Journal of Optimization in Industrial Engineering
Vol.16, Issue 1, Winter & Spring 2023, 119-127
DOI:10.22094/JOIE.2023.1975210.2028

Research Article

Residential Appliance Clustering Based on Their Inherent Characteristics for Optimal Use Based K-Means and Hierarchical Clustering Method

Shima Simsar [a], Mahmood Alborzi [a,*], Ali Rajabzadeh Ghatari [c], Ali Yazdian Varjani [d]

[a] *Department of Information Technology Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran*
[c] *Department of Management, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran*
[d] *Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran*

Received 25 December 2022; Revised 05 May 2023; Accepted 09 May 2023

Abstract

With global warming and energy shortages, smart grids have become a significant issue in the power grid. Demand response is one of the basic factors of smart grids. To enhance the efficiency of demand response, an intelligent home appliance control system is essential, which prioritizes the start-up of electrical appliances according to the necessity of use and efficiency. To properly manage the demand response, utilities use different signals such as price. One of the pricing methods that can be considered is different pricing for electrical appliance clusters. In this article, appliances are clustered by the K-means and hierarchical clustering based on the characteristics of the appliances themselves, such as the appliances' extent of consumption, the type of use of home appliances, how home appliances work, the ability to change the working conditions of home appliances, home appliances usage time, etc. It seems that the K-means clustering method outperforms the hierarchical method in this issue, due to its lower value of DB coefficient. In this method, home appliances were classified into three clusters. The silhouette coefficient was developed as a measure of the K-means clustering model performance, where the average silhouette coefficient of 0.6 indicates the satisfactory value of the model. Based on the results, it was found that the proposed clustering method can rationally classify different types of home appliances by selecting the appropriate characteristics since the appliances in a cluster are very similar to each other and can help users understand the operating conditions of the appliances.

Keywords: Appliance; Demand response; The K-means clustering; Hierarchical clustering

1. Introduction

Today, for various reasons such as population growth, industrial development, and rising living standards, etc., electricity consumption has increased. Inadequate attention to proper management of electricity distribution seriously endangers the socioeconomic development of countries. The electricity distribution model is related to the relationship between production and consumption in the grid. To meet peak load demand, seasonal and daily fluctuations, and ensure the efficiency of power system operations, utilities must maintain a significant amount of their capacity. This capacity, which is often obsolete and very harmful to the environment, affects electricity costs, which indicates a significant increase in the price of wholesale electricity during peak hours (Haider et al., 2016). A typical solution is to match supply with demand, but this approach is unstable and difficult to trust (Kostková et al., 2013). The construction of a traditional power plant is not preferred due to the increase in greenhouse gas emissions (Delmastro et al., 2015).
Further, adding a new power plant is not a worldly-wise solution, as it meets the demand for a limited period and then leads to wasted energy (Haider et al., 2016). On the other hand, wind and solar energy are the least productive have the problem of natural outages and high costs. In addition, they depend on specific geographical conditions and, therefore cannot be developed in every region. These resources also rely on climate, thus further reducing their reliability (Haider et al., 2016).
To meet these challenges, the scientific community as well as industrial entities, are taking steps to upgrade their grid infrastructure and related technologies to ensure energy production and supply over the next century. In this scenario, the smart grid is an emerging technology that is related to different role players in different parts of energy systems such as public industries, transmission and distribution companies, customers, equipment manufacturers, service providers, or electricity traders. While many current solutions that must meet demand are based on the traditional idea of increasing supply to meet demand, demand response by managing demand opposes the idea mentioned above and seeks to match the available energy (Warren, 2014; Olamae, Ashouri, 2015). To properly manage the demand response, utilities use different signals such as price (Muratori, Rizzoni, 2016). One of the pricing methods that can be considered is different pricing for electrical appliance clusters. Thus, this research has clustered home appliances based on their characteristics, based on this, one can examine the optimal management of electricity consumption in the

*Corresponding author Email address: mahmood_alborzi@yahoo.com

smart grid for a different price for each cluster of home appliances in future studies. Regarding organization, section 2 reviews the literature and studies on the classification of home appliances. Section 3 examines the data structure and input of the clustering model. Section 4 describes the methodology used in the research. Section 5 discusses the model and the results obtained. Finally, the paper is concluded in Section 6.

2. Literature Review

Demand response helps utilities and customers reduce peak demand and price volatility (Siano, 2014; Faruqui, 2006). Instead of adapting electricity generation to match changes in demand, demand can be made more flexible to reduce the requirements for electricity generation infrastructure. Demand response is an economical and promising solution that will make electricity demand more flexible and enable private customers to change their demand plan to meet their energy supply needs. In the demand response paradigm, the power system offers incentives and benefits to customers as compensation for their flexibility in consuming electricity. Utilities also provide a signal to their customers (usually the price of electricity) whose goal is to direct electricity consumption so they can better respond to aggregate demand which is commensurate with the generation. Demand response is very effective in shifting consumption from peak hours and causes enhanced productivity and stability while reducing the need to invest in production during peak times and bringing many financial as well as environmental benefits (Muratori, Rizzoni, 2016). To achieve this goal, the right signal must be sent to the end customer to create a win-win situation for customers and utilities. Household energy consumption in green or passive buildings is greatly influenced by occupants' behaviors, appliances and devices and the way they interact with each other (Leroy, Yannou, 2018).

In the paper (Shirazi, Jadid, 2015), scientists described a new method to minimize energy prices in a dynamic pricing scheme named HEMDAS. They divided home appliances into three categories, including controllable and uncontrollable electrical appliances and heat-controllable appliances. A schedule is provided for each category. In this paper, a two-way data flow between the public company and the smart meter is controlled. Smart meters are responsible for measuring the processes transmitted to the public company. The major contribution of this plan is minimizing the cost with a certain level of comfort defined by the residents. In the article (Javaid et al., 2016), first, home appliances are divided into six sub-categories that have the same function or behavior. They are then classified into three general categories, taking into account the similar capabilities of the appliances or purpose of use, including heating purposes (heater) and cooling purposes (cooler) and the general working group (laundry and refrigerator). In the article (Iqbal et al., 2018), which aims to reduce the cost of electricity consumption and minimize the peak consumption load compared to the average rate by considering the user comfort in a smart home, home appliances in the house were classified into three categories. The first category, Non-Deferrable Loads (NDLA) cannot be transferred to another time; NDLAs have a starting and ending point for describing their time, and the consumer cannot adapt to such devices. In the second category, Interruptible Loads consumers can suspend in the middle for a limited period of time to supply electricity. The third category, Must-Run Loads (MRLA) must start immediately at any time. These appliances are not suspensible. They must be implemented at all costs. In the paper (Li et al., 2018), researchers designed an intelligent energy management system based on renewable energy, in which home appliances were divided into three categories: Delay-Tolerant, Delay-Intolerant with Essential Load, and Delay-Intolerant with Flexible Load. In the article (Hassan et al., 2016) to manage electricity demand, home appliance loads were divided into three categories: Delay-Tolerant, Delay-Intolerant with Essential Load, and Delay-Intolerant with Flexible Load. For delay-tolerant appliances, their starting times can be shifted from peak to off-peak hours to reduce peak load. These delay-tolerance devices include electric gas stove, microwave oven, water heaters, dishwashers , washing machine, clothes dryers, and so on. For Delay-Intolerant appliance with Essential Load, it is not allowed to change their starting time, and their load variation between different modes is usually negligible; LED bulbs, fans, televisions, and computers are this types of delay . In the case of Delay-Intolerable with Flexible Load appliances, users can alter the operating modes of their device to save energy and correct the peak load. These types of appliances include refrigerators and air conditioners.

In the paper (Jeong et al., 2021), A systematic method was proposed to group residential energy consumption models using demographic characteristics. The K-means-based features selection method was used to classify energy consumption patterns of residential customers into six types. The analytical results of the article have shown that extreme points were effective in grouping the energy consumption patterns of residential customers.

In the article (Tambunan et al., 2020), a simulation concept was proposed for analysis of peak load data by K-means clustering algorithm based on historical dataset. The results determined that electrical peak loads clustering by K-means algorithm was optimum classified into three clusters. This cluster was evaluated by silhouette scores which were high, intermediate, and low load level interpretation.

In the research (Czétány et al., 2021), the objective was to evaluate database in detail to determine energy consumption profiles from time series of daily and annual electric load. After representativity check of dataset daily and annual energy consumption profiles were developed, applying three different clustering methods (k-means, fuzzy k-means, agglomerative hierarchical) and three different cluster validity indexes (elbow method, silhouette method, Dunn index) in MATLAB environment. The best clustering method for the examination proved to be the k-means clustering technique. Analyses were carried out to identify different consumer groups, as well as to clarify the impact of specific parameters such as meter type in the housing unit (e.g. peak, off-peak meter), day of the week (e.g. weekend, weekday), seasonality, geographical location, settlement type and housing type (single-family house, flat, age class of the building). Finally, four electric user profile types were proposed.

In the paper (Rajabi et al., 2020), a comparative study of different techniques for load pattern clustering was carried out. Different parameters of the methods that affect the clustering results were evaluated and the clustering algorithms were compared for two data sets. In addition, the two suitable and commonly used data size reduction techniques and feature definition/extraction methods for load pattern clustering were analyzed.

In the article (Si et al., 2021), the basic concepts and the general process of electric load clustering were summarized. Several similarity measurements and five major categories in electric load clustering were comprehensively summarized along with their advantages and disadvantages. Afterward, eight indices widely used to evaluate the validity of electric load clustering were described. Finally, vital applications were discussed thoroughly along with future trends including tariff design, anomaly detection, load forecasting, data security and big data, etc.

The study (Cen et al., 2022), proposed clustering algorithms to segment consumers and obtain representative load patterns based on diurnal load profiles. First, the proposed method uses discrete wavelet transform (DWT) to extract features from daily electricity consumption data. Second, the extracted features were reconstructed using a statistical method, combined with Pearson's correlation coefficient and principal component analysis (PCA) for dimensionality reduction. Lastly, three clustering algorithms were employed to segment daily load curves and select the most appropriate algorithm. They experimented with the method on the Manhattan dataset and the results indicated that clustering algorithms, combined with discrete wavelet transform, improve the clustering performance.

In the article (Rasheed et al., 2016), home appliances were divided into three categories: 1. User dependent, 2. Interactive Schedulable, and 3. Unschedulable. The first group is the main requirement of home appliances where energy consumption is based on the main needs of the user based on various factors such as weather conditions, etc. For example, these appliances include refrigerators, lamps with controllable lighting and air conditioning system. In the second group are home appliances that can be scheduled during the period, time T. Energy consumption and user satisfaction with these appliances can be measured by the total amount of energy consumed in each period, time. These appliances are scheduled at low times. Dishwashers and uncontrollable lighting fixtures are kept in this category. The third group are is appliances with a fixed energy consumption, such as entertainment appliances (TVs, music players, etc.). The minimum and maximum energy consumption of these appliances are 0 and 1500 W, respectively.

Based on studies, the existing methods to describe household electricity consumption, in general, can be divided into four categories: statistical, engineering, time series, and clustering. Statistical methods are used for a wide range of purposes such as Time of Use (ToU), which indicates the amount and time of use (residential, commercial, industrial load), though this method is suitable for settlement purposes, in fact, not suitable for how to consume electricity at home on a daily basis; and they merely display the average value for all customers in the same group. Engineering approaches to describe the load characteristics of the home are different but generally describe power consumption as a function of parameters such as location and various appliances. This method usually has a bottom-up approach, since a profile is created for each household. However, engineering methods are difficult to generalize since they require accurate knowledge of home occupants and when to use home appliances. In contrast, time series approaches are used to describe power consumption at the operator level and are limited because they depend on historical data and suffer from the same problem as statistical methods such as aggregating different customer profiles. Clustering data mining approaches are used to group customers who behave similarly electrically through power consumption data. For example, in a study of approximately 3,000 residential customers over one year , SOM, the K-means, and hierarchical methods were used to cluster and construct load profiles (McLoughlin et al., 2015). For this purpose, in this research, the clustering method was used to describe the consumer electrical appliances, with the difference that clustering is based on the specifications of the home appliances themselves to cluster the home appliances in the households.

3. Data Structure

This research used the data related to home appliances recorded on the Daftlogic website to classify home appliances used by home subscribers based on the amount and time of consumption and the inherent characteristics of home appliances. The data of this research were completed by experts. To start clustering, the completed information related to home appliances was pre-fitted to be prepared for clustering. For this purpose, the consumption of households was normalized and prepared for clustering analysis by the desired algorithm.

4. Methodology

In this research, the home appliance detection algorithm based on the K-means clustering model and hierarchical clustering was investigated. Based on the characteristics of home appliances and the opinion of experts, the data expressed in Section 3 should be used for clustering home appliances. R software was used for clustering and model analysis.

The clustering strategy in this research consists of four steps: data collection, data preprocessing, the K-means and hierarchical clustering, and model or pattern recognition. In general, considering the characteristics of home appliances, appliances are divided into several groups; since the K-means clustering method requires determining the correct number of clusters, the hierarchical clustering method was examined to solve this problem. Finally, the evaluation of the clustering results was done by silhouette coefficient and Davies Bouldin Index

4.1. Selecting the Clustering Characteristics

Since clustering is determined based on the similarity and dissimilarity of the characteristics that describe the entities, selecting appropriate characteristics that can help identify different groups is of particular importance (Wang, Srinivasan, 2015). As displayed in Table 1, the following characteristics were selected leading to different performances and working at different times in the model.

Table 1.
Characteristics of home appliances

Amount of consumption	Type of use	Ability to be used by the general public	Type of operation of the appliance	Ability to change the conditions of the appliance	Ability to change the timing of use of the appliance	Type of weather of use of the appliance	Time of use	Days of use	Type of weather conditions
	Personal	Yes	Continuous	Yes	Yes	Cold	Night	Holidays	Cloud
kWh per year	Family	No	Pause if needed	No	No	Hot	Day	Non-holiday	Rainy
									Sunny

4.2. The K-Means clustering

In the first stage, according to the literature review, one of the most widely used clustering techniques (K-means) for the electricity industry was reviewed and selected; the appropriate number of clusters for data categorization was determined. After developing the model, the silhouette coefficient was used to identify the suitability of the model. Clustering is an unsupervised learning method that groups a set of data in several clusters so that the entities in one cluster are very similar, while the entities in the other cluster are less similar (Han et al., 2012). In other words, the clustering technique groups entities based on similarity or dissimilarity, where similarity and dissimilarity are evaluated based on the characteristics used to describe objects. In this paper, a center-based clustering algorithm is used to detect the pattern since it is efficient for multidimensional data. The basic idea of this algorithm is to assign entities to the most similar cluster and update the value of the cluster center frequently to make the assignment stable. There are K similar groups in the K-means algorithm (Bagherighadikolaei et al.,2020) and a cluster is expressed in terms of its center, which is defined as the average value of the entities within the cluster. The difference between each entity in the cluster and the center of the cluster is measured by the Euclidean distance, which is defined as follows:

$$d_{euclidean}(x.C) = \left[\sum_{i=1}^{d}(x_i - c_i)^2\right]^{\frac{1}{2}} \cdot \tag{1}$$

In the above formula, x is the entity and C is the center of the cluster, d represents the dimension of the entity, x_i, is the ith value of the entity characteristics, and c_i shows the ith character of the value of cluster center characteristics (Alborzi, Alikhani, 2016).

4.2.1. Determining the number of K clusters

The number of suitable clusters in this method should be determined in advance (Han et al., 2012) since the appropriate number of clusters should establish a good balance between accuracy and compactness in cluster analysis. In this study, it is important to determine the number of clusters as the data set is unlabeled. Determining the correct number of clusters depends on the shape and scale of the distribution in the data set as well as the clustering resolution required for this research work. An error function is introduced to help find the exact number of clusters for this study. An error function is defined as follows:

$$E = \sum_{i=1}^{k}\sum_{x \in C_i} d(x_i . C_i) \cdot \tag{2}$$

In the k-mean algorithm, like other optimization algorithms, there is a cost function the algorithm tries to optimize, which is defined as follows:

$$J\left(C^{(1)} \dots C^{(m)} . \mu_1 \dots \mu_k\right) = 1/m \sum_{i=1}^{m} x^{(i)} - \mu_{c^{(i)}}^2 \cdot \tag{3}$$

$C^{(i)}$ is the cluster to which $x^{(i)}$ is assigned, μ_k denotes the center of gravity of k clusters that change with each iteration, $\mu_{c^{(i)}}$ reflects the center of gravity of the cluster to which the sample is assigned. Thus, the cost function will be the mean squared of the distances of the samples from the center of gravity. The k-Mean algorithm tries to minimize the average distances to all centers.

$$\min J\left(C^{(1)} \dots C^{(m)} . \mu_1 \dots \mu_k\right) \cdot \tag{4}$$

Sometimes the above function is also called the dispersion function. Since by minimizing it, the dispersion data are assigned to specific clusters. Another point about using the k-Mean algorithm is choosing the number of clusters, as there is no specific and automated formula for this work. Nevertheless, there is a method that is used in many studies to determine the number of clusters, called "elbow". In short, in the elbow method, the number of clusters is assumed to be k times 1 and the cost function J is calculated. Then, the number of clusters is gradually added and the values of the cost function J are plotted. It is often observed that the cost function J declines rapidly, but at a point where the reduction slope abruptly decreases, the shape of the cost diagram resembles an elbow, hence the name elbow method. The value of k corresponds to the point where the slope of the cost diagram J decreases and becomes elbow. Then, it is selected as the most appropriate value of k (Alborzi, Alikhani, 2016).

In this research, R software was used for clustering. The elbow diagram according to the input data is as follows. As displayed in Fig. 1, three clusters would be the most appropriate value for the number of clusters according to the description given above.

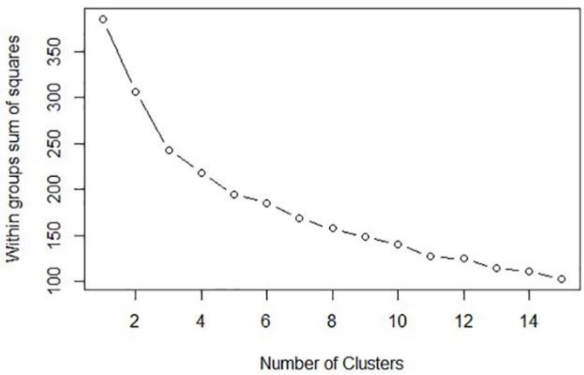

Fig .1. The framework of research

4.2.2. The K-means clustering result

Fig. 2 illustrates the K-means clustering result. To use the model easier, numbers have become appliance names.

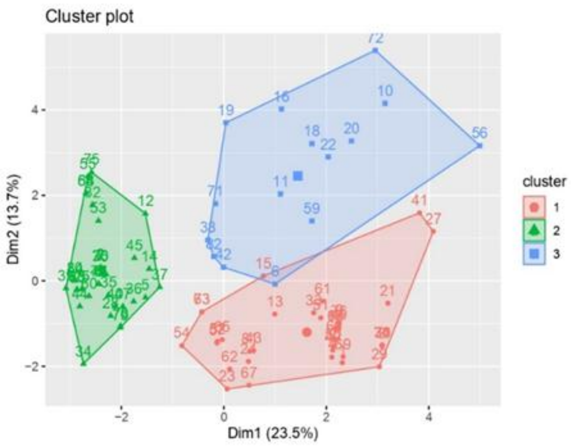

Fig .2. The K-means clustering diagram

4.3. Hierarchical clustering

There are numerous clustering techniques one can encounter in the literature. Most of the existing data clustering algorithms can be classified as hierarchical or partitioning (Abudalfa, Mikki, 2013). A hierarchical clustering algorithm (HCA) is an unsupervised clustering technique that divides things into clusters based on their similarities . The cosine distance is used to calculate the distance between two clusters since it is an accurate indicator of semantic relatedness. The result is a dendrogram, which is a hierarchy of nested clusters where each cluster is distinct from the others and the objects within each cluster are generally comparable to one another (Semeraro et al., 2021). Divisive Analysis (DIANA) and Agglomerative Nesting (AGNES) hierarchical clustering algorithms do not need to specify the number of k clusters as input, but they do require a way to calculate the distance between clusters (Chhabra et al., 2019). The advantage of this type of clustering in comparison with the K-means is that the results are unbiased with initial parameters (faezy, Shadloo, 2017).

4.3.1. DIANA-AGNES hierarchical clustering results

In this research, for hierarchical clustering, first, the data are analyzed by the DIANA method followed by AGNES with the results shown in Fig. 3 and Fig. 4.

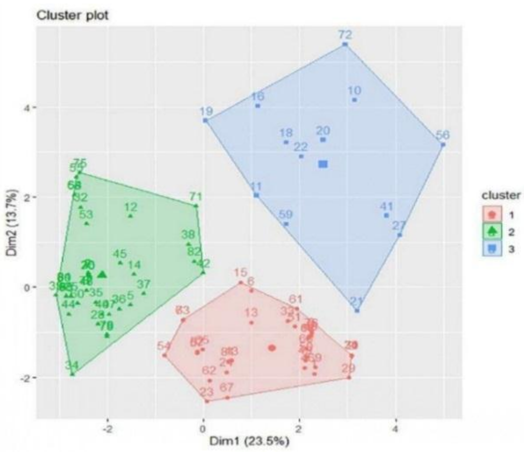

Fig. 3. DIANA hierarchical clustering diagram.

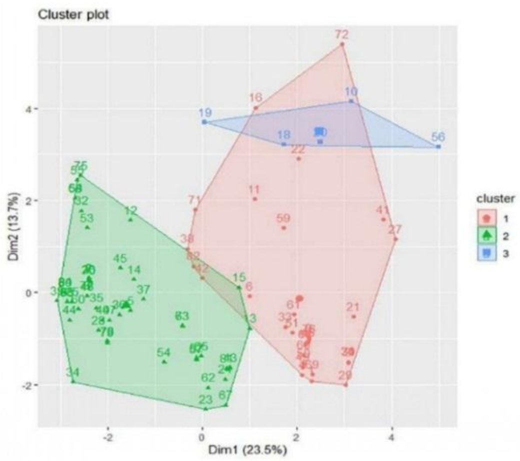

Fig. 4. AGNES hierarchical clustering diagram

4.3.2. Hierarchical clustering validation

The created cluster models (DIANA and AGNES) are compared based on the generated dendrograms. A pair of cluster models are represented in a tanglegram. Tanglegram is used to compare tree diagrams. It measures the quality of the two dendrogram alignments as entanglement (Tamas, 2019). Parallel lines shown in Fig. 5 indicate the validity of the model. This means that two members at the same distance from each other also lie at the same distance in the other diagram.

$$c = cor[h_1(ij).h_2(ij)]\cdot \tag{5}$$

If we calculate the correlation coefficient of distances $h(ij)$ for each of the two points i, j in the same cluster, higher obtained values show the similarity of the two clustering methods. The Pearson correlation coefficient of the two clustering methods of DIANA and AGNES in this issue is 0.85, which indicates the validity of clustering.

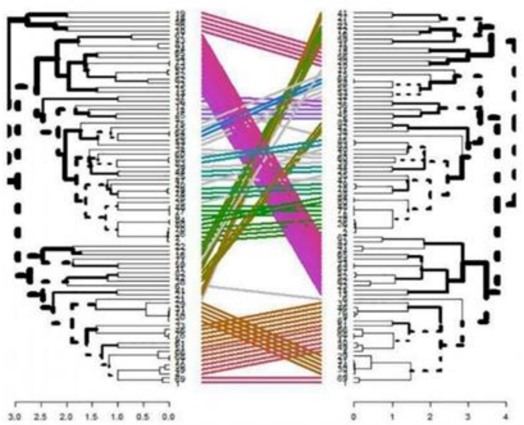

Fig. 5. TANGLEGRAM diagram for DIANA and AGNES clustering.

4.4. Examining the performance of the clusters

4.4.1. Davies bouldin index (DB)

Davies and Bouldin proposed the Davies Bouldin Index (DBI) for evaluating clustering performance, which is an internal evaluation criterion that employs intrinsic data set characteristics and is used as a criterion for evaluating clusters. Intra-cluster dispersion to inter-cluster dispersion ratio is measured by the DB index (Davies, Bouldin, 2019).

$$DB = 1/k\sum_{i=1}^{k}\max_{i\neq j}\big((W_i + W_j)/d(c_i.c_j)\big)\cdot \tag{6}$$

In the above formula, K represents the number of clusters and W_i denotes the average distance of all samples of class c_i to the centers of their clusters, W_j shows the average distance of all samples of class c_i to the center of class c_j, and $d(c_i.c_j)$ is the distance between the centers of classes c_i and c_j (Tian et al., 2019).

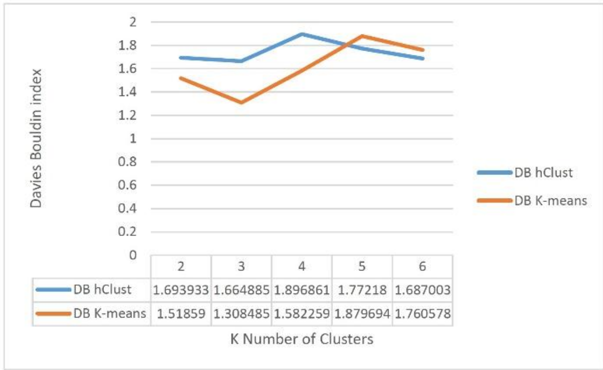

Fig. 6. Dispersion ratio between clusters to cluster dispersion by K-Means and Hierarchical method.

As shown in Fig. 6, the number of three clusters among the selected clusters has the lowest value of DB, which is selected as the optimal value. The K-means clustering method seems to have outperformed the hierarchical method in this regard.

4.4.2. Silhouette coefficients

In this research, the silhouette coefficient method was used to measure the quality of clustering, which was first proposed by Peter J. Rousseeuw. It is a method that shows how well each object lies in its cluster (Han et al., 2012). The silhouette coefficient is defined as follows:

$$s(i) = (b(i) - a(i))/max\big(b(i).a(i)\big)\cdot \tag{7}$$

Where i is an arbitrary point in the data, $a(i)$ denotes the average distance of point i from all entities in its cluster, $b(i)$ shows the minimum mean distance of point i from points in other clusters. The value of the silhouette coefficient is between 1 and -1. Positive values of the silhouette coefficient mean that the cluster containing i is farther away from the other clusters, while negative values of the silhouette coefficient mean that i is closer to other cluster entities than to its own cluster. In general, the larger the silhouette coefficient values, the better, and a value of 1 means extremely suitable conditions.

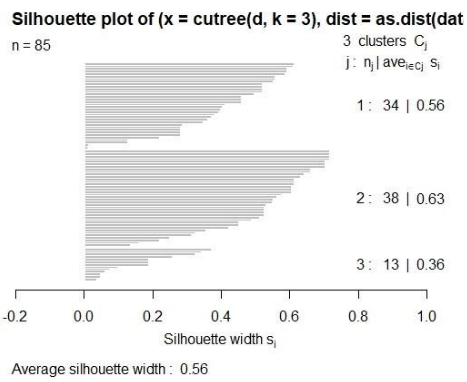

Fig. 7. Hierarchical Silhouette plot.

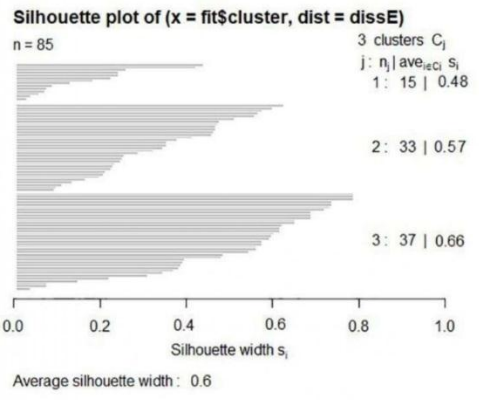

Fig. 8. K-means silhouette plot.

Hierarchical and the K-means silhouette coefficient demonstrated in Fig. 7 and Fig.8. The result of 0.6 silhouette coefficient indicates the satisfactory performance of the model by the K-means method. The K-means clustering method seems to have outperformed the hierarchical method. For this purpose, the clustering result by the K-means method was used to identify the clusters, which are examined in the next section.

5. Result and Discussion

Identifying different clusters is important for users as it describes how home appliances work as different clusters. Appliances were clustered by the K-means and hierarchical clustering, based on the characteristics of the appliances themselves, which indicates in Table 1 finally the K-means method was selected as a better method for this clustering model due to having a lower DB coefficient. In this method, home appliances were classified into three clusters after which the silhouette coefficient was developed as a performance measure . The silhouette coefficient value of 0.6 indicates that each cluster differs from the others. The clustering result shows that the proposed clustering method can rationally classify different types of home appliances by selecting the appropriate characteristics.

As reported in Table 2, the appliances in the first cluster are known as items that are mostly recreational or non-essential, such as outdoor hot tub, straightening iron, water feature, game console, etc., which can be placed in lower priorities. The appliances that are in the second cluster are known in two categories; 1) These loads can be scheduled on the optimal decisions. These devices can shift their run time within user-defined time or optimal defined time such as dishwasher, washing machine etc. 2) Appliances that should always be plugged in (fridge, phone, garage door, Wi-Fi modem, water pump, etc.). The appliances that are in the third cluster are known as essential loads, and should be run immediately; these are the base loads that can be predicted a day ahead including different types of lamps, heating or cooling appliances such as air conditioners or water heaters, and cooking appliances.

Table 2.
Appliance clustering result.

Cluster1	Cluster2		Cluster3	
Clothes Dryers	Air Purifier	Kettle Boiler	Air Conditioner (Room) 6,000 BTU	LED television
Corded Electric Handheld Leaf Blower	Bathroom Towel Heater	Laser Printer	Air Cooler	MI Box
Curling Iron	Coffee Machine	Mobile Phone Charger	AV Receiver	Microwave
Drill	Computer (Monitor & Printer)	Nintendo Switch AC Adapter	Ceiling Fan	Night Light
Electric Blanket	Cordless Drill Charger	Paper Shredder	Cooker Hood	Oven
Electric Mower	Dehumidifier	Power Shower	Crock Pot	Projector
Electric Shaver	Electronic Alarm clock - Radio	Scanner	DVD Player	Rice Cooker
Game Console	Espresso Coffee Machine	Steriliser	Electric Pressure Cooker	Sandwich Maker
Hair blow Dryer	EV Home Charger	Treadmill	Electric Tankless Water Heater	Sewing Machine
Heated Hair Rollers	Flat Screen Computer	Iron	Electrical Stove	Slow Cooker
Outdoor Hot Tub	Fridge A+	Washing Machine A+++	Energy Saver Lamp	Space Heater
Power Saw	Dish Washer	Washing Machine B	Evaporative Air Conditioner	Stereo
Straightening Iron	Furnace Fan Motor (Continuous)	Water Dispenser	Extractor Fan	Stove Hood
Strimmer	Furnace Fan Motor	Water Pump (Deep well – higher powered)	Fan (Portable)	Toaster

Cluster1	Cluster2		Cluster3	
	(Intermittent)			
Water Feature	Garage Door Opener	Water Pump (Deep well – moderate power)	Food Blender	Tower Fan
	Home Phone	WiFi Booster	Fryer	Vacuum Cleaner
	Humidifier (Portable)		Halogen Lamp	Video Cassette Record DVD
			Heater	Water Heater Typical Family(4)
			LCD television	

6. Conclusion

Demand response is one of the basic factors of smart grids. To enhance the efficiency of demand response, an intelligent home appliance control system is essential, which prioritizes the start-up of appliances according to the necessity of use and efficiency. To properly manage the demand response, utilities use different signals such as price. One of the pricing methods that can be considered is different pricing for different appliance clusters. In this research, an unsupervised machine learning model was proposed to cluster home appliances based on the performance and characteristics of the appliances, where a different price for each cluster can be presented. In addition, since clustering is based on the characteristics of the appliances themselves, it can be described by the variations among customers. These characteristics include the number of home appliances used, the type of use of home appliances, how home appliances work, the ability to change the working conditions of home appliances, home appliances usage time, etc.

Examining the results of the study, it was found that the proposed classification method can classify household appliances as the appliances in a cluster are very similar to each other and can help users understand the operating conditions of the appliances. This method is useful for smart grid systems since it can be examined by considering a specific price per cluster of home appliances as a research to provide a solution along with existing methods for monitoring and managing the consumption as well as usage time of each home appliance. A different price is offered for each cluster to provide optimal management of electricity consumption in the smart grid in a new way so that instead of adapting electricity generation to match changes in demand, the demand itself becomes more flexible to reduce the requirements of power generation infrastructure.

References

Azab, A., & Naderi, B. (2014). A variable neighborhood search metaheuristic for cellular manufacturing with multitask machine tools. *Procedia CIRP*, 20, 50-55.

Abudalfa, Sh., & Mikki, M. (2013). K-means algorithm with a novel distance measure. *Turkish Journal of Electrical Engineering and Computer Sciences*, 21(6), 1665-1684. doi:10.3906/elk-1010-869

Alborzi, M., Alikhani, M. (2016). Machine Learning. Sharif University of Technology's Press, Tehran, Iran.

Bagherighadikolaei, S., Ghousi, R., & Haeri, A. (2020). A Data Mining approach for forecasting failure root causes: A case study in an Automated Teller Machine (ATM) manufacturing company. *Journal of Optimization in Industrial Engineering*, 13(2), 101-121. doi: 10.22094/joie.2020.1863364.1630

Cen, S., Yoo, J.H. and Lim, C.G. (2022). Electricity Pattern Analysis by Clustering Domestic Load Profiles Using Discrete Wavelet Transform. *Energies*. 15(4), 1350. doi:10.3390/en15041350

Chhabra,G., Vashisht, V. & Ranjan, J. (2019). Crime Prediction Patterns Using Hybrid K-Means Hierarchical Clustering. *Journal of Advanced Research in Dynamical and Control Systems*, 11, 1249-1258.

Czétány,L., Vámos,V., Horváth,M., SzalayZ., Mota-Babiloni,A., Deme-Bélafi,Z., Csoknyai,T. (2021). Development of electricity consumption profiles of residential buildings based on smart meter data clustering. *Energy and Buildings*, 252, 111376. doi: 10.1016/j.enbuild.2021.111376

Davies, D.L., Bouldin, D.W.(2019). Hierarchical Clustering based on IndoorGML Document. *IEEE Transactions on Pattern Analysis and Machine Intlligence*, 177-182. doi: 10.1109/Informatics47936.2019.9119255

Delmastro, C., Lavagno, E. & Mutani, G. (2015). Chinese residential energy demand: scenarios to 2030 and policies implication. *Energy Build*, 89, 49–60. doi: 10.1016/j.enbuild.2014.12.004

Earle,R., Faruqui,A. (2006). Toward a new paradigm for valuing demand response. *The Electricity Journal*, 19(4), 21–31. doi: 10.1016/j.tej.2006.03.006

faezy razi, F., & Shadloo, N. (2017). A Hybrid Grey based Two Steps Clustering and Firefly Algorithm for Portfolio Selection. *Journal of Optimization in Industrial Engineering*, 10(22), 49-59. doi: 10.22094/joie.2017.276

Haider, H.T., See, O.H. & Elmenreich, W. (2016). A review of residential demand response of smart grid. *Renewable and Sustainable Energy Reviews*, 59, 166–178. doi: 10.1016/j.rser.2016.01.016

Han, J., Kamber, M. & Pei, J. (2012). Data Mining,Concepts and Techniques. Third Edition, Morgan Kaufmann, Waltham, MA, USA.

Hassan, N.U., Khalid, Y.I, Yuen,C., Huang,S., Pasha, M.A. et al. (2016). Framework for minimum user participation rate determination to achieve specific demand response management objectives in the residential smart grids. *International Journal of*

Electrical Power & Energy Systems, 74, 91–103. doi: 10.1016/j.ijepes.2015.07.005

Iqbal, Z., Javaid, N., Iqbal, S., Aslam, Sh., Khan, Z.A. et al. (2018). A Domestic Microgrid with Optimized Home Energy Management System. *Energies*, 11(4). doi: 10.3390/en11041002

Javaid, S., Javaid, N., Javaid, M.Sh., Javaid, S., Qasim, U. et al. (2016). Optimal Scheduling in Smart Homes with Energy Storage Using Appliances' Super-Clustering. 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), IEEE Press, Fukuoka, Japan, 342-348. doi: 10.1109/IMIS.2016.130

Jeong, HC., Jang, M., Kim, T., Joo, S-K. (2021). Clustering of Load Profiles of Residential Customers Using Extreme Points and Demographic Characteristics. *Electronics*, 10(3), 290. doi: 10.3390/electronics10030290

Kostková,K., Omelina,L., Kyčina, P., Jamrich, P. (2013). An introduction to load management. *Electric Power Systems Research*, 95, 184–191. doi: 10.1016/j.epsr.2012.09.006

Leroy,Y., Yannou, B. (2018). An activity-based modelling framework for quantifying occupants' energy consumption in residential buildings. *Computers in Industry*, 103, 1-13. doi: 10.1016/j.compind.2018.08.009

Li, M., Li, G.Y., Chen, H.R., Jiang, C.W. (2018). QoE-Aware Smart Home Energy Management Considering Renewables and Electric Vehicles. *Energies*, 11(9). doi: doi.org/10.3390/en11092304

McLoughlin, F., Duffy, A., Conlon, M. A. (2015). Clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied Energy*, 141, 190-199. doi: 10.1016/j.apenergy.2014.12.039

Muratori, M., Rizzoni, G. (2016). Residential Demand Response: Dynamic Energy Management and Time-Varying Electricity Pricing. *IEEE Transactions on Power Systems*, 31(2), 1108-1117. doi: 10.1109/TPWRS.2015.2414880

Olamaei, J., Ashouri, S. (2015). Demand response in the day-ahead operation of an isolated microgrid in the presence of uncertainty of wind power. *Turkish Journal of Electrical Engineering and Computer Sciences*, 23(2), 491-504. doi:10.3906/elk-1301-164

Rajabi,A., Eskandari,M., Jabbari,M., Li, Li, Zhang, J., Pierluigi, S. (2020). A comparative study of clustering techniques for electrical load pattern segmentation, *Renewable and Sustainable Energy Reviews*, 120, 109628. doi: 10.1016/j.rser.2019.109628

Rasheed, M.B., Javaid, N., Ahmad, A., Jamil, M., Khan, Z.A. et al. (2016). Energy Optimization in Smart Homes Using Customer Preference and Dynamic Pricing. *Energies*, 9(8). doi: 10.3390/en9080593

Semeraro, C., Lezoche, M., Panetto, H., Dassisti, M. (2021). Digital twin paradigm: A systematic literature review. *Computers in Industry*, 130. doi: 10.1016/j.compind.2021.103469

Shirazi, E., Jadid, Sh. (2015). Optimal residential appliance scheduling under dynamic pricing scheme via HEMDAS. *Energy and Buildings*, 93, 40-49. doi: 10.1016/j.enbuild.2015.01.061

Si, C., Xu, S., Wan, C., Chen, D., Cui, W. and Zhao, J. (2021). Electric Load Clustering in Smart Grid: Methodologies, Applications, and Future Trends. *Modern Power Systems and Clean Energy*, 9(2), 237-252. doi:10.35833/MPCE.2020.000472

Siano, P. (2014). Demand response and smart grids—a survey. *Renewable and Sustainable Energy Reviews*, 30, 461-478. doi: 10.1016/j.rser.2013.10.022

Tamas, J. (2019). Hierarchical Clustering based on IndoorGML Document. 15th International Scientific Conference on Informatics, IEEE Press, Poprad, Slovakia, 177-182. doi: 10.1109/Informatics47936.2019.9119255

Tambunan, H. B., Barus, D. H., Hartono, J., Alam, A. S., Nugraha, D. A. and Usman, H. H. H. (2020). Electrical Peak Load Clustering Analysis Using K-Means Algorithm and Silhouette Coefficient. International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP), Bandung, Indonesia, 258-262. doi:10.1109/ICT-PEP50916.2020.9249773

Tian, K., Li, J., Zeng, J., Evans, A., Zhang, L. (2019). Segmentation of tomato leaf images based on adaptive clustering number of K-means algorithm. *Computers and Electronics in Agriculture*, 165. doi: 10.1016/j.compag.2019.104962

Wang, Z., Srinivasan, R.S. (2015). Classification of Household Appliance Operation Cycles: A Case-Study Approach. *Energies*, 8(9), 10522-10536. doi: 10.3390/en80910522

Warren, P. (2014). A review of demand-side management policy in the UK. *Renewable and Sustainable Energy Reviews*, 29, 941–951. doi: 10.1016/j.rser.2013.09.009

Zahedi, A. (2011). A review of drivers, benefits, and challenges in integrating renewable energy sources in to electricity grid. *Renewable and Sustainable Energy Reviews*, 15(9), 4775–4779. doi: 10.1016/j.rser.2011.07.074

This article can be cited : Simsar, S., Alborzi, M., Rajabzadeh Ghatari, A., & Yazdian Varjani, A. (2023). Residential Appliance Clustering Based on Their Inherent Characteristics for Optimal Use Based K-Means and Hierarchical Clustering Method. *Journal of Optimization in Industrial Engineering*, *16*(1), 119-127. doi: 10.22094/joie.2023.1975210.2028