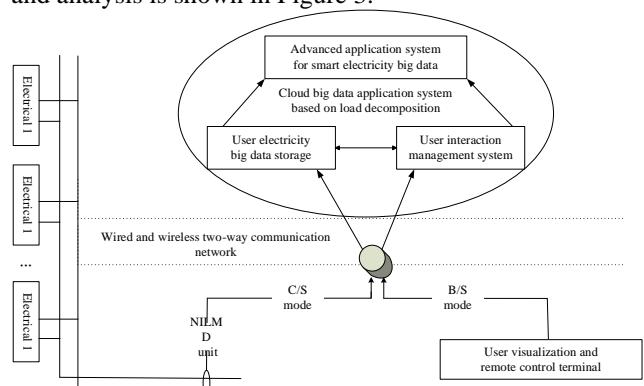


Figure 3 Power consumption information acquisition and analysis architecture based on NILM

3.2.2. Energy consumption information collection and aggregation

The energy use curve is essentially a superposition of multiple energy use behaviours, and the refined classification of user groups helps to identify users' basic energy use patterns, dynamic characteristics, uncertainty and other indicators. In terms of user group classification, most studies are conducted based on user social attributes or energy consumption data. Traditional power user classification methods are usually carried out in accordance with empirical rules. Users are classified according to user attributes such as industry, field, household population, etc., to establish a "user-label" network and construct a relationship weight model

classification [9]. This traditional classification method is simple and fast, but is greatly affected by empirical rules, and the classification results are not accurate enough. Therefore, in recent years, more and more researches have focused on the use of clustering methods in data mining technology to achieve more refined power user classification. The data-based classification model uses big data on the power user side, and based on the selection of features and weights, clustering methods are used to search for similarity of samples and classify them according to the characteristics of electricity consumption, so as to realize the pattern based on electricity consumption. The fine classification. The common way is to calculate the similarity according to the load curve or load characteristic index, and use regression methods, clustering algorithms, fuzzy algorithms and other methods to classify, including k-means [10], K-Medoids [11] based on partitioning clustering methods, COBWEB, self-organizing neural network [12] and other model-based methods. Different user classification methods have their own characteristics, and there is no absolute optimality. K-means and FCM are widely used because of their simple principles, easy implementation and high efficiency. However, traditional clustering algorithms tend to show problems such as unstable clustering results, slow speed, poor results, and excessive memory consumption when targeting high-dimensional, massive, cluster-shaped load curves, in practical applications, the clustering results are affected by multiple links and multiple factors in the extraction process. It is necessary to consider specific data types to determine a suitable user classification method. On the basis of proper classification, the energy consumption behaviour analysis of user groups can bring benefits to power companies.

4 Comprehensive application

4.1. Load forecasting

The traditional load forecasting method is not thorough enough to analyse the influencing factors, and it is easy to cause the consequence of lower forecast accuracy in some periods. In addition, as the scale of data collected by the ubiquitous Internet of Things continues to grow, traditional forecasting methods have defects such as too long calculation time and poor processing performance. There is an urgent need to adopt efficient and accurate forecasting methods for massive data for energy management and energy saving. Provide guarantee for emission reduction.

On the one hand, electricity consumption behaviour analysis has realized the refined classification of users and the identification of related factors. On this basis, individualized research can be carried out on different user groups, and the strong related factors that affect the electricity consumption patterns of various users can be explored, and respectively Establish a load forecasting model, and then aggregate the forecast results of various users to obtain the system-level load [13-14]. On the other hand, for the processing and calculation of massive

data load forecasting, big data processing frameworks can be used for distributed computing processing, including batch processing architectures such as MapReduce and stream processing architectures such as SparkStreaming, Storm, Puma, and S4, combined with advanced load Forecasting algorithm to achieve more efficient and accurate load forecasting.

4.2. Demand response

The response strategy of users in response to demand is not only affected by electricity prices and incentive policies, but also by user types and preferences, weather, social and economic development and other factors [15], while traditional demand response models such as elastic matrix and user psychology Methods such as learning models do not comprehensively consider multiple influencing factors, resulting in insufficient response models. Therefore, a demand response modeling method based on electricity consumption behaviour analysis can be used to analyse the influence of multiple factors on the load transfer rate for different user groups, so as to establish a multi-factor load transfer correlation model for different types of users to explore Demand response mechanism [16-17]. The load dynamic response model can also be established through real-time data analysis, and real-time electricity prices can be introduced for parameter identification and correction.

The regression function constructs a response model or establishes a response model with different time constants and delays. The research framework of this kind of demand response real-time modeling is shown in the following Figure 4.

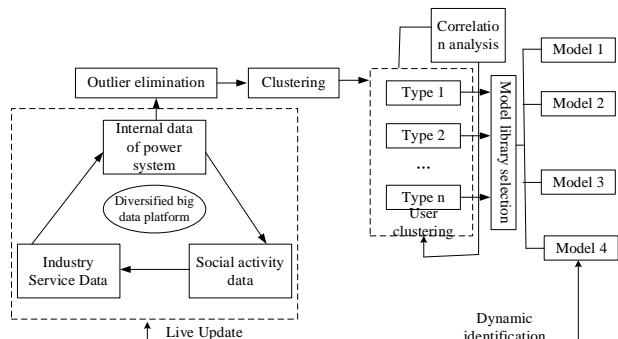


Figure 4 Research framework of real-time demand side response modeling

5 Comprehensive application

Large amounts of user-side data are introduced through big data, which contains a wealth of internal laws and derivative information about user energy use behaviour. This article first summarizes the characteristics of the integrated energy system and builds an integrated energy system service system; secondly, on the one hand, it summarizes the data-driven energy consumption behaviour analysis research model, and on the other hand, it elaborates on the collection and aggregation of electricity consumption information, and the fine-grained users. The key technologies of chemical classification

and other links; again, the comprehensive application of energy consumption behaviour analysis in several typical scenarios such as load forecasting and demand response modeling is deeply analysed.

With the continuous increase of user-side data, further research is to face a large amount of data, how to effectively and rationally use data, how to use big data, artificial intelligence and other technologies to promote the parallelization and distributed processing of user energy behaviour, and continuous improvement, Improve the application of big data analysis algorithms involved in the research and development of integrated energy systems, the use of data analysis technology, data management technology, data processing technology, and data display technology in the integrated energy system.

Acknowledgments

This paper is supported by "Science and technology project of State Grid Tianjin electric power company (KJ20-1-48)".

References

1. Zhao Yang, Hu Shiyao, Yang Shuqiang, et al. Analysis of data-driven based users' electricity consumption behaviour in retail market [J].Power Demand Side Management, 2020, 22 (04) : 45-50.
2. Hao Ran, Ai Qian, Xiao Fei. Architecture based on multivariate big data platform for analysing electricity consumption behaviour [J].Electric Power Automation Equipment, 2017, 37 (8) : 20-27.
3. Gong Gangjun, Chen Zhimin, Lu Jun, et al. Clustering Optimization Strategy for Electricity Consumption Behaviour Analysis in Smart Grid [J].Automation of Electric Power Systems, 2018, 42 (2) : 58-63.
4. Wang Qixin, Liu Dichen, Wu Jun, et al. Comprehensive optimization including user behaviour analysis for supply and demand sides of IES-MEC [J].Power Automation Equipment , 2017, 37 (6) : 179-185.
5. Jin Wende, Jiang Yibao, Ding Yi. Study on Critical Issues Correlated to Optimization and Management in Customer-Centered Integrated Energy System and the Prospect [J].Zhejiang Electric Power, 2016, 35 (10) : 73-80.
6. Ma Tiannan, Wang Chao, Peng Lilin, et al. Predictive analysis of energy consumption behaviour of integrated energy system users under multi-heterogeneous big data[J].Power System Analysis and Research, 2018, 46 (10) : 86-95.
7. Liu Sifang, Deng Chunyu, Zhang Guobin, et al . Research on collaborative architecture for edge computing of residential intelligent usage of electricity[J] . Electric Power Construction , 2018, 39(11) : 69-77(in Chinese)
8. Henriet, Simon & Simsekli, Umut & Fuentes, Benoit & Richard, Gaël . A Generative Model for Non-Intrusive Load Monitoring in Commercial Buildings[J] . arXiv e-prints , 2018 , arXiv : 1803 . 00515 .
9. Xu Changqing , Zhao Huadong , Song Xiaohui . Research on method of power user group identification and analysis based on large data[J] . Journal of Zhengzhou University , 2016, 48(3): 113-117(in Chinese)
10. Chévez , Pedro , Barbero D , Martini I , et al . Application of the K-means clustering method for the detection and analysis of areas of homogeneous residential electricity consumption at the Great La Plata region , Buenos Aires , Argentina[J] . Sustainable Cities and Society , 2017(32): 115-129 .
11. Liu Yang, Xu Lixiong . High-performance back propagation neural network algorithm for classification of mass load data[J] . Automation of Electric Power Systems , 2018 , 42(21) : 131-140(in Chinese) .
12. Oprea S , Bâra A . Electricity load profile calculation using self-organizing maps[C]//20th International Conference on System Theory , Control and Computing (ICSTCC) . Sinaia: ICSTCC, 2016: 860-865 .
13. Quilumba F L, Lee W J, Huang H, et al . Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities[J] . IEEE Transactions on Smart Grid, 2015, 6(2): 911-918 .
14. Zhao Teng , Wang Lintong , Zhang Yan , et al . Relation factor identification of electricity consumption behavior of users and electricity demand forecasting based on mutual information and random forests[J] . Proceedings of the CSEE, 2016, 36(3): 604-614(in Chinese) .
15. Tongzhi L I . Technical implications and development trends of flexible and interactive utilization of intelligent power[J] . Automation of Electric Power Systems, 2012, 36(2): 11-17 .
16. Yang Xuying, Zhou Ming, Li Gengyin . Survey on demand response mechanism and modeling in smart grid[J] . Power System Technology , 2016, 40(1): 220-226(in Chinese) .

17. Cai Long, Gu Jie, Jin Zhijian . Study on factor identification and feature extraction of residential demand response behavior[J] . Power System Technology2017, 41(7): 346-353(in Chinese) .