The Journal of CPRI,
Vol. 13, No. 2, June 2017 pp. 245-252

Customer segmentation based on electrical consumption in a VPP Environment: Technologies and Applications

Vijayalakshmi Anbazhagan*, Shanti Swarup K*, Haile- Selassie Rajamani** and Prashant Pillai**

Virtual Power Plant (VPP) is an important desired new comer to the electricity market. The objective of VPP is to supply green energy by integrating Renewable Energy Sources (RES) like wind and solar, etc., optimize the grid stability, enable customers to participate flexibly in the market operation and maximize the VPP profit. A VPP may consist of number of customers. Handling them as an individual is a difficult task to VPP operator, hence, electricity customer segmentation based on their demand becomes crucial. Nowadays, customers are becoming producers (called as prosumers), VPP helps to plan power transaction from prosumers and/or grid to reduce the loss and congestion, etc. In this paper, a new method has been proposed for electricity customer segmentation and the results are compared with K-means clustering method as well. Electricity customer segmentation plays a vital role in dispatching scheduled power, prioritize the customers in an emerging situation, tariff design, demand response, voltage regulation and planning power transaction from prosumer and/or generation, etc. Here, the segmentation has been used for tariff design.

Keywords: *Virtual power plant, Customer segmentation, Distribution generation, Clustering.*

1.0 INTRODUCTION

The ever-growing energy demand, forcing environmental issues and economic considerations have made a number of Distributed Generation (DG) using renewable resources such as wind energy, solar energy, biomass and hydro to come up. Many countries have set up incentives and policies to promote generation from RES, which results in a significant growth of Distributed Generations in the past decade. The rising number of DGs poses challenges for the power system. With the liberalization of the electricity market, DGs will have the opportunities to participate in electricity markets. The high risk of not meeting the scheduled demand makes the market participation of DG difficult. Also, there are some other integrating challenges like resource intermittency, voltage regulation, reactive power compensation, energy storage and cost benefit analysis [1&8]. A VPP enables the participation of large number of DGs in electricity markets in spite of these problems.

2.0 VIRTUAL POWER PLANT

VPP is an entity which has distributed generation, controllable loads, energy storage system and two way communication systems. Also, it is participating in market activities and encouraging DGs to contribute in market activities irrespective of grid connection and distance as shown in Figure 1

2.1 Definition of VPP

VPP can be defined as "A flexible characterization of a Distributed Energy Resources (DER) that

*Department of Electrical Engineering, IIT-M, Chennai-600036, India.
**Engineering and Informatics, University of Bradford, BD7 1DP, UK

can access and gain knowledge across all energy markets and can get benefit to maximize its revenue opportunities [2&10]."

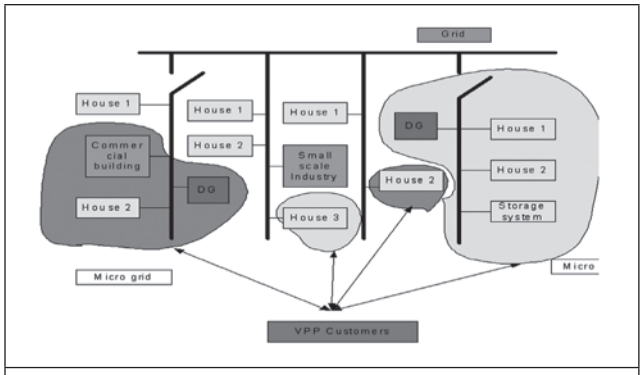

FIG.1 COMPONENTS OF VPP

2.2 Components of VPP

An ideal VPP consists of three main components, Generation Resources, Energy Storage Systems (ESS) and Information and Communication Technologies (ICT) [11].

- Generation Resources

Generation can be either distributed generation or conventional generation already connected to the grid.

- Energy Storage Systems (ESS)

ESS and its elements play an important role in bridging the gap between the supply and demand, particularly in the presence of high penetration of DGs. ESS can store the energy during low demand periods and feed it during the high demand periods. ESS plays an important role in the technical function and financial benefits. Technical role of ESS includes grid frequency support, load leveling/ peak shaving, spinning reserve, power quality improvement and power reliability [3].

- Information and Communication Technology (ICT)

Information and Communication Technology is the important element for VPP. ICT plays an important role in generation scheduling, pricing, grid constraints and grid status, etc.

3.0 CUSTOMER SEGMENTATION TECHNIQUES

Clustering is the basic tool used to segregate the electricity customer belonging to same consumption levels. Segmentation helps not only in competitive energy trading but also in inter and intra regional energy exchange. It also, helps many customers understand the electricity market policies and their consumption patterns. Based on the knowledge of segmentation, a customer can change their electricity utilization plans cost-effectively and optimally [4].

- Methodologies of segmentation:

Electricity customer segmentation is associated with load pattern of the customer. Segmentation is used to segregate the electricity customers belonging to same load pattern into the same clusters. Various clustering techniques like hierarchical clustering, K-means and fuzzy K-means are generally used to segregate customers with similar consumption behaviour[12].

3.1 Hierarchical clustering

It is following bottom-up aggregation or top-down split of clusters for given data until the clustering result is formed. Then, all similar pair of clusters will be forming a single cluster based on some criteria. This procedure will repeat until all the clusters becoming a part of the single large cluster.

The process of hierarchical clustering is easy to implement. However, there are some disadvantages: 1) A few points that outline a link between two clusters make this clustering method to unify these clusters into one. 2) It may cause enlarged clusters to split and for portions of adjacent enlarged clusters to combine [5].

3.2 K-means Clustering

K-means clustering is an unsupervised learning algorithm. It groups a data set of N customers in K clusters. Initially, a guess is made for the k cluster centroids C(k), for k = 1,..., K (generally

preferred at random among the customers of the data). The centroids are used to segregate the customers into clusters, so the better the size of clusters, the better result and values of centroids are recalculated. These steps will be repeated until the data points are formed into clusters [6].

The Procedure of K-means is simple and it uses the sum of the squared distances to the cluster. But, the drawbacks of K-means are: 1) Random selection of cluster size can affect the entire result, 2) Mostly, Normalization is done by standard deviation, 3) Finding the number of clusters is very difficult.

3.3 Fuzzy K-means clustering

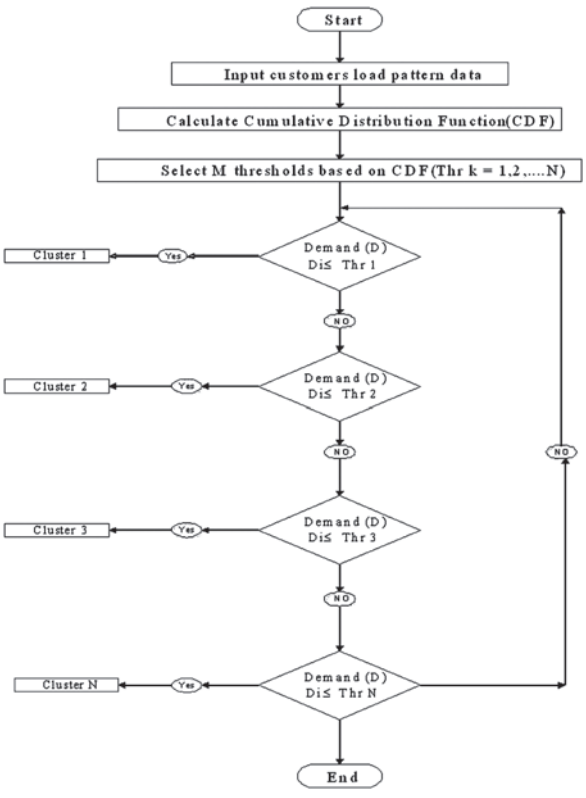

FIG.2 FLOWCHART OF THRESHOLD-BASED CLUSTERING

Fuzzy K-means clustering is relatively connected with k-means clustering, but each customer x (d) has a degree of membership an (d, k) to all possible cluster k. The process is initialized by choosing C(k) as cluster centroids and allotting grade of membership with the k centroids to the N customers of data. Every cluster centroid C(k) is updated with the fuzzy mean of each and every customer regarding the cluster k. The procedure is repeated until stabilization of the cluster centroids. Moreover, the number of clusters and membership criteria are defined by trial-and-error.

There are some shortcomings when using the fuzzy K-means algorithm. 1) Its initial cluster numbers might be inaccurate in an outlier environment. 2) There will be distributions and sparseness on clusters [7].

Customer segmentation process supports the decision-making of different stakeholders like DERs, DSOs and Policy makers. Some papers have discussed about malicious data identification and correction, and load forecasting based on electrical customer segmentation. Also, it helps to plan power transaction from prosumer and/or generation, demand response, it can improve the appropriateness of tariff design, etc [8].

4.0 PROPOSED METHOD AND TECHNIQUES

Two methods have been proposed for segmenting the customers.

- Half-Hourly clustering- in which the customers will change their cluster for every half an hour based on their consumption and electricity prices are calculated individually for each and every half an hour of the day based on their cluster type.
- Daily clustering- in which the customers will be in the same cluster for the whole day and electricity prices are calculated for the whole day based on their cluster type.

4.1 Proposed segmentation method

To overcome the shortcomings of existing clustering methods like determining appropriate cluster number, etc., an easy and less time-consuming segmentation method has been proposed which is called as threshold-segmentation method.

- Procedure for the threshold-segmentation method:

In this paper, a highly accurate and straightforward way to provide better clusters with robustness is proposed. The objective function is proper allocation of customers into their respective cluster. It can be defined as,

$$Diff = (Thr_k - Demand_i); i \in N, k \in K \quad ...(1)$$

$$\begin{Bmatrix} Clustercustomers = \\ Diff = +veor\ 0 \quad \in clusterk \\ Diff = -ve \neq clusterk \end{Bmatrix}$$

Where N= number of customers, K= number of clusters, Thr= Threshold value, Demand belongs to the customer.

To efficiently achieve this objective function the threshold- segmentation has been proposed. The procedure of customer segmentation using threshold-segmentation method is shown in Figure 2.

Finally, the clusters have been evaluated for the electrical energy consumption. Using this information the total cost for every customer in each cluster is calculated using different schemes.

5.0 APPLICATION OF ELECTRICITY CONSUMER SEGMENTATION

In this paper, the electricity customer segmentation has been used to design suitable tariff structure for VPP. The successful outcome of difference between the customer types encourages the association of a specified tariff to each customer type. Initially, customers with the same total consumption and different peak consumption are charged the same. Currently, for electricity billing, the various packages of tariff schemes has been increasing day by day.

The main methodology of price design involves establishing the customer segmentation [9]. This paper considers three different tariff structures, namely Cluster type-based pricing, Cascaded pricing and Flat pricing. Cluster type-based pricing and Cascaded pricing are multiple rate type with rates differentiated on the basis of the hour and day of consumption, while Flat pricing is single as the name suggests.

Every tariff schemes has its unique structure. Tariff offered by a company and/or government are same for all customers within the same region. Likewise, the cluster type-based pricing scheme defines same price for each customer who belongs to a particular cluster type.

If k=1: n

$$Price1_k = \sum_{j=1}^{A_k} D_{jk} \times P_k \quad (2)$$

end

In equation (2), A_k is the number of customers belonging to the k^{th} cluster, n is the total number of clusters.

In Flat pricing, the price remains the same for all the customers.

$$Price2_k = \sum_{j=1}^{A_k} D_{jk} \times P \quad (3)$$

Cascaded Pricing is described as a single customer will pay the different price for the demand in the same hour based on the divided point between everycluster.

$$D_k = \sum_{k=1}^{A_k} [Thr_k - Thr_{k-1}] \quad ...(4)$$

If $D_{jk} \geq D_k$,

$$Price\ 3 = \sum_{j=1}^{A_k} [D_{jk} - D_k] \times P_k \quad ...(5)$$

else

$$\text{Cascadedpricing} = \text{sum}[(\text{Price } 3) + (D_k \times P_k)] \quad ...(6)$$

Using Equations (2) to (6), total prices have been calculated.

6.0 CASE STUDY

6.1 Data set

Electricity consumption pattern data based on 208 domestic customers are collected from the Scottish and Southern Energy (SSE) [13]. The clustering decision is taken in 30-minute resolutions.

Then, the customers are subjected to Data pre-processing method. 185 consumption patterns are considered after the malware detection of data. Therefore, pre-processing ensures that the actual data is of high quality and not erroneous. Segmentation of electrical customers is performed by using the existing and proposed algorithm. They are K-means clustering method, Threshold-segmentation method respectively. Half-hourly and daily clustering has been done in this work.

6.2 Segmentation result

- **K-means clustering method**

The electricity customers have been clustered into 4 groups based on their consumption. For every half anhour, the clustering process will take place continuously. Figure 3 shows that the total number of customers belonging to each cluster for every half an hour. Figure 4 displays the total electricity consumption of particular cluster for every half an hour. From Figure 3 and Figure 4, it can be seen that the number of customers is high in the first cluster compared to other clusters but the total demand of the first cluster and other cluster are almost the same. Sometimes last cluster has high consumption compared to the first cluster, though the first cluster has a high number of customers than the last cluster. This is because of low consuming customers belongs to the first cluster while highly consuming customers fit into the last cluster. [Prosumers can make money by transferring energy from 1st cluster to 4th cluster]

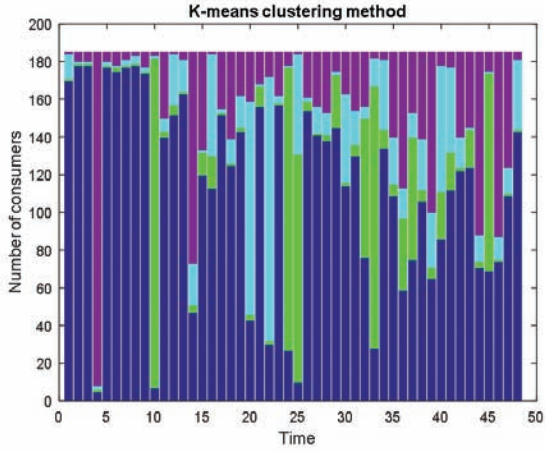

FIG.3 HOURLY CLUSTERING USING K-MEANS CLUSTERING METHOD

As discussed earlier, the clustering results will be applied to the new tariff schemes. Figure 5 shows that the total cost is paid by the customers for the whole day with the proposed tariff structures. Three pricing schemes have been discussed in this paper. Cluster type based pricing method gives the lowest price compared to other methods.

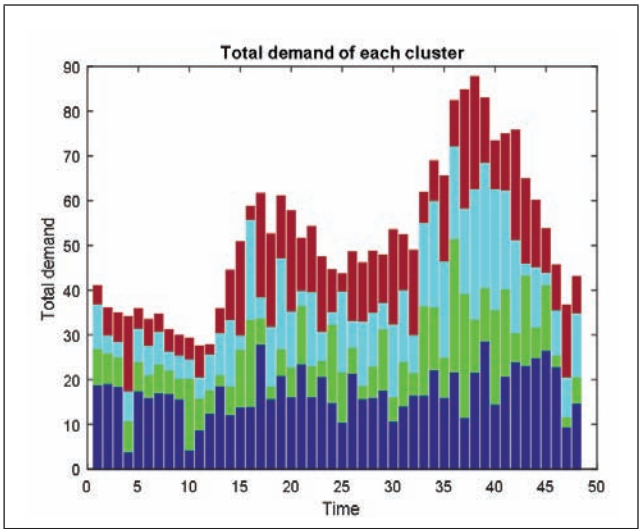

FIG.4 TOTAL DEMAND (W) OF EACH CLUSTER

- **Threshold-segmentation method**

Customers are being clustered into 4 groups by using Threshold-segmentation method based on their power consumption. Thresholds have been selected based on the Cumulative Distribution Function (CDF). Figure 6 shows that the CDF of demand. Figure 7 displays the hourly clustering using Threshold-segmentation method.

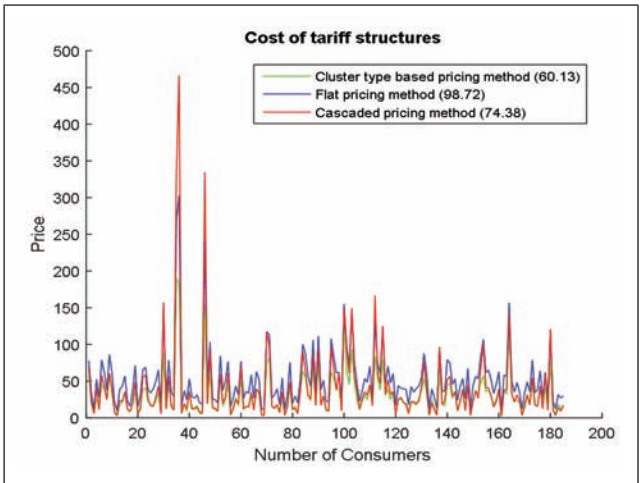

FIG.5 PRICING RESULT WITH NEW TARIFF SCHEMES FOR K-MEANS CLUSTERING METHOD

Total demands of 3rd and 4th clusters are higher than the first cluster, though the first cluster has a high number of customers than 3rd and 4th. Total power consumption of every cluster is shown in Figure 8. The clustering results are applied for pricing. The proposed pricing schemes have been applied to clustered customers and the total cost of each customer is shown in Figure 9.

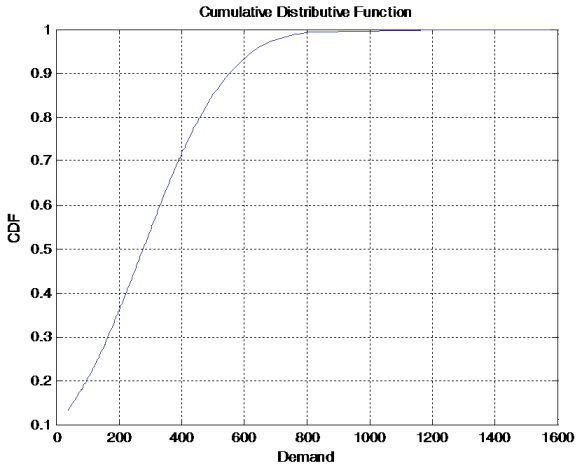

FIG.6 CUMULATIVE DISTRIBUTION FUNCTION OF DEMAND

Total cost paid by all customers for the day using different pricing schemes was shown in Figure 5 and Figure 9. The total cost remains the same in Flat pricing method. Cluster type based pricing method has only 4£ difference. Cascaded pricing method has significant changes. Figure 10 shows the Number of shifts of customers from one cluster to another. Comfort level of customers is better in threshold-based segmentation method compared to K-means clustering method. Here, the comfort level is high if customers have lesser shifts between the clusters and low if they have more shifts.

7.0 DISCUSSION

The total number of customers in a cluster, the total demand of each cluster, pricing and a total number of consumers shift from one cluster to another cluster have been studied for 185 customers. In this paper, threshold-segmentation method has been proposed.

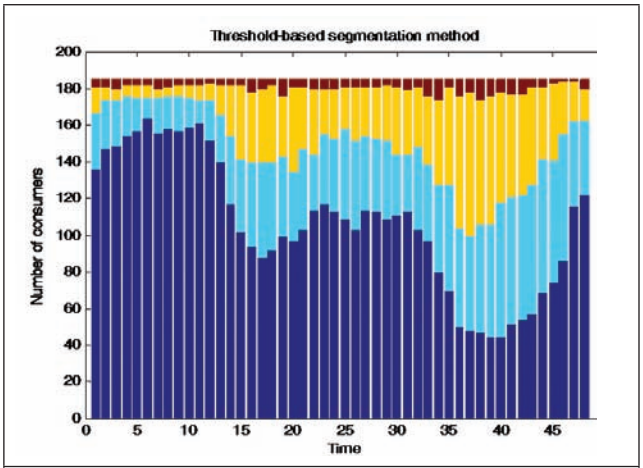

FIG.7 HOURLY CLUSTERING USING THRESHOLD-SEGMENTATION METHOD

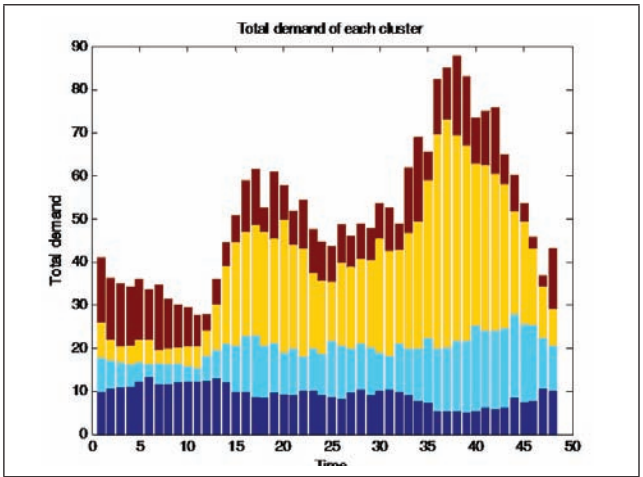

FIG.8 TOTAL DEMAND (W) OF EACH CLUSTER IN THRESHOLD-SEGMENTATION METHOD

Customers have been clustered based on K-means clustering method to compare the performance of threshold-segmentation method because K-means clustering algorithm is a simple and commonly used tool for clustering purposes

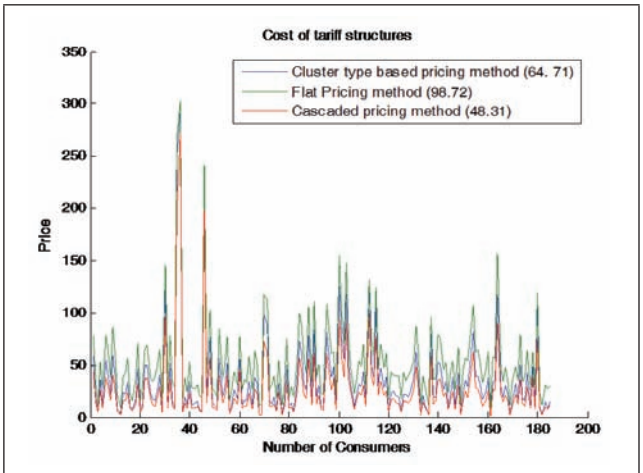

FIG.9 COST WITH NEW TARIFF SCHEMES FOR THRESHOLD- SEGMENTATION METHOD

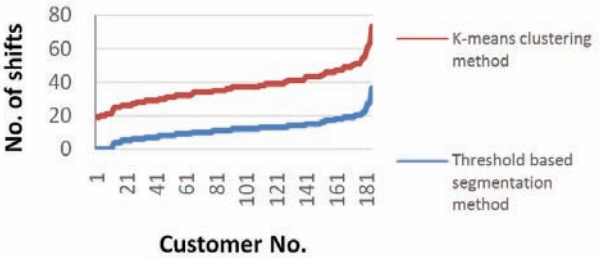

FIG.10 NUMBER OF SHIFTS OF CUSTOMERS FROM ONE CLUSTER TO ANOTHER

Threshold-segmentation method is compared with K-means clustering method based on the following criteria.

1. Total cost
2. Comfort level
3. Clarity of picture

Total Cost –Total cost paid by the customers for a day is less in Threshold-segmentation method compared to K-means clustering method. Cost is one of the key factors to select a method. It can be evident from Figure 5 and Figure 9. Based on this factor, it can be observed that threshold-segmentation method is better.

Comfort level – Customers are changing from one cluster to another cluster, which is happening frequently in K-means clustering method. It might be happening due to the high cost, incentives and demand, etc. Continuous change of customers across the clusters might confuse them with different tariff rates which will make them feel uncomfortable. Figure 10 displays that a total number of customers' change is less in threshold-segmentation method.

Clarity of Picture –K-means clustering result didn't give more clarity and couldn't get any information [Figure 3], meanwhile threshold-segmentation method gives the clarity [Figure 7]. It can be evident from Figure 3 and Figure 7.

As discussed earlier, customer segmentation plays important role in demand side management in a VPP environment. Threshold-segmentation method gives an opportunity to reduce the total cost incurred by the customers. If they are the prosumers, they can sell energy to high consuming customers with a high cost. For example, prosumers belonging to cluster 1 produce electricity for 2p, they can sell this electricity for 3p or 4p to the high consuming customers but they can get electricity for 1p from the grid. Cascaded tariff scheme discussed here, assigns a low price for low consuming customers and high price for high consuming customers to minimize the number of high consuming customers. Hence, the clustering helps VPP to make a profit.

8.0 CONCLUSION

This paper presents an efficient method for electrical customer segmentation. The segmentation method uses SSE data and it is based on Threshold-segmentation method. The proposed method was implemented in MATLAB and was compared with commonly used K-means clustering method. The result showed that the Threshold-segmentation algorithm can classify the demand data into well-separated clusters. The obtained clustering results are used for different pricing schemes. It was proven that the proposed method is simple and accurate method from pricing results.

ACKNOWLEDGEMENT

The research leading to these results has been funded by the British Council and the Department for Business, Innovation and Skills. Smart Grids to harness Satellite based virtual power plants for energy Sustainability (SITARA) is a project under the Global Innovation Initiative in 2015 which is a joint effort of the UK and US to foster multilateral research collaboration with higher education institutions in Brazil, China, India and Indonesia. is a project under the Global Innovation Initiative in 2015 which is a joint effort of the UK and US to foster multilateral research collaboration with higher education institutions in Brazil, China, India and Indonesia.

REFERENCES

[1] Q. Zhao, Y. Shen, and M. Li, Control and Bidding Strategy for Virtual Power Plants with Renewable Generation and Inelastic Demand in Electricity Markets,IEEE Trans. Sustain. Energy, Vol. 7, No. 2, pp. 562–575, 2016.

[2] H .Saboori, M. Mohammadi, and R. Taghe, Virtual Power Plant (VPP), Definition, Concept, Components and Types,IEEE Power and Energy Engineering Conference (APPEEC), 2011 Asia-Pacific.

[3] T. Logenthiran and Dipti Srinivasan, Intelligent Management of Distributed Storage Elements in a Smart Grid, IEEE PEDS Singapore, 5 - 8 December 2011.

[4] K. Le Zhou, S. L. Yang, and C. Shen, A review of electric load classification in smart grid environment, Renew. Sustain. Energy Rev., Vol. 24, pp. 103–110, 2013.

[5] S. V. Verdu, M. O. Garcia, C. Senabre, A. G. Marin, and F. J. G. Franco, Classification, Filtering, and Identification of Electrical Customer Load Patterns Through the Use of Self-Organizing Maps, IEEE Trans. Power Syst., Vol. 21, No. 4, pp. 1672–1682, 2006.

[6] Gianfranco Chicco, Roberto Napoli and Federico Piglione, Application of clustering algorithms and self organising maps to classify electricity customers,lEEE Bologna PowerTech Conference, June 23-26, Bologna, Italy, 2003.

[7] A. Mukhopadhyay, U. Maulik, and S. Bandyopadhyay, Multiobjective genetic algorithm-based fuzzy clustering of categorical attributes, IEEE Trans. Evol. Comput., Vol. 13, No. 5, pp. 991–1005, 2009.

[8] A. Mohd, E. Ortjohann, A. Schmelter, N. Hamsic, D. Morton, S. Westphalia, A. Sciences, D. Soest, and L. Ring, Challenges in integrating distributed Energy storage systems into future smart grid, 2008 IEEE Int. Symp. Ind. Electron., pp. 1627–1632,2008.

[9] Gianfranco Chicco, Roberto Napoli, Petru Postolache, Mircea Scutariu, and Cornel Toader, Customer Characterization Options for Improving the Tariff Offer, IEEE Trans. Power Syst., vol. 18, no. 1, pp. 381-387, 2003.

[10] M. M. Othman, M. M. Othman, and Y. G. Hegazy, "A Review of Virtual power plant Definitions, Components , Framework and Optimization," International Electrical Engineering Journal (IEEJ) Vol. 6 (2015) No.9, pp. 2010-2024.

[11] E. A. Setiawan, Concept and Controllability of Virtual Power Plant, 2007.

[12] Anderberg, M. R., Cluster Analysis for Applications, New York: Academic Press, 1973.

[13] http://www.thamesvalleyvision.co.uk/project-library/research-data/