Time of Use Period Determination for Residential Customers in Peninsular Malaysia: A Case Study

Nurhidayah Khamis[1,a)] and C S Tan[2,b)]

[1]*College of Graduate Studies, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia.*
[2]*Institute of Energy Policy and Research, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia.*

[a)]Corresponding author: hidayahkhamis88@gmail.com
[b)]chingsin@uniten.edu.my

Abstract. Time of Use (TOU) is basically one of the demand response programs that encourage end-user customers to change their electricity usage in response with the changes in the price of electricity over a period of time with an incentive. Generally, Time-of-Use implementation is to reduce the system's maximum demand by transferring some of the demand into different hours. Time-of-Use also is a cost reflective electricity pricing scheme in which days are commonly split into two or three time periods such as peak, mid-peak and off-peak. The residential sector is expected to have the largest volume growth as there will be an increasing population, urbanization and rising living standards that can increase the number of households and the energy demand per household which allows the households and individuals to purchase more electrical appliances. This paper presented a new clustering method called Jenks Natural Breaks to determine the Time of Use period for the residential customers in Peninsular Malaysia. A comparison of K-Means clustering method and the proposed Jenks Natural Breaks method is presented in this paper. These two methods are carried out on the average of six actual residential customers load profile. In this paper, two-part periods (zones) TOU are used for analysis and discussions. The results indicate that the time of use Peak period using the K-Means clustering method is between 10.00am and 8.00pm while for a new proposed Jenks Natural Breaks method the time of use Peak period is between 9.00am and 8.00pm.

INTRODUCTION

In developed countries, Time-of-Use (TOU) electricity pricing is widespread adopted in Australia, America and Europe. Meanwhile in ASEAN, several countries such as Malaysia, Thailand, Singapore and Philippines have begun to implement TOU pricing to provide better tariff options to their customers while other ASEAN countries such as Brunei, Laos, Myanmar, Indonesia, Vietnam and Cambodia are still offering inclining block tariff to their customers[1]. For those countries that have a seasonal weather, time of use may vary based on their seasonal variation. At present, many utilities have offered TOU rates typically as default service to large Commercial and Industrial customers[2]. Introducing TOU for residential customers is tricky since it has a unique load profile. Applying TOU period for residential customers similar to the commercial and industrial TOU period should not be the same since the residential electricity consumption during day time is low while commercial and industrial are high. Generally, TOU period design must benefit both utilities and customers. The utility benefited by gaining reduced electricity generation cost while customers benefited by gaining reduced in their electricity bill[3]. One of the important aspects of Time-of-Use design is the determination of the period such as Peak and Off-peak hours of the day[4].

Generally, cluster analysis is used to group the hours that having a similar consumption of the day into clusters to be used in time of-use rate design. Clustering analysis has been used widely in many fields including Science, Medicine, and Engineering fields. There are numerous clustering techniques and the choice of the clustering techniques depends on the context and purpose of the analysis. There are also variety methods of clustering techniques that has been applied in power systems applications. A recent study conducted by [5] has introduced a new segmentation technique called Jenks Natural Breaks for the residential customers by segmenting customers load profile into two-part periods namely peak and off-peak periods. Most of the literature reviews from the previous studies used Hierarchical clustering, K-Means clustering, Fuzzy K-Means clustering and Self-Organizing Map (SOM) clustering [6-8]. A previous study by [9], has introduced and analyzed the difference between the clustering methods which includes Hierarchical clustering, K-Means clustering, Follow the leader, Fuzzy K-Means and Fuzzy Classification. A good clustering method will produce high quality clusters in which the intra-cluster similarity is high and the inter-class similarity is low. The quality of a clustering result also depends on both the similarity measures used by the method and its implementation and the quality of a clustering method is also measured by its ability to discover some or all of the hidden patterns. Self - Organizing Maps (SOM) and K-Means clustering method have been used in a previous study by [10] and [11] to identify the customers load patterns and to describe a characterization framework for the electricity customers. Two methods for performing electricity customer classification based on their electrical behavior were presented in their studies. The number of clusters shall be determined by the researchers as it cannot be directly fixed. The desired number of clusters may require successive executions of the algorithm by adjusting the threshold value. A two-stage recognition of load curves based on different clustering methods is described by [12] including K-Means clustering. Self-Organizing Maps (SOM) clustering method has been done by [13] based on the database measurements to achieve the classification of clustering and demand pattern of the customers. K-means is a commonly used clustering method with the simplest principle and fast speed of convergence[14]. Other than that, K-Means also has been used to analyze the large scale datasets[15]. Each of the clustering method has its own advantages and disadvantages. Table 1 below shows the advantages and disadvantages of each clustering method.

TABLE 1. Advantages and disadvantages of each clustering method

Clustering Method	Advantages	Disadvantages
Hierarchical	1. Hierarchical algorithm shows high quality 2. Hierarchical algorithm is good for small dataset 3. Easier to decide on the number of clusters by looking at the dendrogram 4. Easy to implement	1. Methods are not necessarily scalable for large datasets 2. The performance of Hierarchical Clustering is less than K-Means Clustering 3. Time complexity is not suitable for large datasets
K-Means	1. The performance of K-Means algorithm is better than Hierarchical Clustering 2. K-Means algorithm is good for large dataset	1. K-Means algorithm shows less quality 2. Difficult to predict the number of clusters (K-Value)
Fuzzy K-Means	1. Unlike k-means where data point must exclusively belong to one cluster center here data point is assigned membership to each cluster center as a result of which data point may belong to more than one cluster center	1. Long computational time 2. Sensitivity to the initial guess (speed, local minima) 3. Sensitivity to noise and One expects low (or even no) membership degree for outliers (noisy points).
Self-Organizing Maps (SOM)	1. Very simple to implement 2. Can be very effective for visualizing high D-spaces 3. Can incorporate new data quickly	1. The output space topology is predefined 2. Can converge to poor clustering depending on initialization and learning rate
Jenks Natural Breaks	1. Mapping values that are not evenly distributed on histogram	1. Class ranges are tailored to one data set, so difficult to compare maps for different data sets 2. Can be difficult to choose optimum number of classes, especially if data are evenly distributed

Jenks Natural Breaks algorithm is also another data clustering method designed to determine the best arrangement of values into different classes which involve iterative process. Historically, Jenks Natural Breaks algorithm was introduced in 1977 as a method for optimal data classification. The design of the algorithm is based primarily upon Fischer's Exact Optimization method that was developed in 1958 [16]. In the previous works, Jenks technique was specifically developed for the use in analysis of geographic data, and has emerged as a standard geographic classification algorithm, as evidenced by its selection as the default classification method in the industry-leading software package ArcGIS from Environmental Research Systems Institute (ESRI). Therefore, in this paper, a continuation study and works from [5] using Jenks Natural Breaks is introduced and presented to determine the two-part periods TOU from the residential customers load profile. A comparison of two-part periods TOU between K-Means Clustering method and Jenks Natural Breaks is presented.

METHODOLOGY

In this study, the Jenks Natural Breaks analysis and K-Means clustering method which are available in the Microsoft Office Excel Real Statistics Resource Pack Adds-In Tools are used and the TOU period is determined. Two classes of periods are chosen for the load profile segmentation which are Peak and Off-Peak periods. This is because the more classes utilized, more complex and time consuming the algorithm will be.

Six (6) load profiles in a billing cycle month were collected from the residential customers in Melaka that are installed with smart meters. The residential load profile was collated between 20th of August 2017 and 20th of September 2017 (31 days). The load profile reading has an interval of 30 minutes. Total dataset points for load profile clustering of 6 residential customers for 31 days are 8928 datasets. Two periods are chosen for the load profile clustering which determine the Peak and Off-Peak periods using the proposed Jenks Natural Breaks method and K-Means clustering method. The analysis was carried out on Weekdays which is Monday until Friday and Weekends which is Saturday and Sunday. The steps of the Jenks Natural Breaks method and also K-Means Clustering method are described below in order to determine the TOU period for residential customers.

K-Means Clustering

The clustering method for K-Means will groups the load profile data by determining a certain number of clusters and a center point for each cluster. After determining the center point of each cluster, each data set should be assigned to the nearest center point then a recalculation of the new center point will be done iteratively until the position of the center point is stable[7]. This method does not create a tree structure to describe the groupings of data, but rather creates a single level of clusters. K-means clustering uses the actual observations of objects or individual data and therefore is more suitable for clustering large amounts of data. The steps of the K-means algorithm are discussed as below:

Step 1: Initialization - Randomly K input vectors (data points) was choose to initialize the clusters.
Step 2: Searching - Find the cluster center that is closest for each input vector and assign that input vector to the corresponding cluster.
Step 3: Updation - The cluster centers will be update in each cluster using the mean (centroid) of the input vectors assigned to that cluster.
Step 4: Continuation - Repeat steps 2 and 3 until no more change in the value of the means.

Jenks Natural Breaks

The Jenks Natural Breaks method is to minimize each class average deviation from the class, while maximizing each class's deviation from the means of the other groups. In other words, the method seeks to reduce the variance within classes and maximize the variance between classes. This method requires an iterative process which means the calculation must be repeated using different breaks in the dataset to determine which set of breaks has the smallest in-class variance. The process is started by dividing the ordered data into groups. There are five steps that must be repeated until the goodness of variance fit (GVF) is maximized as follows:

Step 1: Calculate the sum of Squared Deviations Between Classes (SDBC)
Step 2: Calculate the sum of Squared Deviations from the Array Mean (SDAM)
Step 3: Subtract the SDBC from the SDAM (SDAM - SDBC). This equals the sum of the Squared Deviations from

the Class Means (SDCM)

Step 4: After inspecting each of the SDBC, a decision is made to move one unit from the class with the largest SDBC toward the class with the lowest SDBC

Step 5: New class deviations are then calculated, and the process is repeated until the sum of the within class deviations reaches a minimal value. Alternatively, all break combinations may be examined, SDCM calculated for each combination, and the combination with the lowest SDCM selected. Since all break combinations are examined, this guarantees that the one with the lowest SDCM is found. Finally, the GVF is calculated. GVF is defined as (SDAM - SDCM) / SDAM. GVF ranges from 0 (worst fit) to 1 (perfect fit). The method is achieved when the quantity GVF is maximized.

Figure 1 depicts the flowcharts for both K-Means Method and Jenks Natural Break Method.

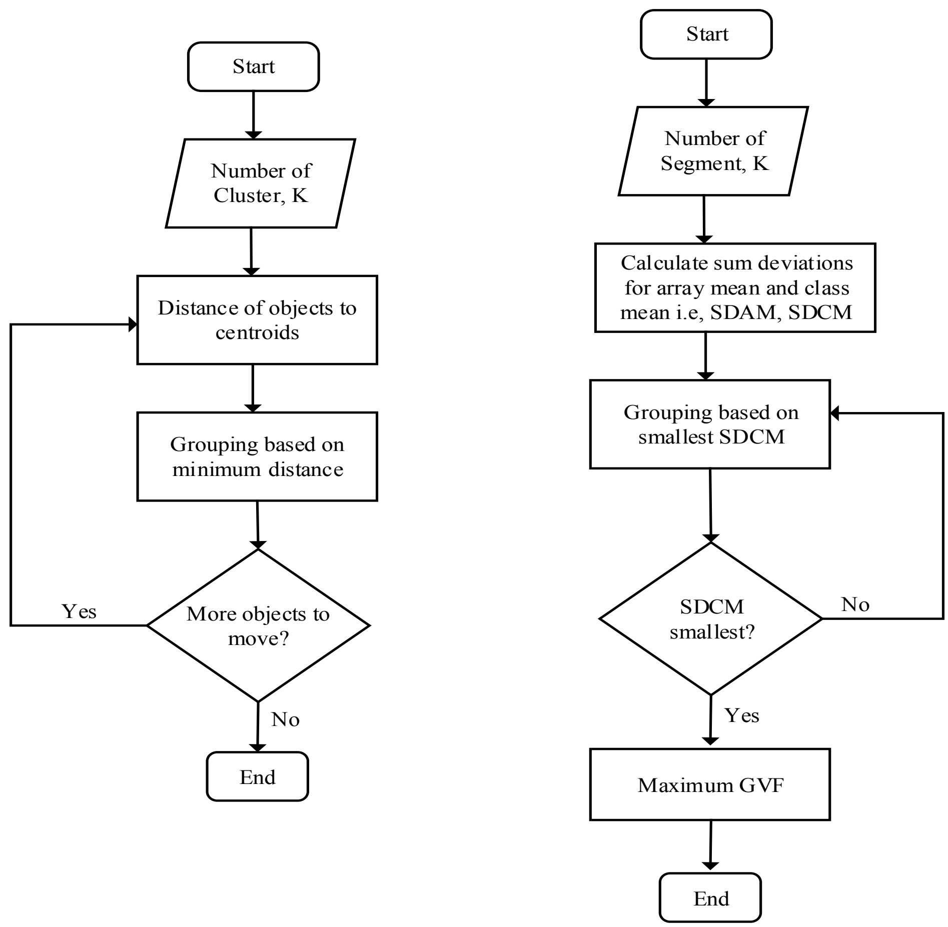

FIGURE 1. Flowchart of the K-Means (Left) and Jenks Natural Breaks (Right) clustering methods

RESULTS AND DISCUSSIONS

In this research study, peak period is defined as the system generation peak hours where the system demand is peaking hence the cost of generation is higher and vice versa. Based on the data analysis in Table 2, all the data elements in the range of 0.0445kWh to 0.7850kWh are considered as peak period. Meanwhile, for the off-peak period, the data elements are from 0.7855kWh the lower data to 4.5530kWh the upper data. The lower data in Jenks Natural Breaks is defined as a minimum data element while the upper data is defined as maximum data element. From the analysis, the six customers load profile gives the total squared deviation of 727.604 while the squared deviation of the input data is 1987.53. The Goodness Variance Fit (GVF) is 63.39% or 0.6339 which is near to 1 (good fit). This shows that the classification using Jenks Natural Breaks method is good.

TABLE 2. Jenks Natural Breaks Optimization Analysis for six actual residential customers

Period	Lower Data (kWh)	Upper Data (kWh)	Count
Peak	0.0445	0.7850	6793
Off-Peak	0.7855	4.5530	2135
GVF	727.604	1987.53	0.6339

Table 3 below shows the comparison of time-of-use period for peak and off-peak using Jenks Natural Breaks and K-Means clustering method. The results from the simulation for average of 6 customers shows the peak period is from 9.00am to 8.00pm while off-peak period is from 8.00pm to 9.00am. This TOU period is from Monday to Sunday. Meanwhile, for K-Means analysis results, it shows that the time of use peak and off-peak period are from 10.00am-8.00pm and 8.00pm-10.00am respectively which are from Monday to Sunday.

TABLE 3. Time-of-Use Peak and Off-Peak period using Jenks Natural Breaks method

Period	K-Means Clustering	Jenks Natural Breaks
Peak	10.00am-8.00pm (10 hours)	9.00am-8.00pm (11 hours)
Off-Peak	8.00pm-10.00am (14 hours)	8.00pm-9.00am (13 hours)

The results tabulated in Table 3 can be depicted as in Figure 2 for the Peak and Off-Peak clusters between K-Means clustering and Jenks Natural Breaks methods. Cluster 1 indicates the Peak period while Cluster 2 indicates Off-Peak period. Based on the data collated, the K-Means clustering method and Jenks Natural Breaks method show TOU time different of one hour in the Peak Period. This indicates that Jenks Natural Breaks method can be used to cluster the customer load profile that having similar consumption to determine the TOU Peak and Off-Peak periods.

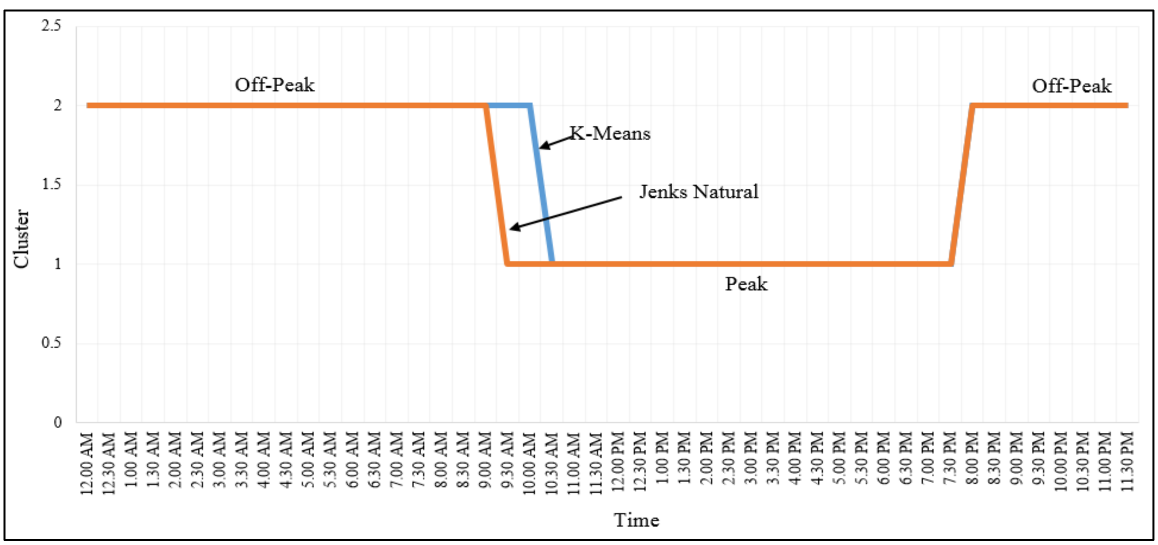

FIGURE 2. K-Means Clustering versus Jenks Natural Breaks

CONCLUSIONS

In this paper, Jenks Natural Breaks method was introduced to perform the Time of Use periods into two periods namely Peak and Off-Peak periods based on the energy consumption of residential customers. The results indicate that the time-of-use Peak period based on the Jenks Natural Breaks is between 9.00am and 8.00pm while the Off-Peak period is between 8.00pm and 9.00am. Meanwhile, K-means clustering method also were discussed in this paper in order to compare the time of use period with the proposed method. The results from the K-Means analysis shows that the Peak period is between 10.00am and 8.00pm and Off-Peak period is from 8.00pm and 10.00am. From this result, it can conclude that Jenks Natural Breaks method can be used to cluster the load profile data in order to determine the

Time of Use period for the residential customers. However, the impact of the TOU period is not discussed in this paper. Further analysis is required to determine the benefits of the TOU period and the rates offered to the customers.

ACKNOWLEDGMENTS

The authors are grateful and appreciative to Tenaga Nasional Berhad (TNB) for sharing few samples of load profile information for the study. This research work is supported and funded by Universiti Tenaga Nasional (Innovation and Research Management Centre, iRMC) internal grant - 10436494/B/2019061. Any of the opinions, findings and conclusions expressed in this paper are those of the authors views and do not necessarily reflect the views of the Tenaga Nasional Berhad or the Universiti Tenaga Nasional.

REFERENCES

1. M. H. Albadi and E. F. El-Saadany, "Demand response in electricity markets: An overview", IEEE Power Engineering Society General Meeting, PES, pp. 1–5 (2007).
2. Q. Zhang and J. Li, "Demand response in electricity markets: A review", 9th International Conference on the European Energy Market, EEM 12, (2012).
3. N. S. Hussin, M. P. Abdullah, A. I. M. Ali, M. Y. Hassan, and F. Hussin, "Residential electricity time of use (ToU) pricing for Malaysia", IEEE Conference on Energy Conversion (CENCON), pp. 1–5 (2014).
4. S. A. Dobrow and B. P. Lingaraj, "Design of time-of-use rate periods for a utility", Journal of Operations Management, Vol. 7, No. 3–4, pp. 25–43 (1988).
5. Khamis Nurhidayah; Tan Ching Sin; Goh Chin Hock, "Segmentation of Residential Customer Load Profile in Peninsular Malaysia using Jenks Natural Breaks", IEEE 7th International Conference on Power and Energy (PECon), pp. 128–131 (2018).
6. G. Chicco, R. Napoli, and F. Piglione, "Comparisons among clustering techniques for electricity customer classification"x IEEE Trans. Power Syst., Vol. 21, No. 2, pp. 933–940 (2006).
7. G. Chicco, R. Napoli, and F. Piglione, "Application of clustering algorithms and Self Organising Maps to classify electricity customers", IEEE Bologna PowerTech - Conference Proceedings, Vol. 1, pp. 373–379 (2003).
8. Jassar, K.K., Dhindsa, K.S, "Comparative Study and Performance Analysis of Clustering Algorithms", International Conference on ICT for Health (ICTHC), pp. 1–6, (2016).
9. I. Prahastono, D. King, and C. S. Ozveren, "A review of Electricity Load Profile Classification methods", 42nd International Universities Power Engineering Conference, pp. 1187–1191 (2007).
10. G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. Toader, "Load pattern-based classification of electricity customers", IEEE Trans. Power Syst., Vol. 19, No. 2, pp. 1232–1239 (2004).
11. V. Figueiredo, F. Rodrigues, Z. Vale, and J. B. Gouveia, "An electric energy consumer characterization framework based on data mining techniques", IEEE Trans. Power Syst., Vol. 20, No. 2, pp. 596–602 (2005).
12. G. J. Tsekouras, N. D. Hatziargyriou, and E. N. Dialynas, "Two-stage pattern recognition of load curves for classification of electricity customers", IEEE Trans. on Power Systems, Vol. 22, No. 3, pp. 1120–1128 (2007).
13. C. Senabre, S. V. Verdu, M. O. Garcia, F. J. G. Franco, and A. G. Marin, "Classification, Filtering, and Identification of Electrical Customer Load Patterns Through the Use of Self-Organizing Maps", IEEE Trans. on Power Systems, Vol. 21, No. 4, pp. 1672–1682 (2006).
14. G. Nourbakhsh, G. Eden, D. McVeigh, and A. Ghosh, "Chronological categorization and decomposition of customer loads", IEEE Trans. on Power Delivery, Vol. 27, No. 4, pp. 2270–2277 (2012).
15. J. Kwac, J. Flora, and R. Rajagopal, "Household energy consumption segmentation using hourly data" IEEE Trans. on Smart Grid, Vol. 5, No. 1, pp. 420–430 (2014).
16. W. D. Fisher, "On Grouping for Maximum Homogeneity", Journal of the America Statistical Association, Vol. 53, No. 284, pp. 789-798 (1958).