Energy Analysis of Integrated Energy Based on Cloud Computing

Du Yan¹, Wang Zhe-long¹, Guo Liang², Zhang Zhi¹, Chen Zhi-ru¹, Zhu hongxia¹ and Yangjie¹

¹State Grid ShanDong Electric Power Research Institute
²State Grid ShanDong Electric Power Company

461849930@qq.com

Abstract. This paper uses big data technology analysis means to abstractly identifies the common characteristics of various energy-using behaviors and analyze their behaviors; establishes user energy behavior analysis based on “multi-meter unification” information acquisition, and analyses and predicts the abnormal energy-using behavior of users according to the analysis results. So then provides a powerful means of monitoring and early warning for energy-using, increasing the level of user energy-using management, and maintaining normal order of user energy-using.

1. Introduction
The transition of energy companies from providing a single energy product to providing integrated energy services has become a trend with the support of big data and the Internet of Things. With the continuous advancement of “multi-meter unification” information acquisition construction, more and more places have been able to realize the joint collection data of electric meters, water meters, gas meters, and heat meters through the collection platform. The collection platform has accumulated a large number of user energy-using data about electricity, water, gas and heat. How to use these massive data to analyze the user's energy-using behavior, and analyses and predicts the abnormal energy-using behavior of users according to the analysis results is one of the problems that need to be solved [1-3].

2. Analysis of user energy use behavior

2.1. Analysis Framework of User Energy-using Behavior
Based on the massive user energy data of the “multi-meter unification” information acquisition system, the K-MEANS algorithm and the entropy weight method are applied. The user can be distinguished by combining the energy characteristics which conclude the energy consumption rate, load rate, trough coefficient and flat quantity of electricity, water, gas and heat. For example, the resident users can be divided into five categories: small business users, vacant houses, elderly families, working people, elderly people plus working people. In-depth analysis of each category of users, combined with the existing energy business, to provide the necessary data support for business development, improve the refinement level of energy business development [4].

The user energy behavior analysis based on “multi-meter unification” information acquisition is mainly composed of four parts: user energy data collection, cloud data processing, data mining and
system application analysis. The user energy data collection mainly completes the collection of user's electricity meter, water meter, gas meter, and heat meter. These data from different devices and different locations are classified into the cloud data processing module, and the parallel storage, conversion and loading are adopted, combined with K-MEANS model and entropy weight algorithm, the energy classification analysis of user behavior is realized efficiently, which lays a foundation for apply of users energy behavior analysis [5-6]. The framework of user energy behaviour analysis system is shown in figure 1.

![Figure 1. Framework of user energy behavior analysis system](image1)

2.2. User energy behavior analysis process

In order to better reflect the predictability of user's energy behavior, the potential characteristics of different user's energy use behavior are identified by collecting and classifying the information of user's energy use behavior. The data of user's energy use are analyzed to obtain the classification of user's energy use behavior. The user energy behavior analysis process is shown in figure 2.

![Figure 2. User energy behavior analysis process](image2)
2.2.1. Obtain data such as meters code and load curve: The user energy data is obtained through the "multi-meter unification" information acquisition system, such as the daily frozen electric energy and the daily power curve data of the low-voltage residents' measuring points, daily frozen heat energy, daily frozen gas energy and other basic energy-using information (all of the above data come from smart energy meter, smart water meter, smart gas meter, intelligent heat meter). At the same time, the real-time data of the energy meters in different time periods are obtained periodically to support the analysis of energy consumption characteristics in different periods.

2.2.2. Data screening: The abnormal checking of the basic energy-using data. If the meter goes backward, flies away, etc., the abnormal acquisition data will be removed from the analysis data set, Which will lay the foundation for accurate energy use behavior analysis.

2.2.3. Calculating energy consumption characteristics: four energy consumption characteristics for each user are calculated. The four energy consumption characteristics are peak-time energy consumption rate, load rate, Valley energy coefficient and flat time consumption energy proportion.

- Peak-time energy consumption rate = energy consumption during peak period/ total energy consumption = peak energy consumption of measurement point/ current day freezing energy. The energy data are calculated by the daily freezing indication and the real-time acquisition indication.

- The load rate includes the calculation of electric and thermal load rate. The calculation formula is the same. It is the average power (heat) / maximum power (heat) of the user. Namely, the daily electric (thermal) power of the measuring point/the maximum daily electric (thermal) power of the measuring point;

- The valley energy coefficient is divided into the valley energy coefficient of electric, water, gas and heat. The calculation formulas is similar, which is the energy consumption in the valley period/the total energy consumption.

- Flat time consumption energy proportion. The flat time consumption energy proportion is divided into the flat time consumption energy proportion of electric, water, gas and heat, and the calculation formula is similar. The flat time consumption energy proportion = The flat time consumption energy in the same day / the total energy consumption in the same day.

2.2.4. Construct the Entropy Model Matrix: Construct the matrix of n (number of users) rows and 4 columns with four energy characteristics and user objects.

2.2.5. Standardization of data: Standardization and normalization using formula (1).

\[
Y_{ij} = \frac{x_{ij} - \min(x_i)}{\max(x_i) - \min(x_i)}
\]

(1)


\(Y_{ij}\): The number of the i-th row and the j-th column of the matrix model after normalization.

\(x_{ij}\): Number of the i-th row and j-column of the original matrix model before normalization. It represents the jth energy characteristic value used by the i-th customer.

\(\min(x_i)\): The minimum value of the i-th row of the original matrix model.

\(\max(x_i)\): The maximum value of the i-th row of the original matrix model.

2.2.6. Calculate the energy use information entropy: Use formula (2)and(3) to calculate the information entropy.
$$E_i = -\ln(n) \sum_{i=1}^n p_{ij} \ln p_{ij}$$  \hspace{1cm} (2)$$

$$p_{ij} = Y_{ij} / \sum_{i=1}^n Y_{ij}$$  \hspace{1cm} (3)$$

If $P_{ij} = 0$, then define $\lim_{P_{ij} \to 0} P_{ij} \ln P_{ij} = 0$.

2.2.7. Calculate the weight of each energy-using characteristic: Use formula (4) to calculate the weight of energy-using characteristic.

$$W_i = \frac{1 - E_i}{k - \sum E_i} (i = 1, 2, \ldots, k)$$  \hspace{1cm} (4)$$

2.2.8. Calculate the energy use characteristic weight value: Combine the energy use characteristic value and the weight of energy use characteristic to calculate the energy use characteristic weight value.

2.2.9. Construct the K-MEANS model matrix: the user is taken as the analysis object, and the four-dimensional coordinate space is constructed with its four energy-using characteristics.

2.2.10. Matrix distribution of cyclic users: Resident users who participate in cyclic analysis are positioned in four-dimensional space.

2.2.11. Classification of object distribution: Classification and aggregation are carried out according to the aggregation degree of analytic objects.

2.2.12. User's energy behavior analysis: The integration statistics of energy-using data are collected for each type of classified objects, and the energy usage rules are analyzed, and the energy usage behaviors are analyzed to realize the user behavior classification of user classification.

3. Abnormal analysis of user energy use

"multi-meter unification" information acquisition system accumulates a large amount of customer energy consumption information, combined with a large number of typical abnormal energy use cases, and abstractly identifies the common factors of abnormal energy use behavior of users, then establishes a mathematical analysis model of abnormal prediction. Through the big data analysis technology, the suspected user of abnormal energy is probability inference and early warned, and the suspected users of abnormal energy use are accurately identified [7]. This stage is divided into two parts. The first stage is the comparison of the user's energy behavior class. Based on collaborative filtering algorithm, the energy use behavior is classified and summarized, and the abnormal energy use behavior is preliminarily distinguished. The second stage is based on Frechet distance algorithm, which further screens the abnormal energy use behavior. The User energy behavior classification and induction diagram is shown in figure 3.
3.1. *Inter class comparison of User Energy Use Behavior*

The inter class comparison of user energy use behavior refers to mining users with the same characteristics from different user behavior analysis. Firstly, the classification model of user's energy use behavior is established, and the similarity of user's attributes and behavior characteristics is calculated by using collaborative filtering algorithm. Then, the commonness of user's energy use behavior is extracted from similar users, and the target user's energy use classification and preference are calculated synthetically. Finally, the classification and the commonness of user's energy use behavior with the greatest similarity are obtained.

Collaborative filtering classification is different from traditional direct analysis classification which based on content filtering. Collaborative filtering analysis of user behavior is to find users with similar behavior in the user group, synthesize the energy characteristics of these similar users, and form the preference degree and behavior classification of the designated users for different energy use modes. In order to realize the recommendation algorithm of collaborative filtering, there are three steps: user single energy data analysis, identifying similar users (including similar behavior, energy preference), classification and preference induction [8].

3.1.1. *User single energy data analysis.* Various kinds of user's energy use history behavior data are obtained based on user's energy use behavior analysis of multi-meter data, such as user's energy use, energy use proportion, single energy use behavior classification, etc. These data can be used as analysis data for collaborative filtering algorithm and serve for collaborative filtering algorithm. It should be specially pointed out that different data have different accuracy and different granularity, and the impact of noise should be taken into account when using them.

3.1.2. *Find similar users (including similar behaviors, energy preferences)*

That is to calculate the similarity between users, including the similarity between behavior and energy preference.

- The similarity of user's energy use behavior is calculated according to the formula (5).

  \[
  \text{sim}(x, y) = \frac{\sum_{i \in I_x} (r_{ix} - \bar{r}_x)(r_{iy} - \bar{r}_y)}{\sqrt{\sum_{i \in I_x} (r_{ix} - \bar{r}_x)^2} \sum_{i \in I_y} (r_{iy} - \bar{r}_y)^2}
  \]

  (5)

  Where, \( I_x \) is a set of user x's energy-using behaviors.
  \( I_y \) is a set of user y's energy-using behaviors.
  \( I_{xy} \) is the union of \( I_x \) and \( I_y \).
  \( \text{sim}(x, y) \) is the similarity of user x and user y using Pearson correlation coefficient.

- The similarity of the user's energy preference is calculated according to the formula (6):
\[
\text{sim}(i, j) = \frac{\sum_x (r_{x,i} - \overline{r}_x)(r_{x,j} - \overline{r}_x)}{\sqrt{\sum_x (r_{x,i} - \overline{r}_x)^2 \sum_x (r_{x,j} - \overline{r}_x)^2}}
\]

(6)

Where,
\( r_{x,i} \) is the energy ratio of the user \( x \) to the energy \( i \).
\( \overline{r}_x \) is the user's average energy use ratio.

- Classification and behavioral induction

In order to better measure the similarity between users and improve the induction of similar behavior, the collaborative filtering algorithm adopted in this project is as follows:

1. Similarity is calculated and user behavior is classified based on the user's collaborative filtering algorithm.
2. According to the distance of similar users, the behavior preference degree of each user's energy consumption is calculated, and the corresponding user behavior preference ranking is given according to the level of each recommendation degree.
3. According to the user behavior preference ranking, the behavior characteristics of each type of user are summarized.

By mining and identifying the user's energy use behavior, a classification model based on collaborative filtering is established, and the corresponding similarity boundaries are determined, so that the best classification and behavior induction can be carried out for users. This link achieves the comparison between different types of energy use. Through identifying users who do not conform to the characteristics of categorized behavior in various types of user energy use behavior, the energy use data is put into the second stage.

3.2. Screening for Abnormal Energy Use Behavior

In order to further improve the accuracy of the prediction of suspected users' abnormal energy use, large data analysis technology is used to further verify whether the user's various energy use trajectories deviate from the normal energy use trajectories. If it deviates from the normal energy use trajectories, it further confirms that the user has the suspected of abnormal energy use.

For the key monitoring users entering the second stage, the Frechet distance discriminant model is used to judge whether the user's energy use behavior deviates from the normal energy use behavior trajectory by comparing the trend between the monitoring data and the user's historical data [9]. The specific process is as follows:

3.2.1. Receive relevant data of abnormal users initially identified in the first stage, including recent data and historical data of water, electricity, gas and heat.

3.2.2. Judge user categories. The typical behavior trend curves of users are established based on user categories. For example, for user's electricity consumption data, it can be divided into special transformer user and public transformer user. The special transformer user generates user's electricity consumption typical behavior trend curve based on user's historical electricity consumption data information; while public transformer user establishes user's electricity consumption typical behavior trend for all users in public transformer station area respectively.

3.2.3. Based on the discrete Frescher distance discriminant algorithm, after the user's recent energy trend curves and various typical behavior trend curves are discretized, the maximum distance between the two curves is taken as the curve similarity. And it is calculated according to formula (7).

\[
\delta_d(P, Q) = \max d\left( P^*_\alpha, Q^*_\beta \right)
\]

(7)
3.2.4. Different similarity thresholds are set based on different categories of users. If the curve similarity exceeds the threshold range, the abnormal energy use of the suspected user is output.

4. Examples of user abnormal electricity behaviour screening
As shown in Figure 4 and Figure 5, there are two typical power consumption curves for educational users during normal power consumption. Class 1 typical power consumption curve, the power consumption time is from 6 o'clock to 22 o'clock, peaking at 9 o'clock to 11 o'clock and 17 o'clock, and the electric current is within 2.5 A; Class 2 typical power consumption curve, the power consumption time is from 6 o'clock to 24 o'clock, between 8 o'clock and 17 o'clock is relatively stable, the electric current is within 8A. From the graph, we can see that there is a certain trend in the change of user's power consumption, and there will be no sudden increase or decrease, and the time of peak and trough is similar. Based on the discrete Frechel distance discriminant algorithm, the maximum distance between the user's recent trend curve and the typical trend curve is taken as the curve similarity after the discretization of the user's recent trend curve and the typical trend curve.

![Figure 4. User Class 1 Normal Electricity Curve](image1)

![Figure 5. User Class 2 Normal Electricity Curve](image2)

As shown in Figure 6, it is the abnormal power consumption data curve of the user. It can be seen that when abnormal power consumption occurs, the current of users becomes extremely high at 10 o'clock. This trend is far away from the typical characteristic curves of the two types, so it can be judged that there is a suspicion of abnormal power consumption.

![Figure 6. Contrast diagram of abnormal power consumption data curve](image3)

5. End
With the rapid advancement of the market economy and the continuous development of science and technology, all kinds of energy consumption information are constantly enriched. At the same time, all kinds of abnormal energy use means will continue to upgrade and change. Simply relying on one abnormal energy use situation may lead to missed judgment. In this paper, based on the user's energy behavior analysis method of big data, according to the characteristics of its energy use data, select their own characteristic data, through the analysis of various kinds of energy use data, get the user's energy use behavior analysis. The results can assist the popularization of safety energy knowledge, the investigation of lines, pipelines and energy-using equipment, and provide data support for future business expansion applications. According to the actual case, the study analysis and summary are carried out. Based on the big data processing technology to continuously improve the discrimination of abnormal energy use behavior, a joint analysis model is established. By mining the potential connection between the data, the failure of field equipment or the abnormal energy use of customers can be judged comprehensively, so as to give full play to the application value of various energy use data.
References

[1] Deng song, Yue dong, Zhu li peng, Hu bin and Zhou ai hua 2016 Framework of intelligent and efficient analysis and mining technology for power big data Journal of Electronic Measurement and Instrumentation vol.30(11) pp. 1679-86

[2] Xue yusheng and Lai yening 2016 Integration of macro energy thinking and big data thinking Automation of Electric Power System Vol.40(1) pp.1-8

[3] Duan Jun-hong, Zhang Nai-dan, Zhao Bo and Yan Xiao-bin 2015 Research on Basic System Architecture and Application of Electric Power Big Data ELECTRIC POWER ICT vol.13(02) pp.92-5

[4] Zhang Lei, Huang Xin and Chen Yiyu 2019 Literature Review on Residential Energy Behavior Urbanism and Architecture Vol.16(02) pp.160-4

[5] Wang saiyi, Yu jianping, Sun fengjie, Wang chengmin and Xie ning 2019 Evaluation and promotion methods with improved k-means for value density of electric power big data Smart Power Vol.47(03) pp.8-15

[6] Zhao Li, Hou Xingzhe, Hu Jun, Fu Hong and Sun Hongliang 2014 Improved K-Means Algorithm Based Analysis on Massive Data of Intelligent Power Utilization Power System Technology Vol.38(10) pp. 2715-20

[7] Ma tiannan, Wang chao, Peng lilin, Guo xiaofan, Du ying and Gou quanfeng 2018 Research on consumer energy use behavior forecasting and analysis of integrated energy system under multi-source heterogeneous data Smart Power Vol.46(10) pp.86-95

[8] Wu Jiajing,He Jianan,Wang Yuequn and Dong Liyan 2018 Research on Collaborative Filtering Algorithm Based on Items’ Attribute Categories Journal of Jilin University(Information Science Edition) Vol.36(04) pp. 470-4

[9] Zhu jie, Huang zhangcan and Peng xiaolin 2019 Curve Similarity Judgment Based on the Discrete Fréchet Distance J. Wuhan Univ.(Nat. Sci.Ed.) vol.55(02) pp.227-32