A Precision Analysis Method of Electrical Characteristics for Demand Response

Chen Wu, Zhenshu Wang*, Jincheng Ding and Zixuan Du
School of Electrical Engineering, Shandong University, Jinan 250000, China

*Corresponding author e-mail: zhenshuwang@sdu.edu.cn, wuchen260123@foxmail.com

Abstract. In order to accurately grasp the users' electricity consumption characteristics during the implementation process of the demand response (DR), a clustering algorithm was used to realize delicate analysis of demand-side users. Firstly, the typical load curve of a single user was extracted by means of averaging. Then, an improved adaptive bilayer spectral clustering algorithm was adopted to cluster typical load curves. The simulation results show that the method can accurately analyze the demand-side power load, the inner and outer spectrum clusters can meet the requirements of different demand response projects, can guide the development of the time-of-use price, and can calculate the interrupted response capacity. At the same time, it can provide a theoretical basis for consumption of renewable energy with DR.

1. Introduction

The construction of the energy Internet has promoted the development of the use of electricity data, making the use of energy-end data continuously accumulate. The massive data of the power grid leads to a simple statistical analysis of the load characteristics has little significance. It is even more unrealistic to use different control strategies for each load. Therefore, it is necessary to classify the load that participates in the demand response. The first step to be made whether it is incentive-based DR or price-based DR is to classify the load. Through the mining of electricity data and the perception of electrical characteristics, it can effectively identify the user's power consumption mode, assess the potential of demand response, guide the formulation of electricity prices, establish a corresponding response model, and provide personalized services to its users.

2. Similarity Measure

The selection of similarity measures is crucial to the clustering result [1~3]. Most of the current load clustering algorithms are clustering of distances, and a single Euclidean distance is often used in the similarity measure [4]. The difference in the recognition of morphology is poor and cannot meet the requirements of demand response. In this paper, a bilayer spectral clustering algorithm is designed to classify the shape of load curves in the outer layer and the inner layer mainly clusters the distance of the curve. The outer-layer spectral clustering uses Cosine distance for similarity measurement, and the inner-layer spectral clustering uses Euclidean distance as a measure of similarity.
3. Bilayer spectral clustering algorithm

3.1. The process of spectral clustering algorithm

The process of spectral clustering algorithm is as follows [5]:

(1) Gaussian kernel functions are used to build a similarity matrix instead of an adjacency matrix. The specific formula is as follows:

\[ H = \exp \left( - \frac{\text{sim}(x_i, x_j)^2}{2\gamma^2} \right) \]  

(1)

Among them, \((x_i, x_j)\) is the element in the matrix obtained by the similarity measure and \(\gamma\) is the parameter of the Gaussian kernel function, which is 1 in the experiment.

(2) Adding each column element in the similarity matrix to a diagonal line gives the degree matrix \(D\). And calculating the Laplacian matrix from the degree matrix according to the following equation:

\[ L = \text{eye} - D^{\frac{1}{2}} \times H \times D^{\frac{1}{2}} \]  

(2)

Obtaining Laplacian Matrices Using Normalized Similarity Transforms. Where \(\text{eye}\) is the unit matrix.

(3) The eigenvector corresponding to the smallest \(K\) eigenvalues of the Laplacian matrix is calculated. And this vector is clustered using k-means algorithm.

3.2. The programming of bilayer spectral clustering algorithm

The block diagram of the bilayer spectral clustering algorithm used in this paper is as shown in Fig. 1.

![Figure 1](image)

Figure 1. The block diagram of the bilayer spectral clustering algorithm.

The inner layer and the outer layer select different similarity measures, which is reflected in the different construction of the similarity matrix when the spectral clustering selects feature values, and the different distance discriminating methods of the invoked k-means function, so as to distinguish the form and the distance.

3.3. Optimization of the number of clusters

The validity index of clustering can be used not only to evaluate the effect of cluster analysis but also to determine the number of clusters. Considering that the DBI index based on Euclidean distance does not apply to cluster validity evaluation based on the cosine distance, the following formula is used to evaluate the similarity of outer clusters:

\[ DB_w = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} (S_i + S_j) \]  

(3)
Where $S_i$ represents the cosine distance error between the data points and the clustering centers in class $i$; $k$ indicates the number of cluster categories. The greater the intra-class similarity, the smaller the angle between the two curves, and the smaller the corresponding index value. The inner layer clustering uses the DBI index as the selection basis for the number of clusters. The DBI indicators also reflect two evaluation criteria: the similarity within the class and the similarity between the classes, and the index is insensitive to boundary points. It is more suitable for the evaluation index of load clustering. Select the number of categories corresponding to the minimum value of the DBI index as the best number of categories \[6\]. The formula is as follows:

$$DBI=\frac{1}{k}\sum_{i=1}^{k}\max_{i\neq j} \frac{S_i + S_j}{d_{ij}}$$  \hspace{1cm} (4)

Where $S_i$ represents the standard Euclidean distance error between the data points and the clustering centers in class $i$; $d_{ij}$ represents the distance between clustering centers of classes $i$ and $j$.

### 3.4. Optimization of the clustering initialization center point

Spectral clustering clusters the eigenvectors corresponding to the maximal eigenvalues of each piece of data. The k-means clustering function that comes with MATLAB is used directly. The initial center point is selected randomly and it is easy to make the class criterion function falls into a local minimum, has a great influence on the clustering result, and the classification result of each running program is occasionally not the same. Therefore, the density method is used to select the initial center point instead of the randomly generated initial center point in the program \[7\]. In this way, the data distribution characteristics of the objects to be classified can be reflected to a certain degree, and it is closer to the actual clustering center.

### 3.5. The algorithm steps

1. Optimize the initial clustering center, and iterate the indicators on the load data to get the best number of classifications.
2. Outer spectral clustering was performed according to the best classification number.
3. The inner layer clustering is performed on each outer layer load data, and the results of the inner layer clustering are fed back to the outer layer for uniform numbering as the result of the inner layer classification.

### 4. Analysis of examples

The dataset DS1 is clustered using bilayer spectral clustering, and the optimal classification number of the inner and outer clusters is determined according to the corresponding indexes. It is as shown in Fig. 2.
The optimal number of classifications for outer clustering is 4, and the corresponding inner classifications are 4, 2, 2 and 2. According to the best selected number of classifications, the outer spectral clustering results is as shown in Fig.3.

The load contained in each cluster is different in load size, but its time scale and shape are similar. When participating in price-based demand response, it can be used as a type of dividing peak-to-valley time and establishing the time-of-use price.

The traditional k-means algorithm divides the dataset DB1 into four classes and adopts Euclidean distance as a similarity criterion. The result is as shown in Fig.4.

In Fig.3, (b) and (d) are classified into different classes to meet the requirement of peak-to-valley time division. At the same time, their load duration is short and they can participate in short-term or fast demand response projects. Clustering using the traditional k-means clustering method will be classified into a cluster, as shown in (d) in Fig.4. Because the two types of load curves use Euclidean
distance to measure the difference is very small, the classification result cannot meet the requirements of the demand response.

The first outer layer is divided into four categories according to the indicators. The results are shown in Fig.5. The rest of the outer layers are no longer detailed.

![Figure 5. Inner spectral clustering results.](image)

According to the clustering result graph, it can be seen that the inner clustering classifies the load size finely to meet the requirements of fine-grained load management, and at the same time, it can obtain the interruptible response capacity of each type of user and can participate in incentive-based demand response project.

(1) Analysis of the running time of the algorithm

This paper uses bilayer spectral clustering algorithm, in this algorithm, k-means clustering is performed by using the eigenvectors corresponding to the first k eigenvalues, which reduces the data amount to a great extent, and thus improves the running speed of the algorithm.

(2) Analysis of the stability of the algorithm

Using k-means for clustering, the results of each operation are different. The bilayer spectral clustering algorithm proposed in this paper optimizes the selection of initial center points. The results of each run are the same, greatly improves the stability of the algorithm.

5. Conclusion

In this paper, a precision analysis method of electrical characteristics for demand response is proposed. Through the bilayer spectral clustering algorithm, the inner and outer layers adopt different similarity measurement methods to achieve effective clustering of load data. Simulation results shows:

(1) The bilayer spectral clustering algorithm can not only identify the loads with similar morphology, but also distinguish different amplitudes in the same modality. So it can meet the requirements of refined load management and adapt to different demand response projects.

(2) The clustering algorithm proposed in this paper is superior to the traditional k-means algorithm in terms of clustering validity, running time and stability of the algorithm.

Acknowledgments
Project Supported by National Natural Science Foundation of China (NSFC) (51377099).

References
[1] Huang Yuteng, Hou Fang, Zhou Qin, et al. A new combinational electrical load analysis method for demand side management [J]. Power System Protection and Control, 2013 (13): 20-25.
[2] ZHANG Zhisheng, SUN Yaming Sun, ZHANG Shiyong, et al. Clustering analysis of electric load series using clustering algorithm of multi-hierarchy and detailed decomposition based on data mining [J]. Power System Technology, 2006, 30 (2): 51-56.
[3] Jia Huimin, He Guangyu, Fang Chaoyong, et al. Load forecasting by multi-hierarchy clustering combining hierarchy clustering with approaching algorithm in two directions [J]. Power System Technology, 2007, 31 (23): 33-36.
[4] Kwac J, Flora J, Rajagopal R. Household Energy Consumption Segmentation Using Hourly Data [J]. IEEE Transactions on Smart Grid, 2014, 5 (1): 420-430.
[5] Lin Shunfu, Tian Erwei, Fu Yang, et al. Power load classification method based on information entropy piecewise aggregate approximation and spectral clustering [J]. Proceedings of the CSEE, 2017, 37 (8): 2242-2252.

[6] Zhao Li, Hou Xingzhe, Hu Jun, et al. Improved k-means algorithm based analysis on massive data of intelligent power utilization [J]. Power System Technology, 2014, 38 (10): 2715-2720.

[7] Han Lingbo, Wang Qiang, Jiang Zhengfeng, et al. Improved k-means initial clustering center selection algorithm [J]. Computer Engineering and Applications, 2010, 46 (17): 150-152.