

## Research of Power Consumers Behavior Using Fuzzy C-means Algorithm

Xiaohui Chang<sup>1</sup>, Lei Yu<sup>1</sup>, Yulong Han<sup>1</sup>, Tao He<sup>1</sup> and Xiayi Hao<sup>2, a</sup>

<sup>1</sup> Zhangjiakou Power supply Co.Ltd, State Grid Jibei Electric Power Company, Hebei, China

<sup>2</sup>Tianjin University of Technology and Education, Tianjin, China

<sup>a</sup>tutezdh1203@163.com

**Keywords:** Clustering, fuzzy C-means algorithm, monthly load curve, concavity degree analysis, power service strategy.

**Abstract.** Fuzzy C-means (FCM) algorithm is used to cluster monthly power consumption data of power consumers, which has been pretreated, in a certain domestic area, and power consumers clusters in three different months and each of clustering centers can be obtained. According to the algorithm, the reliability of the algorithm is evaluated. The clustering center obtained in a certain month is brought into an algorithm for evaluating the degree of coincidence, by which the clustering result of the month can be proved reasonable. Based on the above clustering results, it can be seen that there are some characteristics in different types of power consumers in August and fluctuations within power grid. On the basis of the above described, power companies can make out different power service strategies for all kinds of power consumers so as to reduce fluctuations in power grid and improve the efficiency of power usage.

### Introduction

Power companies can enhance consumers experience of power consumption, adjust the construction of consumption, reduce consumption and improve electricity efficiency by analyzing the behavior of power consumers, which can be a strong support for energy-saving, grid marketing strategy and diversified customer service [1-5]. In general, pattern recognition is often used to analyze the behavior of power consumer. The method usually follows these four criteria [6,7]: 1) Each type of power consumers should have the similar behavior; 2) There is a clear difference between each group; 3) Identification algorithm is easy to implement; 4) Final category should not be too much.

Based on the behavior of power consumption of different consumers, Fuzzy C-means (FCM) algorithm is used to process the data to realize the clustering for different types of consumers. After analyzing the data of power consumption of consumers extracted in August, it is concluded that the types of power consumers in this area can be divided into some types. And based on this conclusion, it can provide guidance for power company to formulate relevant policies.

### Data Preprocessing

The data used for analysis in this paper comes from the Power User Electric Energy Data Acquire System in a common domestic area. Errors such as missed sampling and wrong sampling in data collection are inevitable which results in the necessity to preprocess the data before analysis.

For the missed sampling data, the average of the nearly last collected data and the next data is calculated to be the substitution. When the number of missed sampling data is large, this data will be regarded as a special and removed.

For the wrong sampling data, when there is a negative sampling data, it is feasible to correct. If there is so much negative sampling data, this data will be also removed so as to avoid affecting the stability of the algorithm. And when it happens, it should be checked whether the equipment has been damaged by human/nature or not.

Due to the difference of consumption behavior among customers, the data have different characteristics of multiplicity and diversification. In order to reduce the errors and enhance the degree

of aggregation between the customers of the same consumption behavior, the data of each customer is normalized, that is, the data value is limited to a varying from 0 to 1.

This paper aims to study the customers' power consumption behavior in a specific area. Changing the value of data or removing some abnormal data does not affect the overall implicit information of the original data.

### Clustering Algorithm

Fuzzy C-means (FCM) algorithm is based on the common C-means algorithm and generated by concept of fuzzy theory [8]. It makes the boundaries between different clusters uncertainly, which produces a possibility that one data may have different percentage to every clusters. It greatly reduces the effect of human factors, so that the clusters become more convinced. The basic idea of FCM algorithm is as follows: C value is taken as the clusters, and the data is partitioned by fuzzy affiliation probability generated by membership function; Through making the optimization and adjustment for the clustering centers, the data can be classified many times and the good clusters can be achieved finally.

The basic steps of the algorithm are shown as follow:

- 1) Determine the value of C, which is represented as the number of clusters.
- 2) Generate the initial fuzzy matrix according to the requirement, as shown in Eq. 1.

$$\sum_{j=1}^c u_{ij} = 1, \quad \forall i = 1, \dots, n \quad (1)$$

Wherein,  $u_{ij}$  represents the membership degree of the  $i$ -th data in the  $j$ -th cluster center, which is equal to probability.

- 3) Calculate the initial cluster center using Eq. 2.

$$cluster_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (2)$$

Wherein,  $cluster_j$  represents the  $j$ -th cluster center;  $m$  is a weighted index, generally taking 2;  $x_i$  represents the  $i$ -th data vector.

- 4) The data are classified based on the affiliation within the fuzzy matrix.
- 5) After all the data are classified, the fuzzy matrix is optimized by Eq. 3, and updated.

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{d_{ij}^2}{d_{ik}^2} \right)^{\frac{1}{m-1}}} \quad (3)$$

Wherein,  $d_{ij}^2$  represents the square of "relative distance" between the  $i$ -th data and the  $j$ -th cluster center.

- 6) Calculate the objective function, as shown in Eq. 4;

$$J(U, cluster) = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m d_{ij}^2 = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \|cluster_j - x_i\|^2 \quad (4)$$

Wherein,  $U$  is fuzzy matrix.

- 7) Repeat steps 3) to 6) until the number of iterations overflows or the objective function error is less than the preset parameter.

### Analysis for Behavior of Power Consumers with FCM Clustering Algorithm

In order to analysis the behavior of power consumers under the characteristics of monthly load, the power consumption data are divided into August, September and October three parts according to the month, which are respectively brought into the clustering algorithm for analysis.

The preprocessed data is brought into FCM clustering algorithm and run. It is found that the power consumers in this area can be clustered into 2 categories in August, and the clusters of load curve and clustering centers are shown in Fig. 1. Among them, the clusters are run on MATLAB 2015b version.

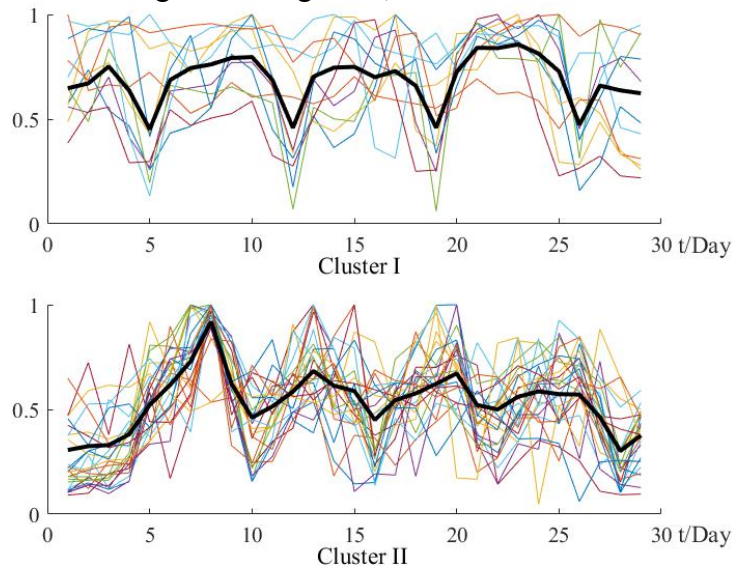


Fig. 1 The clusters of behavior of power consumers in August

Comparing the clusters in Fig. 1, it can be found that there are different types of clustering centers for power consumers in different months.

## Results analysis

Table 1 shows the power consumers belonged to the different clusters obtained by FCM clustering algorithm.

Table. 1 Customers after clustering monthly

Cluster	Customers(in August)
I	23,25,26,28,30,31,32,
	33,34,35,38,40,42,44,
II	45,46,47,49,51,53,56,
	57
II	1,2,3,4,5,6,7,8,9,10,11,
	12,13,14,15,16,17,18,
II	19,20,21,22,24,27,29,
	36,37,39,41,48,50,52

After the clusters and clustering centers are analyzed in Fig. 1, it is found that the values of cluster I are all dropped down with a certain degree every weekend, and the fluctuation of power load at working days does not change much. In addition to the actual situation, the load of enterprise, factory-type power consumption is more affected by the temperature than that of resident, business-type. Therefore, the cluster I can be recognized as the load curve of enterprise, factory-type electricity consumption, the cluster II can be summarized as residents, business-type electricity load curve. Moreover, when resident, business-type of power consumers reach the peak of electricity consumption, enterprises, factory-type power consumers need to reduce their power load during this period due to the relevant policies.

## Conclusions

FCM algorithm is used to process the data to acquire different clusters for power consumption in the studied area. The types of power consumers can be roughly divided into resident, business-type,

enterprise, and factory-type. Firstly, the clusters are analyzed again, and the power consumers are classified more meticulously, so as to realize the purpose of using the algorithm to classify power consumers. Secondly, on the basis of that, it finds that the behavior of power consumption of residents is similar to that of the middle-class, enterprise, factory-type to that of the metallurgical or food processing factories. It will help give the power company a guidance to formulate power service policies in this area so as to reduce the fluctuation of power grids and improve the efficiency of power usage. Lastly, learning that power consumption in the area will enable company to prepare ahead of schedule for dealing with emergency power dispatch.

### Acknowledgements

This work was financially supported by the Science and Technology Project Funded by Zhangjiakou Power supply Co. Ltd, State Grid Jibei Electric Power Company (SGTYHT/16-JS-198).

### References

- [1] K.Y. Liu, W.X. Sheng, D.X. Zhang, et al. Big Data Application Requirements and Scenario Analysis in Smart Distribution Network [J]. *Proceedings of the CSEE*, 2015, 35(2): 287-293.
- [2] C.K. Shi, B. Zhang, W.X. Sheng, et al. A Discussion on Technical Architecture for Flexible Intelligent Interactive Power Utilization [J]. *Power System Technology*, 2013, 37(10): 2868-2874.
- [3] J. Lu, Y.P. Zhu, W.H. Peng, et al. Feature Selection Strategy for Electricity Consumption Behavior Analysis in Smart Grid [J]. *Automation of Electric Power Systems*, 2017, 41(5): 58-63.
- [4] Z.W. Liao, Y.M. Sun. Data Mining Technology and Its Application on Power System [J]. *Automation of Electric Power Systems*, 2001, 25(11): 62-66.
- [5] Y.T. Huang, F. Hou, Q. Zhou, et al. A new combinational electrical load analysis method for demand side management [J]. *Power System Protection and Control*, 2013, 41(13): 20-25.
- [6] BAILEY J.. Load profiling for retail choice: examining a complex and crucial component of settlement [J]. *The Electricity Journal*, 2000, 13(10): 69-74.
- [7] X.G. Peng, J.W. Lai, and Y. Chen. Application of clustering analysis in typical power consumption profile analysis [J]. *Power System Protection and Control*, 2014, 42(19): 68-73.
- [8] Y.H. Zhu. Research of the Related Problems on Fuzzy C-Means[D]. China: Ocean University of China, 2011.