International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-1 Issue-5, November 2011

Determination of Typical Load Profile of Consumers using Fuzzy C-Means Clustering Algorithm

Md. Jahangir Hossain, A. N. M. Enamul Kabir, Md Mostafizur Rahman, Borhan Kabir, Md Rafiqul Islam

Abstract— This paper reports the Typical Load Profile of different types of consumers of distribution feeder which is based on clustering methods. Among many clustering methods, fuzzy c-means has been examined for determination of representative clusters because fuzzy logic is conceptually easy to understand. It is a well known clustering algorithm for typical load profiles determination. The results demonstrate that the proposed method is efficient for assigning Typical Load Profile (TLP) to the consumers. Moreover, the finding shows that the energy consumption can be clustered not only based on the load pattern but also load value. The results demonstrate that the proposed method is efficient for assigning TLP to the consumers.

Key Words— Deregulation, Load Profile, Fuzzy Clustering Algorithm, Optimum Cluster, Probability, Neural network.

I. INTRODUCTION

Utilities would have a better trading and marketing strategies and the ability to design specific tariff options for the various classes of consumers in tune with real operation cost. Load profile of consumer makes more reliable and acceptable in this regards. Consumers can participate in the retail market and keep track of their actual power consumption and distributors could use such information for load management, distribution system planning and developed electricity tariffs, etc.The multiple participants of the electricity market need new business strategies for serving in competitive environments. So, accurate consumer's information is badly needed to fulfillment their electricity demand. Detailed knowledge on consumer's load consumption can facilitate distribution companies in determining specific tariff options for different type of consumers. Initially, the most efficient method of determining electricity consumption would be the direct monitoring. This can be achieved by installing time interval meters, quarter-hourly, half-hourly or hourly at each point of consumption. However, this approach is cost-prohibitive due to the equipment and processing costs. Furthermore, a significant amount of time would be needed to develop such a system. An alternative to this approach is by determining load profiles for consumers.

Manuscript Received October 18, 2011.

Md. Jahangir Hossain, Department of Electrical & Electronic Engineering, Khulna University of Engineering & Technology(KUET), Khulna, Bangladesh, Phone: +88041725375, Mobile No: +8801710862308 (e-mail: jhossain2k2@gmail.com)

A.N.M. Enamul Kabir, Department of Electrical & Electronic Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh, Mobile No.+8801714087197, (e-mail: anmenamulkabir@eee.kuet.ac.bd)

Md. Rafiqul Islam Department of Electrical & Electronic Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh, Mobile No.+8801714087330, (e-mail: rafiq043@yahoo.com)

Small eligible consumers have the possibility to change supplier even they don't have appropriate metering equipment. Therefore, the formation of a feasible and cost effective way to determine consumers' electricity consumption and deviations without installing new time interval meters is an important issue. A load profile based settlement can replace expensive installations of time-interval meters and it ensures a "fair" and accurate billing system and access to the retail market.

Typical Load Profiles (TLP) representing coherent groups of consumers Moreover, the load profiles can be also very helpful in dealing with load management, distribution system planning, and state estimation, distribution transformer loss of life evaluation or distribution system restoration.

There are two different approaches can be used for load profile based settlement; the area model and the category model [1]. The area model includes all those customers that are not metered on time interval basis within the geographic region covered by a network. On the other hand, the category model grouped customers with a similar load pattern into categories. Application of category model requires typical load profiles (TLP) representing coherent groups of consumers. Several papers present work regarding the establishment load profiles for a group of consumers [1], [2], [3]-[7], which are based on field measurements and can be divided into two groups. Typical feature for the first group is that TLPs are derived from load survey systems according to some predefined consumers groups [1], [2], [7]. The second group is obtained by identifying TLPs depending on the shape of the load curves [7], [9]. For this purpose, the use of pattern recognition methods is proposed. Limitations of the first TLP-determination approach are measurements performed over long time period and analyses how to define characteristic groups. Disadvantage of the second approach is formation of typical customer groups represented by TLP.

The aim of the paper is to analyze the Measured Load Profiles (MLP) obtained from the Distribution System Operators (DSO), determine the TLPs and their allocation to the particular group of eligible consumers. Among the popular method are fuzzy clustering, artificial neural network (ANN) and self-organizing map (SOM) as described in [5-7]. Typical load profile has great potential for further improvement in distribution system applications.

II. MODELING

A. Allocation of TPL

Determination and allocation Procedure of Typical Load Profiles for eligible consumers can be divided into three steps.

i) Preprocessing of the data obtained from the DSOs, consisting of arrangement of MLPs according to the season and type of the day, smoothing and normalization.

ii) Classification of the preprocessed MLPs using fuzzy c-means clustering algorithm.

iii) Allocation of the TLPs, obtained in second step, to the particular group of the eligible consumers using the probability neural network.

B. Procedure

The obtained data were preprocessed before clustering, using following steps:

i) Data for each MLP were collected over one month. Therefore they have to separate them regarding the day of the week.

ii) White noise was eliminated from the MLPs by wavelet multi–resolution technique in order to reduce the influence of coincidental electric devices on their shape.

iii) Load data has been recorded at different scales. Therefore data can have quite different ranges. Normalizing converts the MLPs to unit less variables using suitable normalizing factor. The peak power used as a normalizing factor. Moreover, peak power has a certain limitation and could give misleading results, because, the peak value is much higher than all other values of MLP. Thus, in these cases it is necessary to reduce or eliminate this peak value before normalization.

C. Flowchart

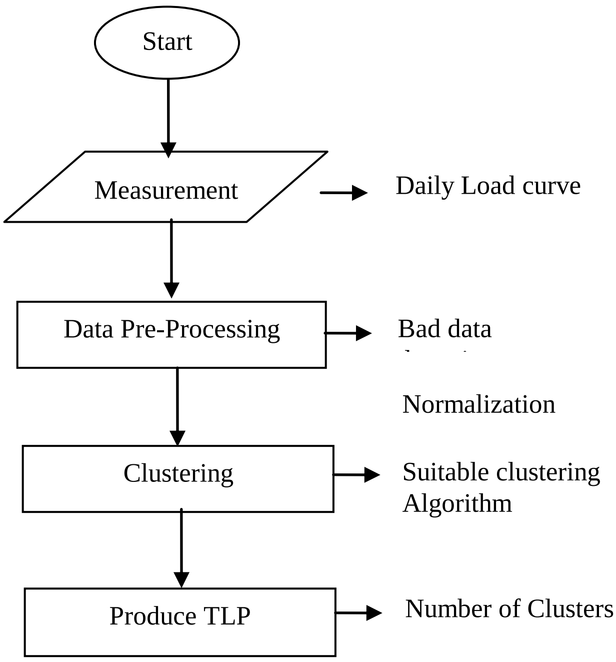

Fig.1: Flow chart of TLP determination

III. FUZZY C-MEANS CLUSTERING

Fuzzy C-means Clustering algorithm (FCM) is a method that is frequently used in pattern recognition. It has the advantage of giving good modeling results in many cases, although, it is not capable of specifying the number of clusters by itself [11]. The FCM can be applied to data that is quantitative (numerical), qualitative (categorical), or a combination of both. In pattern recognition, group of data is called a cluster. The load profile data can be arranged as a matrix where each row contains load profile data corresponding to observed variations of consumers load over time and the column corresponds to the number of evaluated consumers. On the other hand, fuzzy *c*-means (FCM) relaxes this requirement by allowing gradual memberships. In effect, it offers the opportunity to deal with the data that belong to more than one cluster at the same time. It assigns memberships to an object, which is inversely related to the relative distance of the object to cluster prototypes. By using FCM, each data point belongs to a cluster to some degree that is specified by a membership grade. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} (uij)^m \left\| X_i - Cj \right\|^2$$

Where, N denotes the number of load profile; C is the number of cluster, m is weighting parameter; in general m=2; uij is the degree of membership of xi in the cluster j; Xi is the profile of ith feeder of measured data, cj is jth cluster centre; || * || is any norm expressing the similarity between any measured data and the centre.

Assume that a set of N load profiles $X = \{x_1, x_2....x_j.....x_N\}$ to be clustered into C clusters (1<C<N). The steps in this algorithm are given below:

Step 01: Choose C and m, and initialize the partition matrix $U = [u_{ij}] matrix, U(0)$

Step 02: Calculate the cluster centers.

$$C_j = \frac{\sum_{i=1}^{N} uij^m . x_i}{\sum_{i=1}^{N} uij^m}$$

Step 03: Update the partition matrix for the kth step, U(k) as follows:

$$U_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|xi - cj\|^2}{\|xi - cj\|^2} \right)^{\frac{2}{m-1}}} \quad \text{For } 1<i<C$$

Step 04: If $\|u(k+1) - u(k)\| < \varepsilon$ then STOP; otherwise return to *step 02.*

Typically, *K* is chosen to be one; every iteration the updated value of *vij* depends only on the similarity between the object *xi* and *cj*. The data is partitioned into fuzzy subsets, so the objects on the boundaries between several classes are not forced to fully belong to one group, but rather are assigned membership degrees between 0 and 1 indicating their partial membership. Thus, the load profile cluster centre is the mean of all data points in the same observation, weighted by their degree of belonging to the cluster [3]. Updating of the means proceeds exactly the same way as in the case of the FCM.

IV. DETERMINATION OF OPTIMUM CLUSTER

Since clustering algorithms define clusters that are not known a priori, the final partition of the data requires some kind of evaluation. Cluster validity is the term to describe the procedure of this evaluation. One of the most important issues in clustering is to decide the optimal number of clusters that fits a data set. Using a validation index can solve this. Many different indices have been proposed [12]. In this study, some widely used indices are illustrated in Table 1.

Table 1: Clustering Validity Indices

Functional	Symbol	Authors
Nonfuzziness index	NFI	Roubens(1978)
Partition Coefficient	F	Bezdek(1981)
Partition Entropy	H	Bezdek(1981)
Minimum Hard endency	Min HT	Rivera(1990)
Mean Hard Tendency	Mean HT	Rivera(1990)
Separation Index	S	Xie-Beni(1991)

V. RESULTS AND DISCUSSION

The aim of the paper is to determine the typical load profile of consumers using fuzzy c-means clustering algorithm. The FCM clustering was implemented on MATLAB, where the inputs of the normalized value of hourly data of different feeders in matrix form and obtained the results of FCM clustering. The clustering algorithm for different number of clusters has been examined and the MATLAB program was run for several times with different iteration count. Finally the FCM clustering results was plotted against time in hour. The following figures (Fig.2- Fig.7) show the TLP of different feeders using FCM clustering.

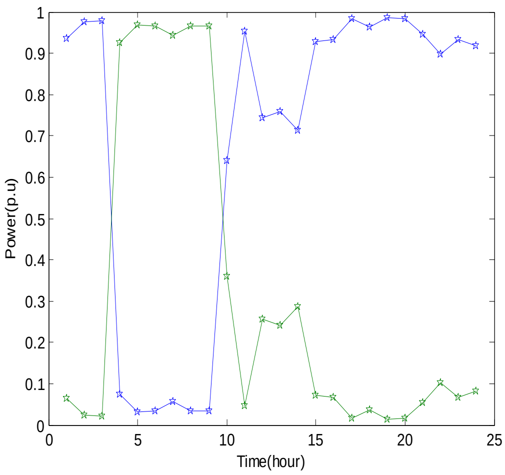

Fig 2:FCM clusrering result for cluster=2

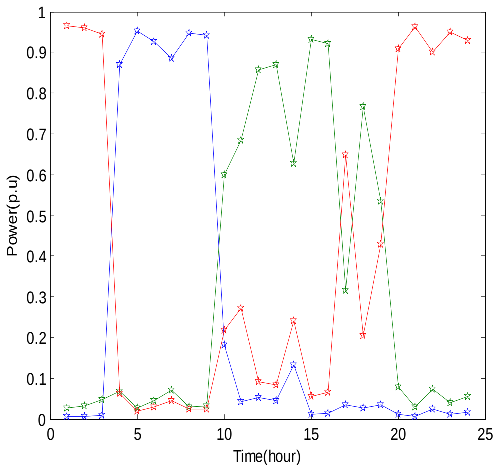

Fig 3:FCM clusrering result for cluster=3

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

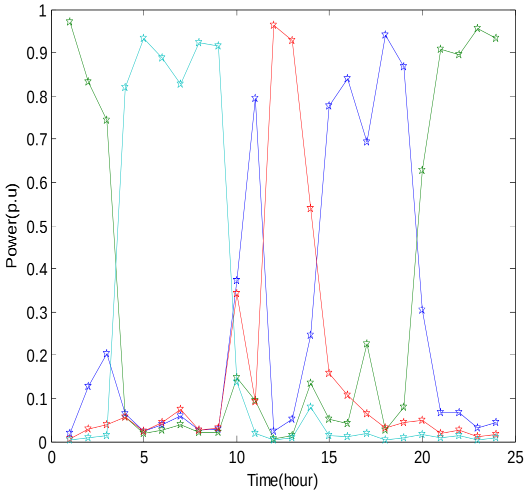

Fig 4:FCM clusrering result for cluster=4

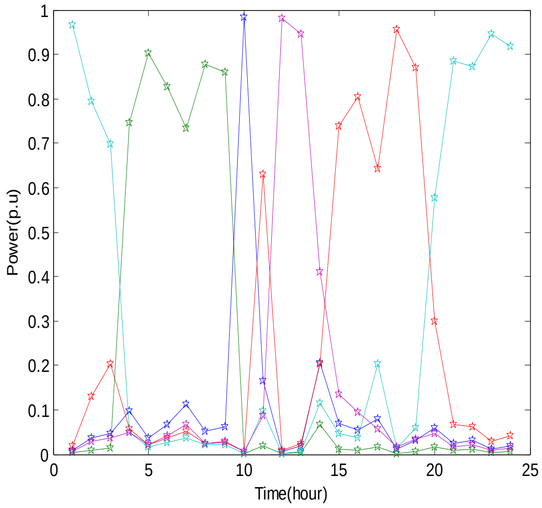

Fig 5:FCM clusrering result for cluster=5

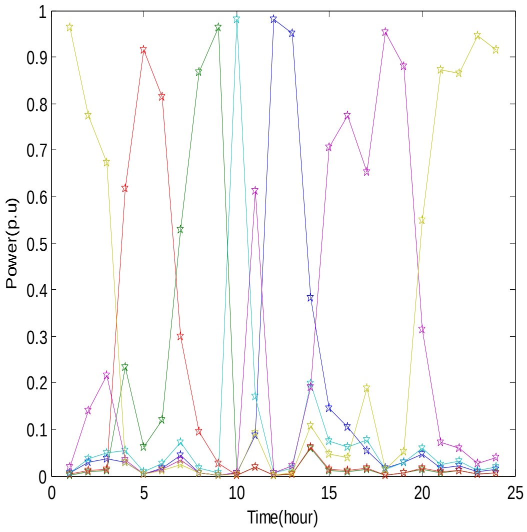

Fig 6:FCM clusrering result for cluster=6

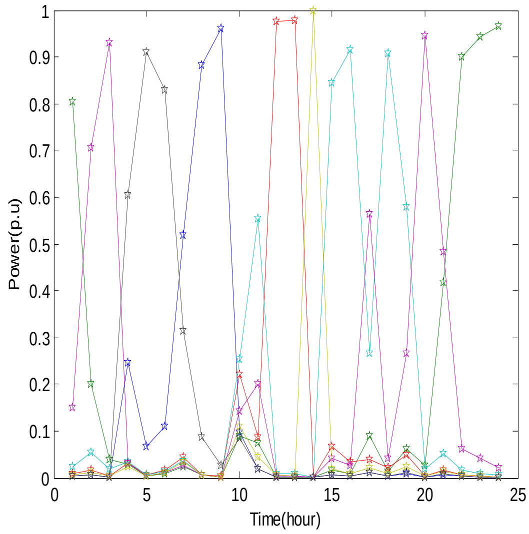

Fig 7:FCM clusrering result for cluster=7

Also to determine the optimum number of clusters the different cluster validity is examined whose values are listed in Table 2 for different number of clusters.

Table 2: Clustering validity indices for different number of clusters

	F	H	NFI	XB
C=2	0.7402	0.4062	0.4841	0.0433
C=3	0.7020	0.4821	0.5501	0.0343
C=4	0.6523	0.5326	0.4909	0.0291
C=5	0.6314	0.5917	0.4686	0.0343
C=6	0.5812	0.6819	0.4285	0.0339
C=7	0.5214	0.8123	0.3829	0.0338

The data consists of daily load consumption and measured for every one hour interval from midnight 12 am to next mid night 12am that gives 24 values for each feeder. The aim of clustering is to discover the natural group of the data; the algorithm was run several times. Based on the number of main consumers connected to the feeders involved, in this paper eight is the maximum number of cluster. Therefore the clustering process was repeated from c=2 until c=7. Cluster validity index is calculated at each value of c and this is shown in Table 2.

To choose the optimal number of cluster, F, NFI, MinHT and MeanHT should be maximum while H and S minimum. From Table 2, it seems that this can be either c=2, 3 or 4. However, it has been proven that the values of F and H always maximum and minimum respectively at c=2, thus these two indices are not suitable to determine the best cluster number [12].

Furthermore, in this case it is not appropriate to just have two clusters since there is prior knowledge that the feeders supplied more than two different types of consumers. Therefore, in this case, XB index is more favorable since four clusters are more appropriate for electricity feeders. This decision is taken in line with the fact that electricity consumers are normally categorized into residential, commercial, street lighting and industrial even though each of these major categories can be further subdivided. Accordingly, in this case, the feeders' data is clustered into 4 clusters. This is the optimum number of cluster.

Comparing the pattern of the TLP to the specific type of consumer will help to visualize the situation if it fits to any type of consumer. From the utility database, it is known that the feeders are connected to the following type of consumer:

- Domestic
- Commercial
- Small Scale Industrial
- Mixed Load (Domestic, Commercial and Small Scale Industrial)
- Mixed Load (Domestic and Commercial)
- Mixed Load (Commercial and Industrial)

VI. CONCLUSION

This paper presents the procedure to determine typical load profile of different types of consumers of distribution feeders which is based on clustering methods. The clustering algorithm for different number of clusters has been examined by MATLAB. The results of the fuzzy c-means clustering algorithm showed that measured load profiles could be classified into 7 well-separated clusters. The final results of classification showed that 4 clusters could be nominated as typical, while the other clusters can be rejected. Typical load profile is obtained by averaging the number of patterns in each cluster. Furthermore, the findings show that the energy consumption can be clustered not only based on the load pattern but also load value. Typical load profiles of different tariff of consumed energy pricing can take major impact on metering and billing systems. In conclusion, typical load profiles established in this paper has great potential for further analysis in distribution system applications.

REFERENCES

1. System Tariff Issues Working Group, "Metering, Load Profiles and Settlement in Deregulated Markets," Union of Electricity Industry 2000-220-0004, 2000.
2. "Load Profiling: An Innovate Alternative to Direct Metering When Implementing Customer Choice in the Electric Utility Industry," AEIC Load Research Committee July 1998.
3. C.S Chen, J.C Hwang, C.W Huang, Y.M Tzeng, and M.Y Chu, "Determination of Customer Load Characteristics by Load Survey System at Taipower," IEEE Transaction on Power Delivery, vol. Vol.11, 1996.
4. C.S Chen, J.C Hwang, and C.W Huang, "Application of Load Survey Systems to Proper Tariff Design," IEEE Transaction in Power System, vol. Vol.12, 1997.
5. K.L Lo and Zuhaina Zakaria, "Electricity Consumer Classification Using Artificial Intelligence," presented at University Power Engineering Conference, Bristol, U.K,vol 1, pp. 443-447, 2004.
6. D Gerbec, S Gasperic, and F Gubina, "Comparison of Different Classification Methods for the Consumers' Load Profile Determination," presented at 17th International Conference on Electricity Distribution, CIRED, Barcelona,vol. Session 6, pp. 2003.
7. Mihai Gavrilas, Viovel Calin Sfintes, and Marius Nelu Filimon, "Identifying Typical Load Profiles Using Neural Fuzzy Model," presented at IEEE Power Engineering Society Transmission and Distribution Conference, vol. 1, pp. 421-426, 2001.
8. Miin-Shen Yang, "A Survey of Fuzzy Clustering," Math. Compute. Modeling, vol. 18, pp. 1-6, 1993.
9. S.V Allera and A.G Horsburgh, "Load Profling for Energy Trading and Settlements in the Uk Electricity Markets," presented at DistribuTECH Europe DA/DSM Conference, London,vol, pp. 1998.
10. R Bellman, R Kalaba, and L Zadeh, "Abstraction and Pattern Classification," Journal of Math Analysis and Applications, vol. 13, pp. 1-7, 1966.
11. MohammadAlata, Mohammad Molhim, and Abdullah Ramini, "Optimizing of Fuzzy C-Means Clustering Algorithm Using GA" World Academy of Science Engineering & Technology,2008.
12. Hyun-Sook Rhee and Kyung-WhanOh, "A Validity Measure for Fuzzy Clustering and Its Use in Selecting Optimal Number of Clusters," presented at 5th IEEE International Conference on Fuzzy Systems, Seoul, Korea,vol 2, pp. 10201025, 1996.

AUTHORS PROFILE

Md. Jahangir Hossain is a Postgraduate student at Faculty of Electrical & Electronic Engineering in Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh. He received his B.Sc. Engineering degree in Electrical and Electronic Engineering from Khulna University of Engineering & Technology (2007). He is working as an Assistant Manager (Assistant Engineer) in West Zone Power Distribution Company Limited, Khulna Bangladesh. He is a member of the Institution of Engineers, Bangladesh (IEB).

A.N.M. Enamul Kabir, received the B. Sc. Engineering degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, He is working as An Associate Professor in Khulna university of Engineering & Technology. He is a member of The Institution of Engineers, Bangladesh (IEB).

Dr. Md. Rafiqul Islam received the B. Sc. Engineering degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 1998, the M. Sc. Engineering degree in Electrical and Electronic Engineering from Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh, in 2006 and the Ph.D. degree in Semiconductor device growth, characterization and fabrication from Fukui University, Fukui, Japan, in 2010. He is working as An Associate Professor in Khulna university of Engineering & Technology.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication