Intelligent Park Automatic Demand Respond Characteristic Analysis Based on Optimal Fuzzy C-means Clustering Algorithm

Jian-cheng YU¹, Si-wei Li², Liang YUE², Jie ZHANG², Yong-li WANG³ and Jia-pu Li³

¹Stage Grid Tianjin Electric Power Company, Hebei District, Tianjin 300010, China
²Fibrlink Corporation Stage Grid Information and Telecommunication Group, Beijing 100070, China
³School of Economics and Management, North China Electric Power University, Changping District, Beijing 102206, China

Keywords: Intelligent park, Automatic demand response, Electric fuzzy clustering, Fuzzy C-means.

Abstract. Firstly, This paper introduces automatic demand response technology and fuzzy C-mean clustering algorithm, then An algorithm based on fuzzy C means clustering is used to cluster analysis of the daily load data of a distribution station in a certain Eco-city, and the optimal number of load clustering is calculated, and the optimal load clustering results are obtained. At last, this paper summarizes the load response characteristics of various types, and provides a research idea for the cluster analysis of demand response to user load characteristics.

Introduction

The global energy Internet is characterized by clean energy, UHV and smart grid technology, Intelligent park is a typical load system for building the energy Internet at the end of the user's terminal. The key to the realization of the automatic demand response technology in the intelligent park is the comprehensive planning of the user's side load and resources, and study the actual curve characteristics of the user load and the load response characteristics of different types of users, implement the automatic demand response scheme for different types of user loads, So as to improve the security, stability and reliability of power supply in the transmission and distribution network, To realize the improvement of energy efficiency of regional users, as well as the safe use of electricity, intelligent electricity, energy saving, energy saving, environmental protection and orderly use of electricity, has become one of the new hot topics to study the technology of energy Internet demand response.

Some scholars combined with the new requirements of smart grid load forecasting, put forward a platform architecture of smart grid load forecasting based on cloud computing platform, makes up the traditional load prediction technology in the unreasonable allocation, data processing and other aspects of the improper[1].In the establishment of the smart grid demand response standard system, the references [2] has done a related research, combing the framework of the relevant standards. In terms of research on demand response technology in intelligent parks, references [3] have proposed a two-way interactive demand response strategy for the daily load demand response problems involving multiple intelligent power schemes. In order to determine the number of clustering, the references [4] proposed a clustering effective discriminant index based on the FCM algorithm. References [5] proposes an optimization algorithm based on K-mean in power load clustering analysis, which improves the efficiency and reliability of output results. Ming-zhao CHEN and Jian MAO applied clustering algorithm to the demand side management of industrial users. Two interrupting capacity types of climate sensitive load and production load were proposed, and the corresponding DSM type was obtained [6].References [7], considering the load curve form of power users, proposes a method of power load model extraction based on cloud model and fuzzy clustering. In references [8],the response incentive mechanism that affects the load of the civil load system and the capacity of the energy storage equipment is proposed in reference to the resident load containing the photovoltaic and energy storage devices.
To sum up, clustering load according to the characteristics of user load is difficult to meet the management needs of lean and complex power demand side only by classifying the load according to the traditional load classification method. Combining the load clustering technique based on the actual load curve and the automatic demand response technology of electric power, using cluster algorithm to cluster analysis of user load according to the characteristic of actual curve, Responding to different load characteristics from different load types, combining the actual demand response characteristics of regional users and the configuration of user side demand respond resources like energy storage device and distributed generation, Develop a different response scheme and implement an automatic demand response in the region. As the result, The power demand response scheme is more scientific and more feasible, and it also makes the power demand side management more systematic and targeted.

**Fuzzy C-means Clustering Algorithm**

Clustering algorithm, which is known as cluster analysis, is a statistical analysis method to study the classification of samples or indicators, and it is also an important algorithm for data mining. The analysis consists of a number of clustering, it can often be a metric vector, or a point in a multidimensional space, clustering analysis based on similarity and clustering in a set, membership between the elements between elements in the same cluster has more similarity. FCM algorithm is a flexible clustering method based on fuzzy partition. It optimizes the objective function to get the membership degree of every sample point to all the class centers, and determines the class of the sample points to achieve the purpose of automatically classifying the sample data. At present, there are many ways to load classification based on actual load curve. As one of the most widely used classification algorithms for power load characteristics, Fuzzy C means clustering algorithm is more ideal than other algorithms in terms of running time, accuracy of algorithm results, algorithm stability and final clustering effect[5].

In FCM algorithm, For the given set of data \( X = \{x_1, x_2, \ldots, x_n\} \), \( x_j \in \mathbb{R}^s \),That is, each sample in the data set is a S dimensional vector, divide X into \( c(2 \leq c \leq n) \) classes, The set of vector sets of each cluster center is set to \( V = \{v_1, v_2, \ldots, v_c\} \),The membership \( x_j \) degree of the data samples belonging to the class cluster center \( i \) is \( u_{ij} \in [0,1] \),\( \sum_{i=1}^{c} u_{ij} = 1, j = 1,2,\ldots,n \).The objective function of FCM is the sum of the product of the degree of membership and the Euclidean distance between the point and the cluster center. The FCM algorithm is the partition matrix U and the cluster center matrix C, which minimizes the clustering objective function. The expression of the objective function of the algorithm:

\[
J(U,V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \quad (1)
\]

In the formula (1), \( d \) is the Euclidean distance between the sample \( x_j \) and the cluster center \( v_i \),the algorithm is \( d = \| x_j - v_i \| \),and \( m \) is the fuzzy weighting coefficient that represents the degree of fuzziness of the classified matrix. The membership degree \( u_{ij} \) and the calculation formula of the cluster center \( v_i \) are shown in the following formula:

\[
u_{ij} = \frac{1}{\sum_{j=1}^{c} (d_{ij} / d_{ij})^{(m-1)}} \quad (2)
\]
\[ v_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \]  \hspace{1cm} (3)

The implementation of the FCM algorithm is to iteratively update the clustering center matrix \(C\) and the membership matrix \(U\). The specific iterative steps of the algorithm are as follows:

**Step 1:** Given the number of load clustering \(c\), fuzzy weighting index \(m\) and iterative stopping threshold, take \(k\) to be 0, initialization of the membership degree matrix \(U^0\) and cluster center matrix \(C^0\).

**Step 2:** Update \(U^k\) according to the formula (2), \(\forall i, j, if \ d_{ij}^k > 0\), then

\[
u_{ij}^k = \left\{ \sum_{r=1}^{c} \left( \frac{d_{ij}^r}{d_{ij}^k} \right)^{2(m-1)} \right\}^{-1} = 0, then \ u_{ij}^k = 1 \ and \ j \neq k, \ u_{ij}^k = 0. \]  \hspace{1cm} (4)

**Step 3:** Update \(C^{k+1}\) according to the formula (4):

\[
C^{k+1}_i = \frac{\sum_{j=1}^{n} (u_{ij}^k)^m x_j}{\sum_{j=1}^{n} (u_{ij}^k)^m}
\]  \hspace{1cm} (5)

**Step 4:** Computational objective function according to the formula (2), if \(\|C^k - C^{k+1}\| < \varepsilon\), then stop the iteration, the clustering results are output. If not, take \(k = k+1\), back to step 2.

The clustering validity of the FCM clustering algorithm needs to be tested to judge the rationality of the results. In this paper, the clustering validity of the algorithm is analyzed from the index method of the possibility division of the clustering algorithm in [13]. The formula is as follows:

\[
P(U, c) = \frac{\min_{j=1}^c \left( \sum_{i=1}^{n} u_{ij} \right)^2 \left( \sum_{i=1}^{n} u_{ij} x_i - \frac{\sum_{i=1}^{n} \sum_{i=1}^{c} u_{ij} x_i - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} u_{ij}} \right)^2}{\max_{j=1}^c \left( \sum_{i=1}^{n} u_{ij} \right)^2 \left( \sum_{i=1}^{n} u_{ij} x_i - \frac{\sum_{i=1}^{n} \sum_{i=1}^{c} u_{ij} x_i - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} u_{ij}} \right)^2}
\]  \hspace{1cm} (6)

\(x_i \in R^m, i = 1, 2, ..., n\) is m-dimensional sample vector;

\(v_i \in R^m, i = 1, 2, ..., n\) is cluster center;

\(U = \{u_{ij}\}\) is Membership degree matrix.

The optimal value of the number of fuzzy clustering is checked and judged by calculating the clustering validity function, and the number of clusters corresponding to maximum values of clustering validity function is the best number of clustering.

**Example Analysis**

**Example Scenario**

The example scene selection is the 96-day load data of the 10KV/400V distribution station in an Intelligent Park. The ecological park has now constructed the energy optimization network and the
intelligent public service network. The energy Internet in the region has diversified resources such as photovoltaic, wind power, unit, energy storage, controlled load that has built smart home, automatic demand response, megawatt level microgrid, smart city comprehensive energy service platform and Integrated storage and storage scheduling platform for smart grid source network, the automatic demand response potential and the facilities conditions are ideal, Therefore, it is used as the selection of clustering data and the example object.

The operation of the automatic demand response system of the park requires the coordinated operation of the three modules of ADR, distribution network and users. The connection among the three modules depends on the power equipment, the distribution grid, the information flow and the energy flow. In these three modules, the user module is the base and the main object of the response, the ADR module is the "nerve center system", which affects the whole response system. The distribution network module is the top operation mechanism of the response system, which is responsible for the operation and promotion of the automatic demand response system and the related information collection work. Government departments play a regulatory role in the top level supervision of distribution networks and service providers, and also formulate relevant policy mechanisms.

**Selection and Preprocessing of Daily Load Curve Data**

In order to reduce the interference of other external conditions, such as time, environment, and other factors, The calculation example uses the 96 point daily load data of the 20 stations in the ecological city for 30 days, The typical daily load curve data are obtained by means of the mean, and the computational procedure are as follows:

- 96 day load curve selected from each distribution station from July 3, 2016 to August 3rd for January, then the mean value of each distribution station 30 days average daily load curve as the typical daily load curve in the time period of each distribution station.

It is expressed as this formula:

\[ x_i = [x_{i1}, x_{i2}, \ldots, x_{im}] \]

Due to the large difference of load data in different scale distribution stations, the extreme sequence of typical daily load curves of each group is normalized (the normalized data \( x_{ij} \in [0,1] \)). The normalization formula is as below:

\[
\hat{x}_{ij} = \frac{x_{ij}}{\max[x_{i1}, x_{i2}, \ldots, x_{ij6}]} \quad (7)
\]

According to the typical daily load curve data calculated by the example, after normalizing the extreme value sequence, 20 normalized daily load curve data are obtained. The greater the value of the fuzzy weighting coefficient is \( m \), the more fuzzy the clustering results are, the less distinct the difference between each class in the results of the FCM classification is obtained. The interval of \( M \) is \([1, +\infty)\). According to the results of literature, Pal et al. From cluster validity aspect, the optimal interval of \( M \) is \([1.5, 2.5]\). When there is no special requirement, \( m \) value is 2. According to the FCM clustering algorithm in the part 3, we classify the 20 load curves. After iterating the number of clusters in FCM, we get the best clustering number is 5 when the clustering validity function \( P(U, C) = 0.438 \) reaches the end, and output the fuzzy clustering iterative results. The result of the final load clustering is as shown in the following figure:
Figure 1. Cluster result type I to V of daily load curve.

From the final clustering results of 20 load data, we can see that each classification can clearly represent the load characteristics of a typical load, and the clustering results are ideal. For example, the load level of class I was stable at 0-6, and kept at a lower level, at about 6, the load level rose rapidly and reached the peak from 12 o'clock to 13 o'clock, then keeps down to 20 o'clock and keep low load level. The overall trend of class II load to rise after the first drop further, from 0 o'clock to 8 o'clock, the overall load tends to decline, and gradually increased from 16 o'clock to 20 o'clock to the load peak and begin to decline to close to 24 o'clock. Class III load increased steadily from 6 to 22, and reached a peak at 17 to 18 o'clock. The load type is similar to the distribution station of residential area, and the load level is at the highest level before and after dinner time. Group IV reached peak load at 8 o'clock and 17 o'clock respectively, the lowest load time was 4 to 6 AM, and the total load level of class V increased at first and then decreased, and reached the peak load at 5 PM.

According to the analysis of the above load cluster data, it can be concluded that for the whole system, in 0:00 to 4:00, it is necessary to respond to the low power load of class I and v when time during 0 o'clock to 4 o'clock. From 6 am to 4pm, 8pm to 24pm, the response measures to type IV and type I load response are needed respectively. In the peak response of the power load, The peak shaving response should be carried out for class II load at about 8 am, and the user load of class V should be involved in the peak shaving response from 10 o'clock to 17 o'clock. From 12 to 14 o'clock, class I load will also respond to the peak shaving response of the power load. From 16 to 17 pm, the users of class II and class III need to take part in the power load shaving response, from 17 to 22 PM, the demand response to the power load peeling is the user of the fourth class.

In terms of the automatic demand response strategy, we can design the electricity demand plan of the park's automatic demand response according to the user types of the eco city system and the diversified resources such as photo voltaic, wind power generation, triple supply unit, energy storage and controllable load on the user side scale: large industrial users can reduce peak load by using the configured storage installed in the electric load peak discharge to carry out spontaneous self use, and commercial users take part in the demand response by reducing the flexibility of the interruptible load (such as central air conditioning),The power supply side can give the appropriate price response, demand subsidies and other ways to participate in the electric power peak shaving, or establish typical load multi time scale electricity price demand response model, the electricity price mechanism including multi time load and TOU price should be designed. As well as the Analysis of applicability of incentive measures in different stages of power market, and the interest
game relationship of load response is analyzed, and incentive measures are established. Similarly, we can encourage the residents to use load response devices such as washing machines and water heaters to fulfill the load filling in the low load period by charging the energy storage devices and subsidized incentives in the low load period, and through the cogeneration unit to adjust the cold, heat and electricity load peak, distributed photovoltaic power consumptive area. While ensuring the stability and reliability of the operation of the integrated energy system in the park, it plays an automatic demand response to the controllable load of the user.

By clustering analysis of the above user load, the power supply side makes different time sharing incentive strategies or time sharing optimization schemes for users with different load curve characteristics and different load response characteristics. It encourages users to participate in demand response, so as to achieve the goal of system planning and operation to improve the energy efficiency level of regional power users, and enhance the safety, stability and reliability of power grid operation.

**Conclusion**

Based on the FCM algorithm, this paper cluster analysis on the power load characteristics of the automatic demand response users in the Intelligent Park, introduces the automatic demand response technology and the fuzzy C mean clustering algorithm and implementation process, and applies the algorithm based on fuzzy C means clustering to cluster analysis of the daily load data of distribution stations in an eco city. The number of optimal load clusters is calculated by iteration, and the optimal load clustering results are obtained. The load response characteristics of each type are summarized, which provides a research idea for clustering analysis of load characteristics of users in response to demand.

In this paper, we use fuzzy C mean clustering algorithm to cluster the analysis of the characteristics of user power load, and use clustering validity function to optimize the number of clustering C, and get the optimal number of clusters, the clustering results of 96 point daily load curves of 21 stations in an ecological park from July 3, 2016 to August 3rd are output by iteration. Clustering results divide user load into 5 categories. According to different types of load characteristics, this paper puts forward differentiated time-sharing demand response intervals of different types of loads in the park system. The clustering results have ideal typicality and discrimination. Different clustering results can represent the user load of different curve characteristics, which can support the research on the load characteristics of the power users, and provide the basis and reference for the realization of the park's automatic demand response technology.

**References**


Classification Technology Research Based on an Optimal Fuzzy C-means Clustering Algorithm\textit{[J]}\textit{.} Journal of Zhengzhou University (Medical Sciences), 2015, 36(06):56-60.

