Journal of Operation and Automation in Power Engineering
Vol. 11, No. 3, Oct.. 2023, Pages: 182-192
http://joape.uma.ac.ir

Improving the Effect of Electric Vehicle Charging on Imbalance Index in the Unbalanced Distribution Network Using Demand Response Considering Data Mining Techniques

M. A. Baherifard[1], R. Kazemzadeh[1,], A.S. Yazdankhah[1], M. Marzband[2]*

[1]*Faculty of Electrical Engineering, Sahand University of Technology, Tabriz, Iran*
[2] *Northumbria University, Electrical Power and Control Systems Research Group, Ellison Place NE1 8ST, Newcastle upon Tyne, United Kingdom*

Abstract-*With the development of electrical network infrastructure and the emergence of concepts such as demand response and using electric vehicles for purposes other than transportation, knowing the behavioral patterns of network technical specifications to manage electrical systems has become very important optimally. One of the critical parameters in the electrical system management is the distribution network imbalance. There are several ways to improve and control network imbalances. One of these ways is to detect the behavior of bus imbalance profiles in the network using data analysis. In the past, data analysis was performed for large environments such as states and countries. However, after the emergence of smart grids, behavioral study and recognition of these patterns in small-scale environments has found a fundamental and essential role in the deep management of these networks. One of the appropriate methods in identifying behavioral patterns is data mining. This paper uses the concepts of hierarchical and k-means clustering methods to identify the behavioral pattern of the imbalance index in an unbalanced distribution network. For this purpose, first, in an unbalanced network without the electric vehicle parking, the imbalance profile for all busses is estimated. Then, by applying the penetration coefficient of 25% and 75% for electric vehicles in the network, charging/discharging effects on the imbalance profile is determined. Then, by determining the target cluster and using demand response, the imbalance index is improved. This method reduces the number of busses competing in demand response programs. Next, using the concept of classification, a decision tree is constructed to minimize metering time.*

Keyword: Classification, data mining, decision tree, demand response, hierarchical clustering, k-means, electric vehicle, unbalanced distribution network.

1. INTRODUCTION

Data analyzing tools are called data mining to discover valid patterns and relationships that have been unknown until subjected to such an analysis. Time Series Clustering is one of the essential concepts of data mining, based on the idea of grouping and separating an unlabeled time dataset into several separate clusters so that similar sequences fall into the same group [1]. Time-series clustering aims to find hidden patterns, look for similarities and predict the future value of time series data [2]. Data of a time series nature have distinct characteristics over the rest of the data, which challenges their clustering [3]. Time series are long strings of numbers that represent different features of an event at one time. Clustering data into relevant and proper categories is one of the most critical topics in pattern recognition and the artificial intelligence domain. It is crucial to reach the clusters where the data closely resemble each other [4].

Time-series clustering has been applied in a variety of areas, such as in the financial field (for instance for pattern discovery in stock exchange time series) [5], in the medical field (such as for determining brain activity) [6], in psychology (such as in-depth analysis of human behaviour [7], in the field of robotics (such as timely decision-making by systems using data sent by rescue robots) [8], in the field of voice recognition (such as phone monitoring, gender classification in biosecurity applications) [9], in the field of bioinformatics and

Received: 12 Jul. 2021
Revised: 03 Dec. 2021
Accepted: 02 Feb. 2022
*Corresponding author:
E-mail: r.kazemzadeh@sut.ac.ir (R. Kazemzadeh)
DOI: 10.22098/joape.2023.9194.1640
Research Paper

discussion of genes and microarrays (such as improved data mining methods) [10] and in the area of environmental monitoring (such as determining the status of seas and tides) [11] applies.

Data mining concepts have also found many applications in electrical engineering. The authors of [12] use k-means clustering analysis of travel survey data from the UK to identify typical conventional vehicle usage profiles. The resulting clusters were examined, and a smaller EV data set was used to compare the behaviors of the electric and national fleets. The purpose of applying time series clustering can be for pattern discovery, data prediction, and suggestion based on clustering results in solving problems for the proposed areas, as in Ref. [13] and [14]. A clustering method based on the k-shape algorithm is proposed in Ref. [15] to identify shape patterns that are time series sequences. The authors of [15] suggest determining energy consumption patterns in buildings to improve buildings' energy consumption prediction, which reduces the forecast error. In Ref. [16], the technique of locating and determining distributed generation (DG) augmentation based on Loss Sensitivity Factor (LSF) and bus voltage using the k-means clustering method is presented. In Ref. [17], the load curves are clustered by the k-means method, then an algorithm for finding out the appropriate number of clusters is given. Another paper [18] deals with the classification of customer load patterns whose classification scales derive from the properties of k-means clustering results. The authors of [19] used a set of genetic algorithms and k-means clustering. A collection of suitable properties is obtained to construct a binary decision tree to classify the data. The authors of [20] present a model for using the inverter capability of photovoltaic units to provide reactive power. In the proposed model, to reduce the regional effect of reactive power, the study distribution network is clustered into smaller areas using the clustering of the k-means method. The authors of [21] used the k-means method to generate random variables of renewable units. In Ref. [22], a multilayer perceptron-based deep learning model for stability prediction is presented, which simultaneously considers transient stability status and small signal. In other words, the proposed method, which is a proper tool for supervised classification and regression problems in nonlinear and complex systems, can comprehensively predict the angular stability of the rotor after turbulence. Much research has been done in this context; however, the following gaps have been identified:

1. One of the clustering applications in electrical engineering is recognizing behavioural patterns usually implemented on users' consumption. In this paper, the technical behaviour of the imbalance index is clustered.
2. Although the decision tree is used to detect errors or critical areas and disturbances, the target cluster's occurrence is detected to reduce the number of measurements of technical parameters in this paper.
3. Various methods are used to control the imbalance of a network, such as demand response, which is usually applied to the whole network. In this paper, the clustering method is used to reduce the size of target buses to use demand response.

In this paper, two concepts of clustering and data classification have been used. The k-means clustering method and the hierarchical clustering method are used in this paper to investigate the networks' imbalance index behavior. On the other hand, there is a set of pre-defined classes in the data classification, and it is determined based on the object belonging to a particular class. It tries to group a set of objects and find out if there is a relationship between them. Therefore, the relationship between the clusters obtained from clustering is determined using the decision tree. Some more objectives of this paper can be expressed as follows:

1. Data mining concepts such as clustering and classification are used to identify the behavioral pattern of network technical characteristics such as the imbalance index.
2. A part of the network participates in the demand response program by determining the target cluster.

2. DATA MINING [23]

Data mining refers to the extraction of knowledge from a large amount of data; hence, many people use it as a synonym for the word knowledge discovery. The different stages of data mining are shown in Fig. 1. In short, using data analysis tools to discover valid patterns and relationships that were previously unknown is called data mining. The purpose of applying data mining techniques is to find hidden functional patterns in the data. In fact, by employing appropriate data mining algorithms, an attempt is made to acquire new knowledge from the data. Data mining involves techniques that can be used in various scientific fields such as databases, statistics, machine learning, neural networks, and information retrieval. However, the comprehensiveness of these methods and the ability to integrate techniques into them have made data mining popular. The basic idea of data mining is that existing

data contains information that will be used in the future and will be helpful. Data mining aims to find patterns in previous data that clarify needs, preferences, and desires. The fact that patterns are not always explicit, and that the signals received from the data are sometimes vague and confusing, makes things even more difficult. Therefore, data mining is about separating the signs from the useless (so-called) noise; identifying the fundamental patterns in the ventricles of seemingly random variables is critical data mining maps.

Data mining comes in two types, guided and unguided. Guided data mining has a specific and predetermined purpose that seeks a specific pattern. In contrast, unguided data mining aims to find patterns or similarities between sets of information without having a specific purpose or set of pre-defined categories and patterns. A target variable must always be categorized, estimated, or predicted in guided data mining. In unguided data mining, there is no target variable, and the task of data mining is to find general patterns that do not belong to a particular variable. Data mining is mainly concerned with building models. A model is an algorithm or set of rules that relates a set of inputs to a specific purpose or destination. Under the right conditions, a model can lead to the proper insight. Many environmental issues can be incorporated into one of the following six actions to turn a problem into a data-mining problem, which must be turned into a data mining activity, as shown in Fig. 2. These activities are 1- Association rules 2- Sequence 3- Prediction 4- Classification 5- Clustering 6- Visualization.

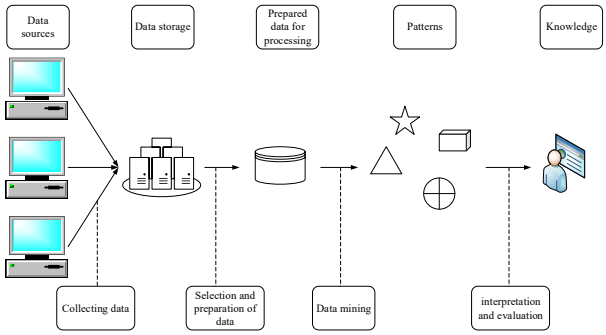

Fig. 1. Knowledge discovery steps

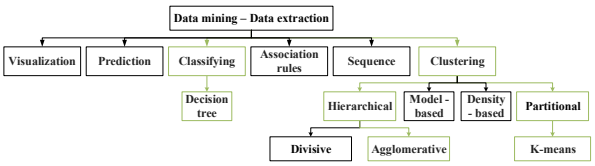

Fig. 2. Classification of data mining

Classification is the process of finding a model that defines and differentiates the categories in the data and using this model to predict the category of records whose category label (target variable) is unknown. In fact, in classification, as opposed to prediction, the goal is to predict the value of a discrete variable. The methods used in forecasting and classification are generally the same. A graphical representation of the relationship between data is called illustration. Data series clustering is one of the essential concepts of data mining. Its central concept is to group and separate an unlabeled data set into several separate clusters, so similar sequences are in the same cluster. In classification based on a model, each data is assigned to a pre-defined category. These categories have been determined through the findings of previous research. However, in the clustering method, there is no pre-defined category, the data is grouped solely based on similarity, and the user determines the titles of each group.

Since the data used in this article are discrete and tabular, the purpose is to discover the patterns of imbalance index profile in a distribution network. Accordingly, in the first step, clustering techniques are used to find patterns. After performing the clustering, because the group labels are specified, the same clusters are reconstructed using the classification (discrete prediction) to use the separation coefficients of the categories into important parameters and effects, which are essential times to metering to calculate imbalance index.

2.1. Clustering

Data clustering aims to find hidden patterns, search for similarities, and predict the future value of data series. Data in series form has different features than other data, which challenges their clustering. Data series are long strings of numbers represening different characteristics of an event. Clustering data into appropriate categories is one of the most important and widely used topics in pattern recognition and artificial intelligence. In this regard, it is essential to reach the categories in which the data are similar. Therefore, it can be said that data may contain complex structures from which even the best data mining techniques cannot extract meaningful patterns. Clustering provides a way to find complex data structures and differentiate uncoordinated competitive signals into their components.

Clustering is dividing a heterogeneous population into several homogeneous subsets or clusters. Clustering is finding the structure in a data set that is not classified. In other words, clustering is the placement of data in groups where the members of each group are similar to

others from a certain angle and have no resemblance to the members of other clusters, or at least less similar to the members of their cluster while Other members have clusters. The similarity criterion here is the distance, and how this distance is calculated is very important in clustering. The distance representing the difference, causes clusters to move in the data space. One of the challenges of clustering is dealing with different data types. Most clustering algorithms are suitable for working with numerical data because the meaning and concept of distance and similarity in this data type can be well defined. However, samples are not recommended only with this data type, and clustering algorithm-based classification must be able to deal with other data types, such as nominal data. The extraction of clusters in any desired shape is also one of the challenges of clustering. The use of distance measurement parameters plays a vital role in determining clusters. Some of these measurement criteria, such as Euclidean distance, lead to finding clusters with a specific shape, such as spherical. However, many data can be placed in clusters of unequal size and density, and it seems very important that clustering algorithms can detect these clusters as well. Clustering methods on static data, i.e. data whose attribute values do not change over time, can be divided into four main categories: 1- Partitional Clustering 2- Hierarchical Clustering 3- Density-Based Clustering 4- Model-Based Clustering.

2.1.1. Partitional clustering

As the name suggests, in the partition-based data clustering algorithm, data is partitioned into multiple clusters. It means that the number of groups is usually specified in this algorithm as inputs, and each data sample can only be a member of a cluster. The similarity criteria evaluate the placement of each sample in a group that the algorithm specifies. These similarity criteria significantly impact the performance of clustering algorithms, and therefore partial clustering techniques greatly emphasize these criteria. One of the partial clustering methods is the k-means algorithm. In the k-means clustering method, the optimisation of an objective function is used. The answers obtained from clustering in this method may be used to minimise or maximise the objective function. It means that if the criterion is the distance between objects, the objective function will be based on minimisation. The result of clustering is to find clusters where the distance between objects in each group is minimal. In contrast, if a similar function is used to measure the similarity of objects, the target function is selected so that the clustering response maximises its value in each cluster.

Since partition-based clustering techniques are limited to finding clusters in n-dimensional space as well as calculating peer-to-peer distances in data series, other techniques that cover this defect should be sought. One of these techniques is the use of density-based clustering or the use of dynamic time warping (DTW) instead of Euclidean distance in the similarity criterion. Many partial methods cluster objects based on their distance towards each other. Some methods only find spherical-shaped clusters and face problems against clusters of arbitrary shapes. In contrast, some other clustering methods have been developed based on density. The general idea of these methods is to develop clusters based on their neighborhood density, meaning that for each data point in a particular cluster, a neighbor with a given radius is considered. This type of clustering is used to smooth disturbances and discover clusters with arbitrary shapes. In useing the DTW algorithm, the goal is to compare two data sets that compute the appropriate distance between the series by swinging the opposite data.

2.1.2. Hierarchical clustering

Unlike partial clustering that divides objects into separate groups, "hierarchical clustering" at any level of distance shows the clustering result. These levels are hierarchical. The "tree" style is used to display hierarchical clustering results. This method is an effective way to display hierarchical clustering results. Hierarchical clustering methods are divided into two categories, which are shown in Fig. 3.

2.1.2.1. Agglomerative hierarchical clustering

The approach of this method is "bottom-up" so that agglomerative clustering starts with one cluster per observation. At each stage, the two clusters that have minuscule differences are combined and form a more massive cluster. This method causes a smaller cluster to exist at a higher level. At each level, this aggregation continues until the number of clusters reaches the desired number.

2.1.2.2. Divisive hierarchical clustering

The approach of this method is "top-down" so that divisive clustering starts with a cluster that includes all observations and then recursively divides one of the existing clusters into smaller clusters at each stage. Different algorithms, such as k-means, can be used to break down each cluster. In general, any clustering algorithm that produces at least two output clusters can be used in decomposing clustering. Since in the divisive hierarchical clustering method, the same k-means

method is used for separation, it is accordingly not suitable for comparison with the k-means clustering method, so the agglomerative hierarchical method has been used. On the other hand, because using Euclidean distance in k-means is a challenge, it is better to use DTW methods to compute the distance in the data series or density-based methods in n-dimensional data. Since the data does not have the distribution density and are given a series of data, there is no need to use the model-based method and density-based method.

2.2. Classification

The purpose of data classification is to organize and allocation of data to separate classes. A primary model is created based on a training data set in this process. This model is used for the classification of new data. Therefore, new data belonging to the specific category is predictable using the model obtained. In other words, the classification includes reviewing the characteristics of a new object and allowing it to be one of the predetermined sets. One division of data mining is the classification that operates using the If-Then rule. Its purpose is to predict a new data attribute based on other attributes known as a predictor. In classification, the data are divided into two categories of training and testing and can be extracted by providing training data mining algorithms. Data mining algorithms should place the objective feature and the number of forecasting properties. In the classification process, two steps must be implemented: In the first step, the model algorithm should be made, and in the second step, the data should be predicted. KNN algorithms, SVM, decision tree and ANN, are among the training-based classification algorithms.

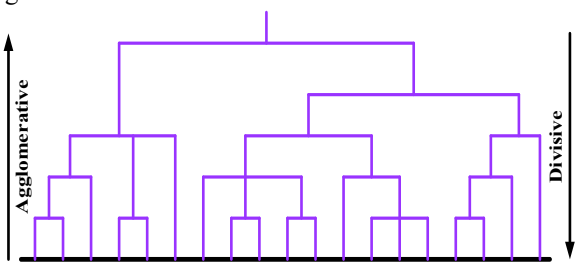

Fig. 3. Hierarchical clustering methods

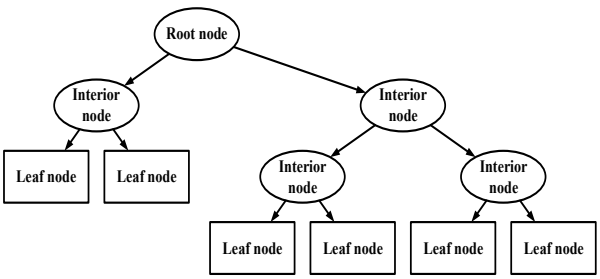

Fig. 4. Outline of the decision tree

2.2.1. Decision trees

The decision tree is one of the data mining models for the classification execution. A particular tree structure is built to assist in decision-making and divides a large set of data collected into smaller sets of data chains. This practice is done through a series of simple decision-making rules. In each consecutive division, the members of the resulting collections are more similar than ever. Decision trees perform classification by separating samples from the root of the tree to its leaves. Each node on the tree specifies an experiment that includes several inputs for the properties of the samples, and each branch coming out of it represents one of the possible values of that property. In other words, first, a sample is classified by being in the tree's root using the specific feature for that node. Branches are considered in accordance with the possible values of that characteristic, and other samples are classified in each branch in the same way. In binary decision trees, each sample can be assigned two values of yes or no. the decision tree has three main components and is shown in Fig. 4. Root Nodes: The starting node of a tree is called the root. Leaf Nodes: The leaf contains questions and criteria that must be answered. Branches: are like arrows that connect nodes.

The advantage of using a decision tree algorithm compared to other data mining algorithms is that its model is faster and easier to interpret. Predictions are more effective and efficient based on a decision tree. The decision tree algorithm is a classification and regression model generation algorithm provided by analytical services for use in predictive models of discrete and continuous characteristics. The algorithm uses the linear regression method to determine where the decision tree is divided regarding continuous characteristics. The decision tree among classification algorithms is a powerful method whose popularity increases with the growth of data mining. The decision tree is used to approximate discrete functions. It is resistant to noise input data and is efficient for high volume data; hence it is used in data mining. So far, the decision tree has been used as a data mining method to extract hidden patterns in medical data. Data mining is a significant branch of a more in-depth understanding of medical data that seeks to solve problems in diagnosing and treating diseases.

In contrast, disadvantages can also be pointed out alongside these advantages, such as not matching the continuous characteristics in the decision tree. This

structure can only work with features that include discrete values (with limited numbers). Many methods have been proposed to divide continuous amounts into smaller clusters to solve this problem. In addition, instead of using each feature's continuous values, the clustering characteristic that contains this value will be used in a decision tree structure for decision-making.

3. SUGGESTED FRAMEWORK

Fig. 5 shows the general framework of project implementation, which includes four steps as follows:

1) In the first stage, residential and commercial loads are connected to an unbalanced distribution network. Then, by adding electric vehicles to the network with 25% and 75% penetration scenarios, unbalanced load flow is implemented for 24 hours. Next, using the obtained results, the imbalance index profile is calculated.
2) Considering that one of the methods to control the imbalance in the distribution network is power injection, so in the next step, using the demand response method, the unbalanced control of the network will be addressed. Also, the incentive load response method is used for the control to have a reasonable rate. At this stage, all busloads of the network are involved in demand response.
3) In the third stage, the distribution network buses are divided into several categories by comparing the imbalance index profiles and using the clustering technique. In this step, two techniques, k-mean, and hierarchy, are applied. Then, by designating one of the clusters as the target cluster, only the buses present in that cluster participate in the demand response program.
4) A decision tree is constructed since all the buses of the distribution network have a specific label using the clusters obtained in the previous step. Building a decision tree will reduce network metering time.

4. PROBLEM FORMULATION

4.1. K-means clustering

In the k-means method, clustering operations are performed based on n-observations and k groups. Thus, the number of clusters or groups in this algorithm is already known. Each object will belong only to one cluster during the partial clustering process, and no cluster will remain without a member. If l_{ab} indicates the status of x_a belonging to cluster c_b, it only accepts values of 0 or 1. This case is shown by (1) and (2):

$$l_{ab} = 1, \ \ x_a \in c_b \tag{1}$$

$$l_{ab} = 0, x_a \notin c_b \tag{2}$$

However, these rules change when the "fuzzy clustering" method is used, and the membership dignity is used to indicate that each object belongs to each cluster. Thus, the membership rate of x_a to cluster c_b is between 0 and 1 as (3):

$$l_{ab} \in [0,1] \tag{3}$$

Partial algorithms usually operate based on optimising an objective function. Partial algorithms usually operate based on the optimization of an objective function based on the repetition of steps from different optimization algorithms. For example, the k-means algorithm operates by specifying the objective function based on the average distance of each cluster's members relative to their average. This algorithm places objects in clusters so that the average sum of squares of distances in clusters has the lowest value. If the observations are denoted by x, the k-means algorithm's objective function can be written as (4):

$$\kappa = \sum_{b=1}^{k} \sum_{x \in c_b} d\left(x, \mu_{c_b}\right) \tag{4}$$

where μ_{c_j} is the mean of clusters c_j, and d is the square of the Euclidean distance.

4.2. Decision tree

One of the tasks of data mining is classification. Classification employs various techniques that are used in different researches. The most common classification techniques are K- nearest neighbor, decision trees, neural networks, backup vector machine, Bayes classification, regression, coarse seed set theory, state-based logic, expert systems, fuzzy logic and genetic algorithms. Decision trees are prevalent due to their simplicity and high comprehensibility. This technique falls into the category of classification trees. Classification trees predict the values of the properties of dependent and discrete variables. Classification trees predict the values of dependent and discrete variables. Decision trees only predict discrete properties. A class variable makes this prediction, also called target feature or dependent feature. Equation (5) is used to make a decision tree [24].

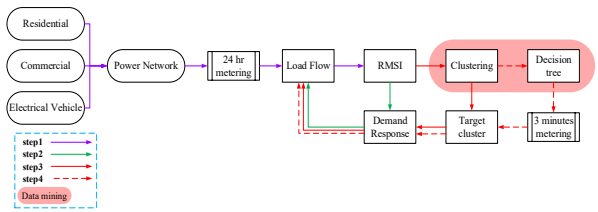

Fig. 5. Structure of proposed framework

$$\beta = \sqrt{\frac{\sum_{z=1}^{h}\left(N_z - \frac{N}{h}\right)^2}{\left(\frac{N}{h}\right)^2}} \tag{5}$$

in which β is the balance coefficient between

classification groups, N is the number of input patterns in the decision tree, Nz is the number of the patterns included the z-th node and h is the number of branches separated from each node, which in this article is 2 due to the binary decision tree. The fitness function that includes the balance coefficient is defined as (6):

$$F = \frac{1}{1+\omega.E+\varpi.\beta} \quad (6)$$

where E is a classification error between two groups; ω is the weighting factor of the classification error, and ϖ is the weighting factor for β. These coefficients indicate the importance of the parameters in separating the groups. If both the classification error and the balance coefficient have, a value of 0, F has the highest possible value of 1.

4.3. Dynamic Time Warping (DTW)

In this section, dynamic time warping is defined for two data series O and S with lengths $|O|$ and $|S|$. The components of these two data series are the following Equations (7) and (8).

$$O = o_1, o_2, \dots, o_i, \dots, o_{|O|} \quad (7)$$

$$S = s_1, s_2, \dots, s_i, \dots, s_{|S|} \quad (8)$$

The warp path (i.e. W) is represented by (9), where G is the length of the warp path and its value is in the range shown in (10). The g-th element of the warp path is defined according to (11). In this regard, i is an index of the O data series and j is an index of the S data series.

$$W = w_1, w_2, \dots, w_G \quad (9)$$

$$\max(|O|, |S|) \le G \le |O| + |S| \quad (10)$$

$$w_g = (i, j) \quad (11)$$

The warp path must start at the beginning of both data sets, $w_1 = (1,1)$ and continue to the end of both data sets, $w_g = (|O|, |S|)$. This ensures that all indices of both data sets participate in the warp path. Equation (12) indicates the same thing.

$$w_g = (i, j), w_{g+1} = (i', j') \quad i \le i' < i+1, j \le j' < j+1 \quad (12)$$

The optimal warp path is the minimum path in which the distance of the warp path W is defined according to (13).

$$\lambda(W) = \sum_{g=1}^{g=G} \lambda(w_{gi}, w_{gj}) \quad (13)$$

$\lambda(W)$ is the distance (usually the Euclidean distance) of the warp path W and $\lambda(w_{gi}, w_{gj})$ is the distance between two data indices (one from the O series and the other from the S series). The goal of the dynamic time-warping algorithm is to find the shortest path for warping the O data series into the S data series. To solve this problem, it is necessary to construct a two-dimensional cost matrix D of size $|O|$ in $|S|$, in which $D(i,j)$is the optimal warp path that can be constructed from two sets of data, $O = o_1, o_2, \dots, o_i, \dots, o_{|O|}$ and $S = s_1, s_2, \dots, s_i, \dots, s_{|S|}$. Therefore, the value in $D(|O|, |S|)$ is the minimum warp distance between the two data sets O and S. To calculate the minimum route size, all the cost matrix cells must be filled. Since the value $D(i,j)$ is the length of the optimal path between the two data sets with length i and j, the optimal path length for the values prior to i and j in both sets of data is specified . Therefore, according to (14) , $D(i,j)$ will be equal to the length of the minimal paths in one point less than i and j in addition to the distance between two points O and S.

$$D(i,j) = dist(i,j) + \min[D(i-1,j), D(i,j-1), D(i-1,j-1)] \quad (14)$$

4.4. Unbalace indices

Different indices are defined in various standards to determine the distribution network imbalance. The authors of [25] have proposed a new definition for calculating current imbalances by considering the 0 sequences shown in (15).

$$RMSI = \sqrt{\frac{|I_0|^2 + |I_2|^2}{|I_1|}} \quad (15)$$

In distribution systems, when an imbalance occurs, the power loss difference is imposed on the network, which is obtained from (16).

$$\Delta P = \frac{(RMSI)^2}{1+(RMSI)^2} \times Uloss \quad (16)$$

which $Uloss$ is unbalanced power losses obtained in power flow.

4.5. Demand response

The authors of [26] have proposed a demand response model that sets the incentive rate for each consumer in exchange for changing their load. This model considering the average demand elasticity for household consumers equal to -0.11 and commercial consumers to be -0.14, is shown in (17).

$$\pi_{inc} = \frac{-\Delta P}{El.P_0}\pi_0 \quad (17)$$

where π_{inc} is incentive price to change ΔP in consumers demand while the base demand of P_0 at electricity price of π_0.

5. DISCUTION AND RESULTS

The system studied in this paper is an IEEE-37 bus unbalanced distribution network [27]. The network has 25 buses with commercial and residential loads, as shown in Fig. 6, and the information on this network is given in Table 1. Buses 713, 720, 722, 730, 733, and 734 are considered commercial

buses connected to the commercial electric vehicle parking lots (CPL).

Table 1. IEEE-37 bus unbalanced distribution network's load data

Node	Spot Loads						
	Load Model	Ph-1 kW	Ph-1 kVAr	Ph-2 kW	Ph-2 kVAr	Ph-3 kW	Ph-3 kVAr
701	D-PQ	140	70	140	70	350	175
712	D-PQ	0	0	0	0	85	40
713	D-PQ	0	0	0	0	85	40
714	D-I	17	8	21	10	0	0
718	D-Z	85	40	0	0	0	0
720	D-PQ	0	0	0	0	85	40
722	D-I	0	0	140	70	21	10
724	D-Z	0	0	42	21	0	0
725	D-PQ	0	0	42	21	0	0
727	D-PQ	0	0	0	0	42	21
728	D-PQ	42	21	42	21	42	21
729	D-I	42	21	0	0	0	0
730	D-Z	0	0	0	0	85	40
731	D-Z	0	0	85	40	0	0
732	D-PQ	0	0	0	0	42	21
733	D-I	85	40	0	0	0	0
734	D-PQ	0	0	0	0	42	21
735	D-PQ	0	0	0	0	85	40
736	D-Z	0	0	42	21	0	0
737	D-I	140	70	0	0	0	0
738	D-PQ	126	62	0	0	0	0
740	D-PQ	0	0	0	0	85	40
741	D-I	0	0	0	0	42	21
742	D-Z	8	4	85	40	0	0
744	D-PQ	42	21	0	0	0	0
Total		727	357	639	314	1091	530

Only commercial and residential loads are connected to the network in the study network, and CPLs not considered. In this case, unbalanced load flow is performed for 24 hours. Later, the bus voltages obtained by the Fortescue matrix are converted into positive, negative, and zero sequences. Then the current imbalance index (RMSI) is calculated. The steps are performed again by connecting the CPLs with a penetration coefficient of 25% and 75% capacity of commercial buses. The obtained RSMI profile for all three modes is shown in Fig. 7.

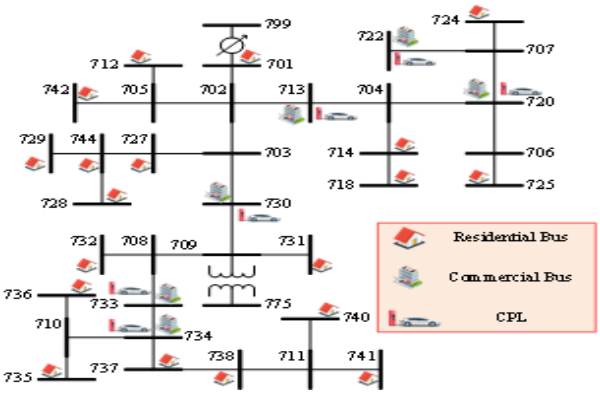

Fig. 6. Case study: IEEE-37 bus unbalanced distribution network

As shown in Fig. 7, after connecting the CPL to the power grid, the RMSI imbalance index increases between 10 to 12 and 14 to 16. The maximum presence of EV in CPL is from 8 to 20. Comparing RMSI when the penetration rate of EVs in the network is 25% compared to the penetration rate of zero percent, it can be noticed that as the number of EVs increases, RMSI becomes more unbalanced. Then, after all the network buses participate in the DR program, the imbalance index profile is shown in Fig. 8. As illustrated in Figure 8, the RMSI index improves during the CPL's entry and exit hours.

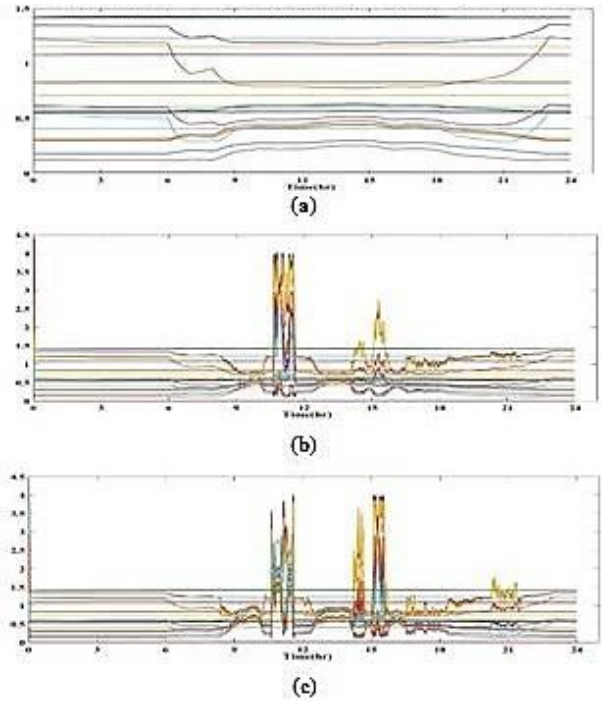

Fig. 7. Daily RMSI profiles with EV penetration rate of 0% (a), 25% (b) and 75% (c)

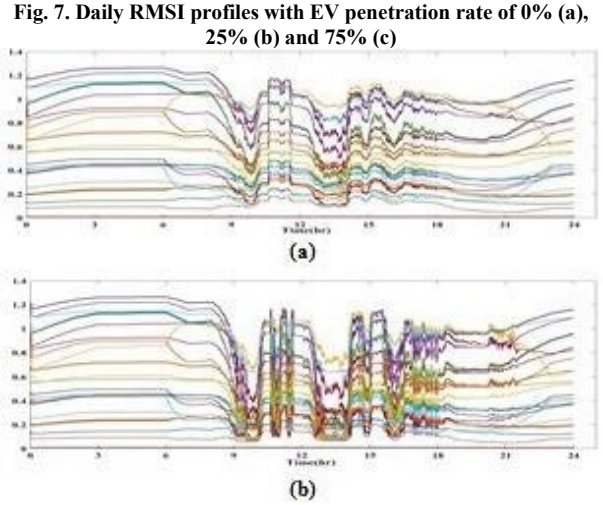

Fig. 8. RMSI profile with participation of 100% of all busses in DR with EV penetration rate of 25% (a) and 75% (b)

In the next step, two typical clustering methods are used to ensure the accuracy of the results. Therefore, the RMSI profiles of all network buses are divided into three categories using hierarchical clustering and k-means techniques considering DTW as distance measurement for similarity, which are shown in Fig. 9

for a 25% EV penetration coefficient.

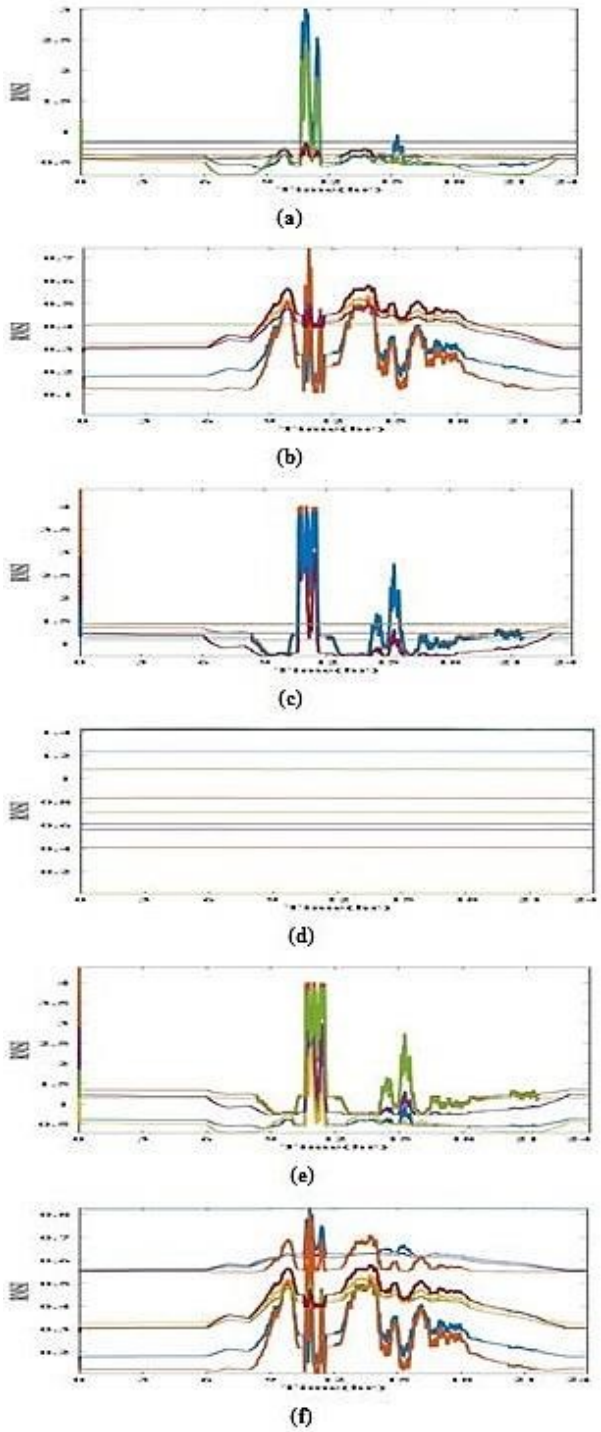

Fig. 9. RMSI clustering with hierarchical methode: cluster 1(a), cluster 2(b) and cluster 3(c) and k-means methode: cluster1(d), cluster 2(e) and cluster 3 (f) with EV penetration rate of 25%

As shown in Fig. 9 , the RMSI profiles categorized by the k-means method in each cluster have more appropriate similarities than the hierarchical technique. After performing clustering, it is visually understandable that cluster 3 in the hierarchical method and cluster 2 in the k-means method have the greatest effect on the imbalance index after connecting the EVs to the network. Therefore, according to the hierarchical method, 14 buses (714, 718, 720, 722, 724, 725, 729, 731, 732, 732, 735, 736, 740, 741, 742) participate in the demand response program. While according to the k-mean method, five busses (704, 707, 713, 720, 722) participate in DR as target busses. By increasing the penetration coefficient of EVs in the network to 75%, based on the k-means method, four buses (704, 707, 720, 722) are selected for the DR program. Then the k-means method is selected as the most appropriate method for clustering.

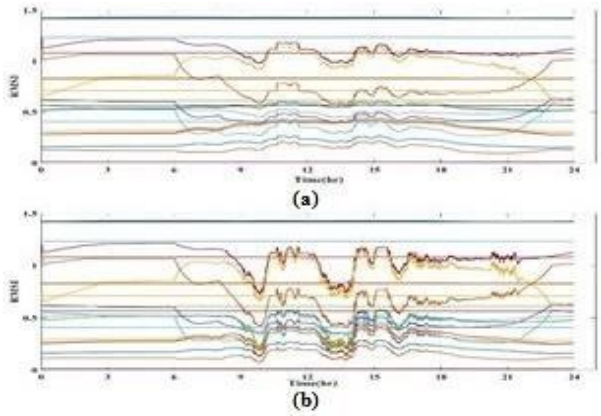

Fig. 10 RMSI profile with participation of busses of cluster 2 in DR with EV penetration rate of 25% (a) and 75% (b)

Furthermore, the DR program runs only for cluster 2. The imbalance profile for the two distribution modes of 25% and 75% is shown in Fig. 10. It is also noteworthy that some buses with CPL, which wrosens the imbalance index, have been selected as the target bus for DR. For example, in the k-mean method, only 720 and 722 buses are included in the DR among commercial buses.

As shown in Fig. 10, despite the decrease in the number of busses participating in the DR, the daily profile of the RMSI index improves. On the other hand, using clustering and reducing the number of busses reduces the power required to run the DR program, which is shown in Fig. 11.

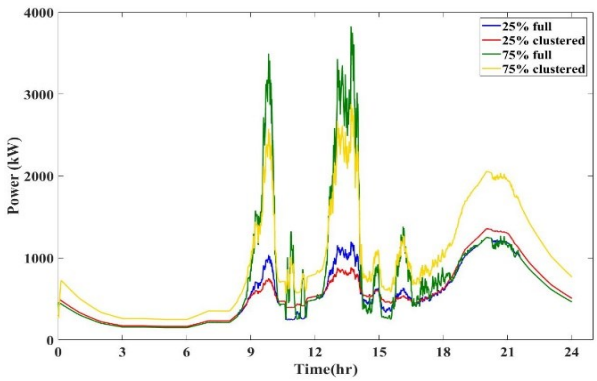

Fig. 11. Power injected to grid by DR

As shown in Fig. 11, as the EV penetration in the

network increases, the power required to control the RMSI is increased by the DR program. Also, when EVs are in the CPL, the power required to run by the DR program is reduced by performing clustering compared to the participation of all buses. Then, after labeling the network buses based on the clusters obtained by the k-means method, the decision tree is constructed according to Fig. 12.

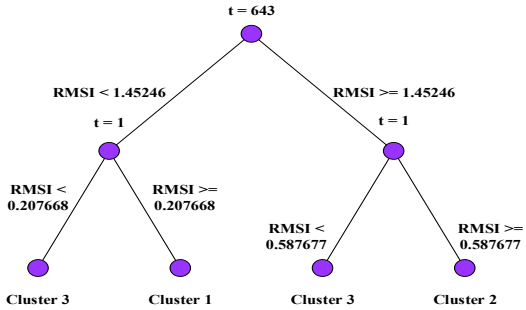

Fig. 12. Decision tree for clusters achieved from k-means method

Based on Fig. 12, it is clear that only two measurements per day are needed to find the right cluster for each profile. At the same time, it is clear that to have an RMSI profile and perform the necessary calculations, 24 hours of consumer metering is required. The error of the constructed decision tree is equal to 14%. First, the RMSI is calculated at minutes 643, and if it is less than 1.45246, the left branch path is selected. Then in minute 1, if the RMSI is minor than 0.207668, the RMSI profile will be a member of cluster 3, and if it is more than 0.207668, it will be a member of cluster 1. The right path also requires RMS calculations in 1 minute. If it is less than 0.587677, it will be a member of cluster 3; otherwise, it will be a member of cluster 2.

6. CONCLUSIONS

In the proposed paper, first, the RMSI imbalance index, which is calculated using current sequences, is obtained for a non-equilibrium IEEE-37 bus distribution network without the presence of electric vehicles. Then, by increasing the penetration of EVs by 25% and 75%, the RMSI profile of the network is calculated. Then, to control the network imbalance, the DR-based incentive program participates in all network buses. Then two methods of hierarchical clustering and k-means are used to separate RMSI profiles into three different groups. The results show that the number of basses in the target cluster based on the k-means method is less than the hierarchical method. Also, the similarity of RMSI profiles classified in each cluster in the k-means method is more than the hierarchical method. The following results are obtained by participating the target cluster member bass in the DR program:

1) Reduce the number of buses participating in the DR program in order to control the network imbalance
2) Reduce the amount of power exchanged with the network in the DR program compared to the case where all buses participate in the DR.
3) Build a decision tree to reduce the number of RMSI metering.

REFERENCES

[1] S. Singh and A. Yassine, "Big data mining of energy time series for behavioral analytics and energy consumption forecasting", *Energies*, vol. 11, 2018.
[2] L. Ruiz et al, "A time-series clustering methodology for knowledge extraction in energy consumption data", *Expert Syst. Appl.*, vol. 160, p. 113731, 2020.
[3] A. Saas, A. Guitart, and A. Perianez, "Discovering playing patterns: Time series clustering of free-to-play game data", I*EEE Conf. Comput. Intell. Games,* 2016.
[4] T. Räsänen and M. Kolehmainen, "Feature-based clustering for electricity use time series data", *Lecture Notes Comput. Sci.*, 2009.
[5] M. Wen et al, "Stock market trend prediction using high-order information of time series", *IEEE Access*, vol. 7, pp. 28299-308, 2019.
[6] A. Wismüller et al., "Cluster analysis of biomedical image time-series", *Int. J. Comput. Vis.*, vol. 46, no. 2, pp. 103-28, 2002.
[7] V. Kurbalija et al, "Time-series mining in a psychological domain", *ACM Int. Conf. Proc. Series*, pp. 58-63, 2012.
[8] R. Gopalapillai, D. Gupta, and T. Sudarshan, "Experimentation and analysis of time series data for rescue robotics", *Adv. Intell. Syst. Comput.*, vol. 235, pp. 443-53, 2014.
[9] S. Fong, "Using hierarchical time series clustering algorithm and wavelet classifier for biometric voice classification", *J. Biomed. Biotechnol.*, vol. 2012, 2012.
[10] N. Subhani et al, "Multiple gene expression profile alignment for microarray time-series data clustering", *Bioinformatics*, vol. 27, no. 13, pp. 2281-88, 2011.
[11] M. Scotto, A. Alonso, and S. Barbosa, "Clustering time series of sea levels: Extreme value approach", *J. Waterw. Port, Coast. Ocean Eng.*, vol. 136, no. 4, pp. 215-25, 2010.
[12] C. Crozier, D. Apostolopoulou, and M. McCulloch, "Clustering of usage profiles for electric vehicle behaviour analysis", *2018 IEEE PES Innov. Smart Grid Tech. Conf. Europe,* 2018.
[13] A. Sfetsos and C. Siriopoulos, "Time series forecasting with a hybrid clustering scheme and pattern recognition", *IEEE Trans. Syst. Man, Cybern. Part ASystems Humans.*, vol. 34, no. 3, pp. 399-405, 2004.
[14] W. He et al, "A new method for abrupt dynamic change detection of correlated time series", *Int. J. Climatol.*, vol. 32, no. 10, pp. 1604-14, 2012.
[15] J. Yang et al., "k-Shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement", *Energy Build.*, vol. 146, pp. 27-37, 2017.
[16] O. Penangsang, D. Putra, and T. Kurniawan, "Optimal placement and sizing of distributed generation in radial distribution system using K-means clustering method", *2017 Int. Seminar Intell. Tech. Its Appl.,* 2017.

[17] Y. Ji et al, "Clustering algorithm of power load curves in distribution network based on analysis of matrix characteristic roots", *Asia-Pacific Power Energy Eng. Conf.*, 2016.

[18] P. Liu et al., "Hybrid Features based k-means clustering algorithm for use in electricity customer load pattern analysis", *Chinese Control Conf.*, 2018.

[19] K. Kim et al, "Binary decision tree using k-means and genetic algorithm for recognizing defect patterns of cold mill strip", pp. 341-350, 2004.

[20] H. Yousefi, S. Gholamian, and A. Zakariazadeh, "Distributed voltage control in distribution networks with high penetration of photovoltaic systems", *J. Oper. Autom. Power Eng.*, vol. 8, no. 2, pp. 164-71, 2020.

[21] H. Shayeghi and M. Alilou, "Multi-objective demand side management to improve economic and environmental issues of a smart microgrid ", *J. Oper. Autom. Power Eng.*, vol. 9, no. 3, pp. 182-92, 2021.

[22] M. Shahriyari and H. Khoshkhoo, "A deep learning-based approach for comprehensive rotor angle stability assessment ", *J. Oper. Autom. Power Eng.*, vol. 10, no. 2, pp. 105-12, 2022.

[23] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Elsevier Inc., 2012.

[24] F. García Márquez, I. Segovia Ramírez, and A. Pliego Marugán, "Decision making using logical decision tree and binary decision diagrams: a real case study of wind turbine manufacturing", *Energies*, vol. 12, no. 9, 2019.

[25] M. Tavakoli Bina and A. Kashefi Kaviani, "Erratum to 'Three-phase unbalance of distribution systems: Complementary analysis and experimental case study", *Int. J. Electr. Power Energy Syst.*, vol. 35, no. 1, p. 201, 2012.

[26] A. Niromandfam, A. Yazdankhah, and R. Kazemzadeh, "Modeling demand response based on utility function considering wind profit maximization in the day-ahead market", *J. Clean. Prod.*, vol. 251, p. 119317, 2020.

[27] M. Baherifard et al., "Intelligent charging planning for electric vehicle commercial parking lots and its impact on distribution network's imbalance indices", *Sustain. Energy, Grids Net.*, vol. 30, p. 100620, 2022.