Research Article *ISSN: 2319-507X*
Rupali Meshram, IJPRET, 2013; Volume 1(8): 884-890 *IJPRET*

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

LOAD PROFILE CLUSTERING AND FORCASTING OF ELECTRICITY CUSTOMERS

RUPALI MESHRAM[1], PROF. A. V. DEORANKAR[2], DR. P. N. CHATUR[3]

1. M. Tech 2nd year, Department of Computer Science and Engineering, Government College of Engineering, Amravati (Maharashtra), India.
2. Asst. Professor, Department of Computer Science and Engineering, Government College of Engineering, Amravati (Maharashtra), India.
3. Head of Department, Department of Computer Science and Engineering, Government College of Engineering, Amravati (Maharashtra), India.

Accepted Date:
27/02/2013

Publish Date:
01/04/2013

Keywords

Forecasting electricity,
Classification,
Load Pattern Analysis,
Clustering,
Typical Load Profile.

Corresponding Author
Ms. Rupali Meshram

Abstract

This paper shows the Forecasting electricity demand for future years is an essential step in resource planning. A common approach is for the system operator to predict future demand from the estimates of individual distribution companies. The results demonstrate that the proposed method is efficient for assigning Typical Load Profile (TLP) to the consumers. Clustering the huge amount of data k-means clustering and extended k-means clustering algorithms are used. Moreover, the finding shows that the energy consumption can be clustered not only based on the load pattern but also load value. This paper shows work on extended k-means clustering algorithm which helpful for clustering same type of customers into one cluster.

INTRODUCTION

Electricity load forecasting has been an important risk management and planning tool for electric utilities ever since the conception of forecasting. Load forecasting is necessary for economic generation of power. Load serving entities use load forecasts for system security, to schedule generator maintenance, to make long-term investments in generation, and to plan the most cost-effective merit order dispatch. Over the last decade, as electricity markets have deregulated, the importance of load forecast accuracy has become even more evident. Without an optimal load forecast, utilities are subject to the risk of over- or under- purchasing in the day-ahead market. While an entity can buy or sell power in the real time market to correct for forecast inaccuracy, it comes at the expense of higher real time prices. The aim is to classify the load pattern of different types of customers. Conducting load pattern analysis is an important task in obtaining typical load profiles (TLPs) of customers and grouping them into classes according to their load characteristics. When using clustering techniques to obtain the load patterns of electricity customers, choosing a suitable clustering algorithm and determining an appropriate cluster number are always important and difficult issues. Even if the customer information needed in the classification is correct, some of the customers can simply have such an irregular behaviour pattern that they do not fit in any of the predefined customer class load profiles. The predefined customer class load profiles also include some inaccuracy due to geographical generalization. The most widespread customer class load profiles are created to model the average Finnish electricity consumption. They do not take into account the regional differences in electricity consumption, which originate from different climate conditions and socio-economic factors. The objective is utilization of electricity, developing Tariff on different types of electricity customers, Selection of generators. In the approach, all load curves of customers are first clustered with the clustering algorithms under a serial given number of clusters. When we are using clustering techniques to obtain the load patterns of electricity customers, choosing a suitable clustering algorithm and determining an appropriate cluster number are always important and difficult issues.

Many methods or techniques for clustering load curves have been proposed in the literature. Some clustering methods are: k-means (J. A. Hartigan et al, 1979), (D. Gerbec et al,2005), modified follow-the-leader (G. Chicco et al, 2006) ,(G. Chicco et al,2003) average and Ward hierarchical methods (N. M. Kohan et al, 2008), fuzzy c-means (FCM) (D. Gerbec et al,2005), statistic-fuzzy technique [14], the self-organizing map (SOM) (D. Gerbec et al,2005), [15], support vector machines (SVM) (G. Chicco et al, 2009), and extreme learning machine (A. H. Nizar et al,2008). Some hybrid techniques (S. C. Cerchiari et al, 2006) have been proposed to improve the clustering effect.

The purpose of this paper is to generate typical load profile and classify electric load using k-means clustering and modified k-means clustering. Researchers had worked on to decide which clustering algorithm is best for classification of load pattern analysis for different types of electricity customers. They compared most of the clustering algorithms.

TYPICAL LOAD PROFILE GENERATION

The classification of different types of electricity customers are achieved by applying clustering techniques, figure 1. Shows the flow chart of classification and load profile generation of large electricity customers which include the following basic steps:

Data Selection

The power consumption data of customers can be recorded by an automatic meter reading system with time periods in steps of 15 min, 30 min, or 1 h. The daily chronological load curves for each individual customer are determined for each study period (month, season, and year). Data contains the different types (residential, commercial, industrial, street light etc) of customer

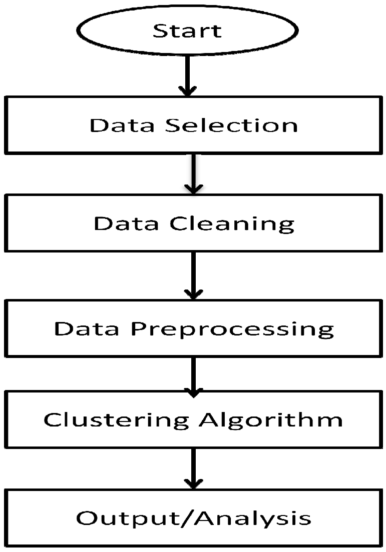

Figure 1 shows flow chart of typical load profile generation of electricity customers.

Data Cleaning

The load curves of each customer are examined for normality, in order to modify or delete the values that are obviously wrong (noise suppression). For example, we remove those daily load curves with 0 MW values and unreasonable load curves for known reasons (such as network failure or meter error).

Data Pre-processing

Clustering load curves are based on the shape of a load curve but not by absolute MW values, so the data should be normalized. Normalization is particularly useful for classification algorithms involving distance measurements. The different methods for data normalization are: min-max normalization, z-score normalization, and normalization by decimal scaling. The data were normalized in the range of (0, 1) by using as the normalizing factor the peak value.

Clustering Algorithm

Various clustering algorithms are used to cluster the normalized load curves. Clustering algorithm forms the load pattern clusters in results.

Output: Customer Classification

Output contains the typical load profile of different type of customers. Each load pattern contains a certain number of customers; no “empty” patterns exist. The customer classes can be obtained according to the load patterns. The typical load pattern for each customer can then be generated by the load curves belonging to the same load pattern; each typical load pattern is a centroid curve of a cluster of load curves connected with a load pattern.

APPLICATION DESCRIPTION

Difficult to select number of clusters for huge amount of electricity data. In this work, we focus on the k-means clustering algorithm and extended k-means clustering algorithm. Figure 2 shows the system diagram of generating load profile of electricity customer.

Data cleaning routines work to "clean" the data by filling in missing values, smoothing noisy data, identifying or removing outliers, and resolving inconsistencies. If users believe the data are dirty, they are unlikely to trust the results of any data mining that has been applied to it. Also, dirty data can cause confusion for the mining procedure, resulting in unreliable output. But, they are not always robust. Following two method shows easy way of filling the missing values.

- Fill in the missing value manually: This approach is time-consuming and may not be feasible given a large data set with many missing values.
- Use the attribute average value to fill in the missing value: This approach is not time-consuming and feasible given a large data set with many missing values.

In Data preprocessing module normalization the given data. In this paper, the data is normalized by min-max normalization (J. Han et al, 2006). The data were normalized in the range of (0.0, 1.0) by using as the normalizing factor the peak value. The data preprocessing output stored into database and then given to clustering module for clustering process.

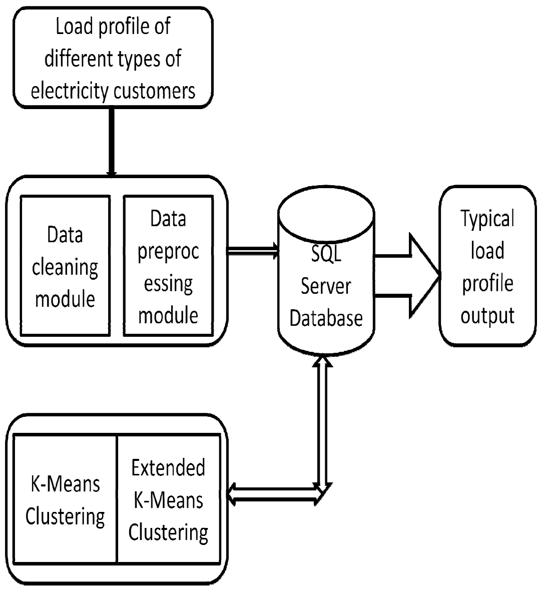

Figure 2 shows the System Diagram.

In clustering module two (k-means algorithm and extended k-means algorithm) algorithms used to form five clusters. Classical k-means clustering (J. A. Hartigan et al, 1979) groups a data set of $\mathbf{x}^{(n)}$ (n = 1, …, N) samples in k = 1, …, K clusters by means of an iterative procedure. A first guess is made for the K cluster centres $\mathbf{c}^{(k)}$ (usually chosen in a random fashion among the samples of the data set). The K centres classify the samples in the sense that the sample $\mathbf{x}^{(n)}$ belongs to cluster k if the distance $|| \mathbf{x}^{(n)} - \mathbf{c}^{(k)} ||$ is the minimum of all the K distances. The estimated centres are used to classify the samples into clusters (usually by Euclidean norm) and their values $\mathbf{c}^{(k)}$ are recalculated. The procedure is repeated until stabilization of the cluster centers. Clearly, the optimal number of clusters is not known a priori and the clustering quality depends on the value of K. The clusters formed by k-means clustering algorithm contain the two different types of customers in same cluster.

The extended k-means algorithm helpful for calculating consumption of particular type of customer or particular region. The clusters which are formed by k-means algorithm is given to extended k-means algorithm. The extended k-means algorithm find maximum number of same type of customers from cluster and transfer the different type of customer to their respective cluster. The extended k-means algorithm useful in load scheduling problem.

RESULT AND DISCUSSION

Table 1 shows the result of centroid values of clusters before clustering and after clustering. This centroid belongs to k-means clustering algorithm.

Table 1. Result of centriod values

No. of Clusters	Censored values before Clustering	Censored values after Clustering
1	(0.008, 0.007)	(0.008,0.008)
2	(0.001,0.012)	(0.000,0.020)
3	(1.000,1.000)	(0.250,1.000)
4	(0.001,0.016)	(0.020,0.018)
5	(0.480,0.500)	(1.000,0.930)

CONCLUSION

The paper presents the procedure how to determine the typical load profile based on the clustering methods and generation typical load profile of different types of customers. Data normalization is also a very important issue in clustering techniques.

Different normalization methods may cause different clustering results.

In this work, we classify a load pattern of different types of electricity customers. The extended k-means clustering helpful to form appropriate clusters. In k-means algorithm less number of cluster improve the classification rate.

REFERENCE

1. A. H. Nizar, Z. Y. Dong, and Y. Wang, Power utility nontechnical loss analysis with extreme learning machine method, IEEE Trans. Power Syst., vol. 23, no. 3, pp. 946–955, Aug. 2008.

2. G. Chicco and I. S. Ilie, Support vector clustering of electrical load pattern data, IEEE Trans. Power Syst., vol. 24, no. 3, pp. 1619–1628, Aug. 2009.

3. G. Chicco, R. Napoli, and F. Piglione, Comparisons among clustering techniques for electricity customer classification, IEEE Trans. Power Syst., vol. 21, no. 2, pp. 933–940, May 2006.

4. G. Chicco, R. Napoli, and F. Piglione, Application of clustering algorithms and self organising maps to classify electricity customers, in Proc. IEEE Bologna Power Tech Conf., Jun. 23–26, 2003, vol. 1.

5. N. M. Kohan, M. P. Moghaddam, S. M. Bidaki, and G. R. Yousefi, Comparison of modified k-means and hierarchical algorithms in customers load curves clustering for designing suitable tariffs in electricity market, in Proc. 43rd Int. Universities Power Engineering Conf., Padova, Italy, Sep. 1–4, 2008, pp. 1–5.

6. J. Han and M. Kamber, Data Mining: Concepts and Techniques, 2nd ed. San Francisco, CA: Morgan Kaufmann, 2006.

7. Tiefeng Zhang, Guangquan Zhang, Jie Lu, Xiaopu eng, and Wanchun Yang, A New Index and Classification Approach for Load Pattern Analysis of Large Electricity Customers, IEEE Trans. Power Syst., vol. 27, NO. 1, Feb. 2012

8. J. A. Hartigan and M. A. Wong, Algorithm AS 136: A k-means clustering algorithm, J. Roy. Statist. Soc. Series C (Appl. Statist.), vol. 28, no. 1, pp. 100–108, 1979.

9. F. Rodrigues, J. Duarte, V. Figueiredo, Z. Vale, and M. Cordeiro, A comparative

analysis of clustering algorithms applied to load profiling, Mach. Learn. Data Mining Pattern Recognit. Lecture Notes Comput. Sci., vol. 2734, no. 2003, pp. 73–85, 2003.

10. S. C. Cerchiari, A. Teurya, J. O. P. Pinto, G. Lambert-Torres, L. Sauer, and E. H. Zorzate, Data mining in distribution consumer database using rough sets and self-organizing maps, in Proc. IEEE/PES Power Systems Conf. Expo., Atlanta, GA, Oct. 29–Nov. 1 2006, pp. 38–43.

11. D. Hand, H. Manilla, and P. Smyth, Principles of Data Mining. Cambridge, MA: MIT Press, 2001.