4th REGIONAL CONFERENCE ON
ELECTRICITY DISTRIBUTION
Tehran - Jan. 2016
چهارمین کنفرانس منطقه‌ای سیرد
تهران، دی ماه ۱۳۹۴

کد مقاله:	03140	وضعیت مقاله:	پذیرش - پوستر
عنوان مقاله:	تخمین بار پست‌های توزیع از طریق جایابی دستگاه‌های اندازه‌گیری با استفاده از روش خوشه‌بندی میانگین-سی فازی		
محورهاي علمي مقاله:	بهره برداری، کنترل و حفاظت شامل: بهره برداری - کنترل - حفاظت - برنامه‌ریزی و توسعه شبکه شامل: مدیریت سرمایه، استراتژی‌های نگهداری - توسعه شبکه - برنامه‌ریزی شبکه‌های توزیع - روش‌ها و ابزارها	تاریخ ارسال:	29/08/1394
مولف مسئول:	هومن فیروزی	تاریخ اصلاح:	04/10/1394
مولف 2:	حبیب رجبی مشهدی	پست الکترونیك مولف 2:	h_mashhadi@um.ac.ir
مولف 3:	محمد عظیم آرمین	پست الکترونیك مولف 3:	az.armin@gmail.com
مولف 4:		پست الکترونیك مولف 4:	
مولف 5:		پست الکترونیك مولف 5:	
تماس گیرنده:	هومن فیروزی	ارائه کننده:	هومن فیروزی
کلمات کليدي:	تخمین بار، خوشه بندی، میانگین سی فازی، شبکه توزیع، پروفایل بار		
چکیده مقاله:	اطلاع از بار پست‌های توزیع به‌منظور توسعه، بهره‌برداری و کنترل شبکه، همواره از اهمیت ویژه‌ای برخوردار بوده است. شبکه‌های توزیع به دلیل پیکربندی خاص، با چالش کمبود تجهیزات اندازه‌گیری مواجه هستند. تخمین بار پست‌های توزیع به‌عنوان راهکاری برای مقابله با این مشکل پیشنهادشده است. در این مقاله یک روش جدید برای تخمین بار پست‌های توزیع بیست کیلوولت ارائه می‌شود. ابتدا از خوشه‌بندی میانگین-سی فازی در جهت تعیین نقاط مناسب برای نصب ثبات‌های الکتریکی استفاده‌شده، سپس با قرار دادن ثبات‌ها بر روی ترانسفورماتورهای توزیع برای یک دوره مشخص، پروفایل بار ساعتی آن‌ها به‌دست‌آمده است. در مرحله بعد با داشتن پروفایل بار ساعتی قرائت‌شده از دو ترانسفورماتور نماینده، دو پروفایل بار نمونه با استفاده از یک روش خطی استخراج‌شده است. در مرحله آخر، پروفایل بار کلیه ترانسفورماتورها را با استفاده از این دو پروفایل بار نمونه تخمین زده می‌شود. در این مطالعه تعداد تجهیزات اندازه‌گیری در دسترس حداقل فرض شده است. مطالعه مذکور بر روی یک شبکه 33 باسه استاندارد انجام گرفته است.		
فايل مقاله:	FullPaper.doc FullPaper.pdf		

4th REGIONAL CONFERENCE ON ELECTRICITY DISTRIBUTION Tehran 12 & 13 Jan. 2016
چهارمین کنفرانس منطقه‌ای سیرد تهران، ۲۲ و ۲۳ دی ماه ۱۳۹۴

Distribution Substations Load Estimation with Minimal Meter Placement through a Fuzzy C-Means Clustering Approach

Hooman Firoozi, Habib Rajabi Mashhadi, M. Azim Armin
Ferdowsi University of Mashhad
Mashhad, Iran
hoomanfiroozi@gmail.com , h_mashhadi@um.ac.ir , az.armin@gmail.com

Abstract - The accurate information about distribution substation's load plays an important role in network expansion, operation and control. Distribution networks are usually confronted with lack of measurement devices due to their special configurations. Load estimation is suggested as one of the ways dealing with this problem. This paper proposes a new method of load estimation related to 20kv distribution substations. In the first step, fuzzy c-means clustering (FCM) is employed in order to determine the most appropriate places of meters to install. Then, hourly loads in a specific period is obtained using measurement devices installed on distribution substations. In the next step, the load profiles of representative substations are used to construct pattern load profiles (PLPs). Finally, load profiles of all substations are estimated from extracted PLPs. The paper assumes the minimum number of measurement devices. In addition, the analyses are performed for a 33-bus standard system in order to approve method's transparent application.

Keywords– Load profile, clustering, fuzzy c-means, meter placement, distribution system

1- Introduction

For technical or operational issues measurements have not been placed on every node in a distribution system. Although with sufficient number of measurement, data can be noisy, corrupted or missed. Hence, unavailable data have to be estimated to provide a real time or detail view of system. Unknown parameters can be estimated finding a relation between measured and unknown parameters. At most of the time, estimation error cannot be eliminated but it can be minimizing by measuring significant and sufficient data.

In literature estimation of customer’s load profile has been practiced [1]. Various clustering technique is very common in customer profile estimation. Hierarchical clustering, k-means clustering and fuzzy c-mean are most popular [2]. Although, Time series and neural network method are used but mostly for forecasting load behavior [3]. Using fuzzy numbers and fuzzy clustering is one of the most famous method that have acceptable errors but this error is dependent on the customer behavior and method has to be set for any particular data [4]. The clustering method is widely studied and categorize in literature [5].

In this paper a clustering method has been presented based on the fuzzy c-mean algorithm to estimate the load profile for each distribution substation. According to local distribution company’s policy in Iran, two different customer type (residential and commercial) has been considered. Unlike similar research in this field, each substation considered to feed mixture of residential or commercial customer.

This paper has been organized as follows: FCM Algorithm and network data have been presented in section 2. Main contribution has been presented in section 3. Simulation and the results is presented in section 4. Finally, this paper concluded in section 5.

2- Clustering Using FCM Method

2-1- Fuzzy C-Means Algorithm

The purpose of clustering is to identify groupings of a data set to show system's behavior briefly. In this procedure, data will be grouped to a specific number of clusters in a way that the members of each cluster have the maximum similarity to that cluster and minimum similarity to other clusters [6].

Many clustering methods have been introduced till now [7]. One of the most useful clustering methods is Fuzzy c-means. FCM is a clustering technique which groups data into *c* clusters with every data point in the dataset belonging to every cluster to a certain degree. Today, FCM algorithm has many usages in issues such as feature analysis, pattern recognition, image processing and classifier design [8, 9]. Although FCM method has some disadvantages like sensitivity to noise, heavy computational burden and falling in local minima somehow, there are some brilliant benefits such as being supervision and imminent convergence which assure researchers to utilize it.

The FCM procedure is actually a minimization problem which is based on the following objective function [10]:

$$J_m = \sum_{i=1}^{n} \sum_{j=1}^{N_c} u_{ij}^m . d\left(x_i, c_j\right) \tag{1}$$

Where
x – is the features vector
c – is the center of ith cluster
$d(x_i,c_j)$ – is the distance between feature vector x_k and the center of cluster
u_{ij} – is the degree of membership of vector x_k to cluster i
N_c – number of clusters
n – number of data vectors
m – level of fuzzy overlap between clusters

FCM algorithm has the following steps:

Step 1) Initializing the value of m, n and c. Usually, number of the clusters depending on the work is determined, otherwise it is necessary to identify the best value of it. Besides, the reasonable value of m obtained $m=2$ in [11].

Step 2) Calculating the center of clusters.

$$c_j = \frac{\sum_{j=1}^{n} (u_{ij})^m . x_i}{\sum_{j=1}^{n} (u_{ij})^m} \tag{2}$$

Step 3) Calculating the membership matrix from clusters' center.

$$u_{ij} = \left(\sum_{k=1}^{c} \frac{d(x_i, c_i)^{\frac{1}{m-1}}}{d(x_i, c_k)} \right)^{-1} \tag{3}$$

Step 4) If $\left| U^{(k+1)} - U^{(k)} \right| \le \varepsilon$ the algorithm will be stopped; otherwise return to step 2.

2-2- Distribution Substations Clustering

The case study examined in this research (shown in fig. 1) is a 33-bus radial distribution system. It is assumed there is a MV/LV (20kv/0.4kv) distribution transformer at each node of the network which is connected to specific amount of customers. There are just two class of consumers: residential and commercial.

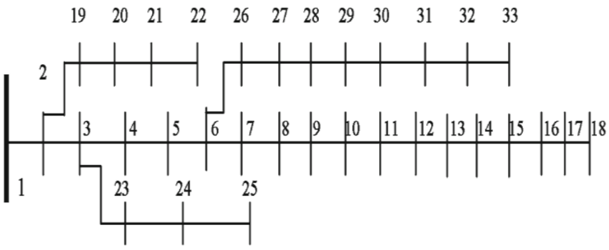

Fig. 1. 33-bus radial network

In order to cluster a set of data, some features are required. To obtain a desirable result of clustering, a proper choice of features is necessary. Here, the only available information of transformers is the number of users of each consumption class, so these features could be chosen to use in FCM as input data. It has been realized from the results of many trials, utilizing the number of each class of consumers does not lead to a reasonable output but, percent of consumers of each class is more suitable and seems to separate data properly. Accordingly, values that should be adopted from transformers data obtain through the following formula:

$$F_{i,j} = \frac{n_i}{\sum_{i=1}^{cc} n_i} \tag{4}$$

Where
$F_{i,j}$ – is the feature i of transformer j
cc – number of consumption classes
n_i – number of users of class i

The proposed methodology in this paper suggests the number of clusters equal to number of consumption classes, and that is for the sake of estimation method which is mentioned in the following sections. Thus, the number of clusters evolved for this study is 2. It means that, if there were four classes of consumers (e.g. residential, commercial, industrial and agricultural) the number of clusters should be adopted 4. Then, FCM algorithm have been run in 100 iterations, values of minimum improvement in objective function between two consecutive iterations and the level of fuzzy overlap between clusters are 10^{-5} and 2, respectively. The results of FCM convergence and transformers grouping objective function are shown in Fig. 2 and Table 1, respectively.

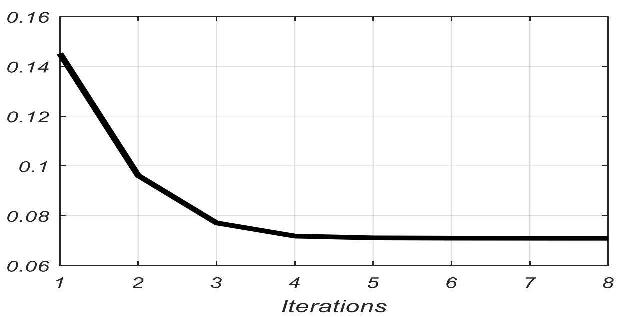

Fig. 2. FCM Algorithm Convergence

Table 1 Clustering Result

Cluster No.	Substation No.
1	2,4,6,8,10,17,18,19,22,25,28,31
2	1,3,5,7,9,11,12,13,14,15,16,20, 21,23,24,26,27,29,30,32,33

Since, percent of variety of each class of users were adopted as clustering features and the number of consumption class of users is 2, clustering procedure changes to one-dimension problem. For instance, for a specific transformer, if the percent of residential users was *% a* then the percent of commercial users would be *% (1-a)*. Generally, in this method the dimension of problem reduces one time. Accordingly, to realize the result of grouping of transformers with a better sensation, transformers are shown by their number of residential and commercial users which is shown in Fig. 3. Note that the black points shown in Fig. 3 named "Candidates" are transformers which their features have the maximum similarity to the center of their clusters because the center of cluster is a paragon representing the characteristics of that cluster.

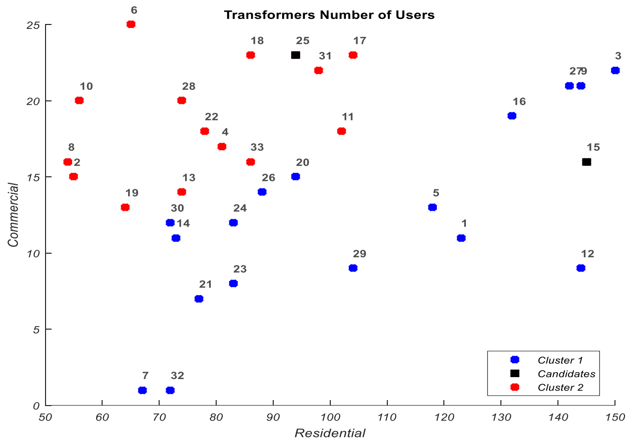

Fig. 3. Transformers Grouping

3- Pattern Load Profile Extraction

3-1- Meter Placement

As aforesaid, one of the assumptions undertaken in this study is lack of metering devices. In this network there are two classes of consumers therefore the minimum number of required measurements in order to attain a probable load estimation must be 2. Metering devices have the ability to record active power in 1 hour intervals. For the aim of obtaining an acceptable load pattern (LP) and then load estimation, a suitable meter placement matters. Means that it is obligatory to install measurement devices at substations which are more representing the characteristics of their cluster. By regarding this point, meter placement approach found the center of each cluster best location to install the meters as if the center of a cluster is a paragon of data in that cluster. However, the center of cluster does not overlap a member of dataset. To overcome this problem this method proposed the nearest member to the center of same cluster. Hence, in each cluster the Euclidian distance between transformers and their centers are calculated. Transformers which have the minimum Euclidian distance to the center of each cluster are elected as candidates for installation (Table 2 and Fig. 4).

Table 2. Candidate Transformers

Cluster No.	Candidate Substation
1	15
2	25

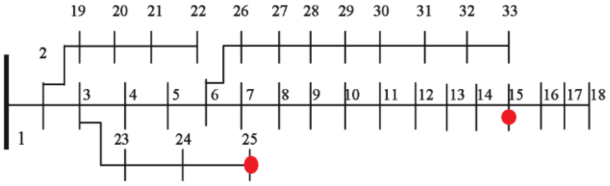
Fig. 4. Candidate Buses

3-2- Pattern Extraction

Measurement devices have been installed and recorded hourly active power for 30 days. A joint representation of average typical daily loads of candidate transformers in order to depicture the consumption pattern is shown in Fig. 5. The point is comprehended from Fig. 5 is high similarity of load profiles to a residential user consumption pattern. This is obvious because this case study is an urban feeder which the majority of its consumers are residential users.

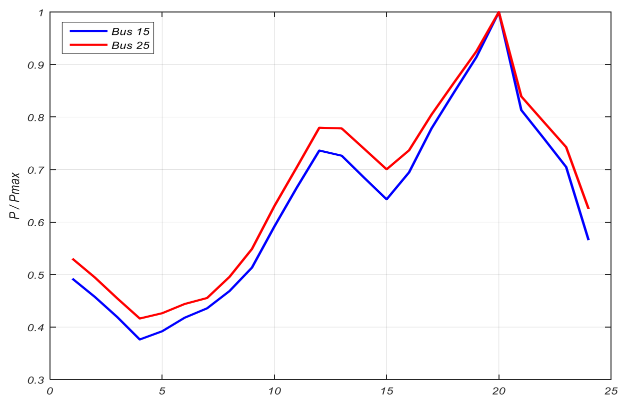

Fig. 5. Daily Average of Measured Loads

In this stage, a linear procedure for pattern load profile extraction is proposed. This method uses two measured profiles to obtain PLPs of residential and commercial consumers. For this aim, residential and commercial LP for each hour obtain through solving the following linear equations:

$$\begin{aligned} \alpha_{c_1}.P_{\alpha}^{t} + \beta_{c_1}.P_{\beta}^{t} &= P_{c_1}^{t} \\ \alpha_{c_2}.P_{\alpha}^{t} + \beta_{c_2}.P_{\beta}^{t} &= P_{c_2}^{t} \end{aligned} \quad (5)$$

Where

α_{c_i} and β_{c_i} - are number of residential and commercial users of candidate transformer of cluster *i*, respectively

$P^t_{c_i}$ - is measured load of candidate transformer *i* at hour *t*

P^t_α and P^t_β - are the LP of residential and commercial users at hour *t*, respectively

The extracted PLPs are shown in Fig. 6.

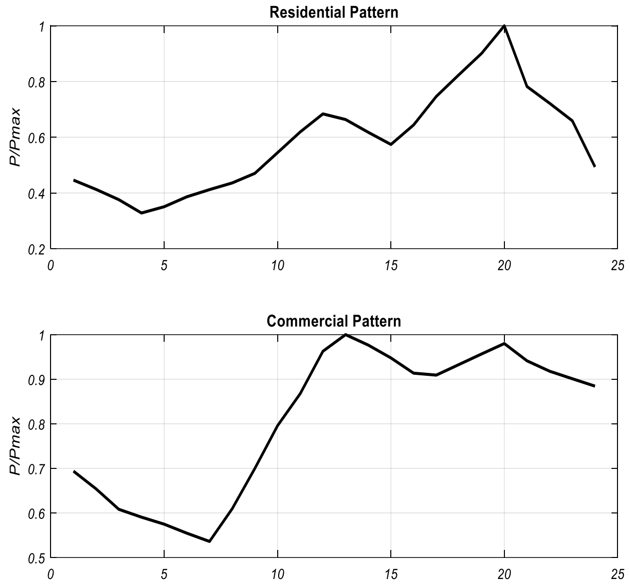

Fig. 6. Daily Average of PLPs

4- Load Estimation

In this section, the pattern loads of residential and commercial users are employed to estimate the hourly load of distribution substations. The actual values of substations' loads are available and will be used to calculate the accuracy of estimation.

The estimated loads of distribution substations are obtained via following equation:

$$\alpha_{t_i}.P^t_\alpha + \beta_{t_i}.P^t_\beta = P^t_{t_i} \quad (6)$$

Where

α_{t_i} and β_{t_i} - are number of residential and commercial users of transformer *i*, respectively

$P^t_{t_i}$ - is estimated load of transformer *i* at hour *t*

P^t_α and P^t_β - are the LP of residential and commercial

In order to show the results of load estimation and precision of method, the percent of maximum and mean error of estimated hourly loads are presented in Fig. 7.

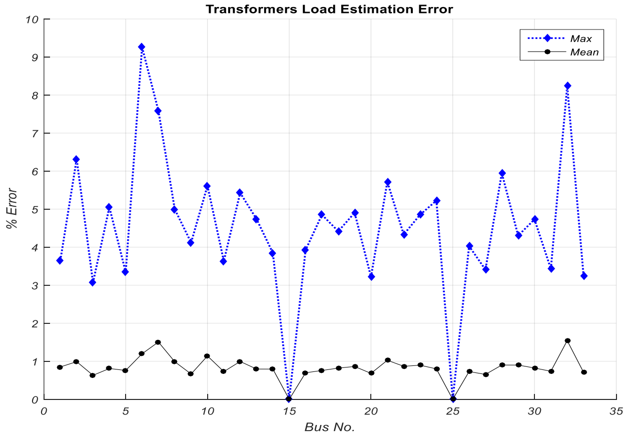

Fig. 7. Estimation Error for each substation

It was expected that value of estimated loads for bus 15 and 25 which have measurement devices contains no error. According to the results, the mean and the maximum value of hourly errors is near %1 and %5, respectively. In addition, the results of some random meter placement are presented in Table 3.

Table 3. Results of Random Meter Placement

Selected Buses	Max Errors
9 , 19	% 20.8
16 , 4	% 22
24 , 33	% 25
30 , 22	% 29
26 , 17	% 31.5

5- Conclusion

In this paper, a novel approach for the load estimation of distribution substations based on pattern load profiles was proposed. This approach divided into three step. In step 1, fuzzy c-means clustering method was employed to determine the meter placement truly. In step 2, a heuristic method was used to extract the pattern load profiles from the recorded values. In step 3, load profiles of all substations were calculated from patterns. Finally, results of load estimation properly affirmed that the accuracy of estimation with purposeful meter placement is higher than random meter placement.

References

[1] G. Chicco, R. Napoli, and F. Piglione, "Comparisons among clustering techniques for electricity customer classification," *Power Systems, IEEE Transactions on,* vol. 21, pp. 933-940, 2006.

[2] B. Akperi and P. Matthews, "Analysis of clustering techniques on load profiles for electrical distribution," in *Power System Technology (POWERCON), 2014 International Conference on*, 2014, pp. 1142-1149.

[3] M. Espinoza, C. Joye, R. Belmans, and B. De Moor, "Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1622-1630, 2005.

[4] F. Daneshvar, M. R. Haghifam, and M. S. Vojdani, "Clustering load distribution substation based on similarity of load curves using

statistic-fuzzy methods," in *Electrical Power Distribution Networks (EPDC), 2011 16th Conference on*, 2011, pp. 1-6.

[5] A. Mutanen, M. Ruska, S. Repo, and P. Jarventausta, "Customer Classification and Load Profiling Method for Distribution Systems," *Power Delivery, IEEE Transactions on,* vol. 26, pp. 1755-1763, 2011.

[6] T. Kwok, K. Smith, S. Lozano, and D. Taniar, "Parallel fuzzy c-means clustering for large data sets," in *Euro-Par 2002 Parallel Processing*, ed: Springer, 2002, pp. 365-374.

[7] F. Höppner, Fuzzy cluster analysis: methods for classification, data analysis and image recognition: John Wiley & Sons, 1999.

[8] M. C. Clark, L. O. Hall, D. B. Goldgof, L. P. Clarke, R. P. Velthuizen, and M. S. Silbiger, "MRI segmentation using fuzzy clustering techniques," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 13, pp. 730-742, 1994.

[9] Y. W. Lim and S. U. Lee, "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques," *Pattern recognition,* vol. 23, pp. 935-952, 1990.

[10] Z. Zakaria, K. L. Lo, and M. H. Sohod, "Application of Fuzzy Clustering to Determine Electricity Consumers' Load Profiles," in *Power and Energy Conference, 2006. PECon '06. IEEE International*, 2006, pp. 99-103.

[11] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *Fuzzy Systems, IEEE Transactions on,* vol. 3, pp. 370-379, 1995.