Electric Power Systems Research 141 (2016) 191–201

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Classification for consumption data in smart grid based on forecasting time series

Kálmán Tornai [a,*], Lóránt Kovács [b], András Oláh [a], Rajmund Drenyovszki [b], István Pintér [b], Dávid Tisza [a], János Levendovszky [c]

[a] Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, Hungary
[b] Kecskemet College, Department of Informatics, Hungary
[c] Department of Networked Systems and Services, Budapest University of Technology and Economics, Hungary

ARTICLE INFO

Article history:
Received 21 September 2015
Received in revised form 16 June 2016
Accepted 18 July 2016

Keywords:
Time series forecast
Feedforward Neural Network
Classification methods
Clustering methods

ABSTRACT

One of the most important tasks of present day smart grid implementations is to classify different types of consumers (households, office buildings and industrial plants) because they may be served by the power supplier with different parameters, rates, contracts.

In this paper, we propose a novel classification scheme for smart grid systems where the collected data are processed in order to increase the efficiency of electricity transportation as well as demand-supply management. The new scheme is based on forecasting the consumption time series obtained from a smart meter. Class assignment is determined using the forecast error. Different linear and non-linear methods were tested based on the corresponding assumptions on the statistical behavior of the underlying consumption time series.

Performance tests were carried out with simulations in order to demonstrate the capabilities and to compare the achieved performance of the proposed scheme with existing solutions. The simulations have been executed using (i) artificially generated consumption data, which data came from a bottom-up semi-Markov model and (ii) real, measured power consumption data as well. The parameters of the model have been validated on real data. The numerical results have demonstrated that our method can better model and classify the consumption patterns of office-buildings than the existing methods. As a result, the proposed method may prove to be a promising classification tool.

1. Introduction

Smart grids can be seen as a new vision about the future evolution of power systems. In contrast to the present system, a smart grid can efficiently incorporate the renewable energy sources and can adaptively manage the balance between supply and demand. The new features are based on two-way communications and intelligent measurement as an integrated part of the energy system. In the presence of smart meters a huge amount of data will be acquired. The information can only be obtained by sophisticated signal processing algorithms. Thus there are several possibilities of applying the system in automatic identification of consumer categories, identification of outliers in certain groups, and identification of the misuse of services.

This paper deals with the automated classification of consumers which is the necessary tool to recognize category changes, consumption behavior changes or responsiveness to real time prices. Before starting the classification process the underlying time series have to be broken into stationary segments. In this paper, we focus on stationary time series and we refer only to the literature in order to detect the statistical change points [1].

The proposed method can be directly applied in Demand Side Management (DSM). The goal of DSM is manipulation of load shape and it aims to reduce the peak [2]. DSM commonly refers to programs organized by the utility that affect the amount and timing of consumer use (e.g. real time pricing (RTP) and direct load control (DLC)). The classification of load curves is very important in the implementation of DSM policies because it allows the optimization of system management from the selection of the most appropriate actions for each set of the loads with the same characteristics [3,4]. Our method can ensure a more accurate analysis of consumer behavior which is needed to improve the DSM programs.

As a result our objective is to develop a new classification scheme that is capable of classifying power consumers based on their consumption patterns. The classes are supposed to originate from having a different mixture of appliances. The novelty in our

* Corresponding author at: Address: P.O. Box 278, Budapest H-1444, Hungary.
E-mail address: tornai.kalman@itk.ppke.hu (K. Tornai).

http://dx.doi.org/10.1016/j.epsr.2016.07.018
0378-7796/

scheme is the usage of the error level of forecasting of consumption time series, obtained from different classes of consumers as a decision variable. As a result, due to the more sophisticated modeling and forecasting of the underlying consumption time series, the proposed scheme is able to distinguish among time series differing only in higher order statistical descriptors but having the same at the level of expected value or variance.

Three different types of methods will be tested, such as Linear Approximation (LA), Radial Basis Function Neural Network (RBF) and Feedforward Neural Network (FFNN) from which the FFNN will be proven to be the most promising tool. This new forecast-based classification scheme will be compared to other existing solutions.

The different methods will be tested by artificially generated consumption data as follows. To model consumption time series two different approaches were used: (i) as an initial level the consumption data is supposed to be an autoregressive process; (ii) as a more realistic scenario consumption data has been generated by a bottom-up model. In the second case the total consumption of building is generated as the sum of independent two-state semi-Markov models due to the lack of enough, *well-annotated*[1] measured power consumption data. These semi-Markov processes represent all the appliance types located in the modeled buildings. The parameters of the semi-Markov model were fitted on measured power consumption data. In this way we construct a huge amount of realistic test data.

The determination of the consumer classes is out of scope of this paper, however the scheme of the method proposed in this paper implicates the possibility of an automated clustering method where there is no prior knowledge about the classes of consumers. The clustering capability of the proposed method is a target of further research as it requires additional modifications and algorithmic steps.

The paper will be organized as follows: in Section 2 existing and referred works are summarized. In Section 3, the basic concept of the clustering scheme and the applied model of the system will be introduced. In Section 4, the performance of the different methods will be analyzed, finally in Section 5, conclusions will be drawn.

2. Related works

In this section we overview the existing classification and clustering methods, which will be compared to the proposed scheme in the numerical analysis. Also the forecast methods for time series are briefly summarized.

2.1. Classification methods

In the followings we summarize the existing approaches for data classification in order to have a firm basis for comparing and introducing our new scheme. The algorithms are categorized as (i) distance based classification; (ii) feature based classification; and (iii) model based classification. Algorithms, which are designed for classification of time series have the following two approaches: (i) redesigning and modifying existing algorithms in order to handle the sequential data where the aim is to exploit the sequential nature of the data; (ii) the sequential data is transformed to fit to the existing algorithms [5].

2.1.1. Distance based classification

In case of distance based classification the decision depends on the distance between data elements. For a new value the distances between the existing values and new measurement have to be calculated. The new measurement is assigned to the class where the values with the smallest distance belong to. The performance of the method is highly influenced by the applied metric (measure of distance) [6]. In most cases the Euclidean distance is used, but special problems require more complicated metrics such as dynamic time warping distance [7,8]. This method is applicable to sequences with non-numerical values, but the data has to be transformed or sequence alignment (such as Needleman–Wunsch) has to performed.

2.1.2. Feature based classification

The feature based method, such as artificial neural networks (ANN) and decision trees, transform the sequential data into a feature-set. The performance of these methods depends on the selection of the appropriate features [9]. The most common transformation is the wavelet decomposition of pattern matching.

The number of dimensions of data has to be reduced by the transformation into frequency domain such as DFT (Discrete Fourier Transform), DWT (Discrete Wavelet Transform), SVD (Singular Value Decomposition) [10]. In addition, kernel methods such as Support Vector Machines (SVMs) can be used to reduce the number and to extract the features [11].

2.1.3. Model based classification

The model based classification methods construct a model for all classes using training sets. The incoming data is classified upon the best fitting model. These methods can be divided into statistical and neural network based algorithms [5].

The most popular statistical models are the: (i) Gaussian, (ii) Poisson, (iii) Markov Model and (iv) Hidden Markov Model where the parameters of the models are fitted to the probability distribution of the data [12]. The models also can be predictive and descriptive [13], where the predictive model tries to estimate the unavailable data from the existing ones; while in the descriptive case the patterns and relationships in the data are mined.

The ANNs, especially recurrent neural networks (RNN) are close to the statistical models [14] however, they do not require a priori knowledge about the data. The main benefit of using this method is that the input noise has less influence on the performance [15].

2.2. Clustering methods

While preliminary tests suggest that our method is capable of using for clustering the time series we briefly summarized the clustering methods and their categorization based on [16,17].

Hierarchical clustering algorithms organize data into a hierarchical structure. The results of clustering can usually be depicted as a dendrogram of binary tree. These methods ban be further divided into two groups (i) as agglomerative methods and (ii) divisive methods. The agglomerative approach initially has as many clusters as the number of data elements. When the distance between two clusters is small enough, then the clusters are merged. The divisive clustering proceeds in an opposite way as all the objects belong to one cluster, and the existing clusters are divided into smaller groups [18].

Clustering methods based on squared error assign a set of objects to clusters without any hierarchical structures. The optimal partition can be found by enumerating all possibilities however, the exhaustive search is unfeasible because of the computational complexity [19]. As a result, heuristic algorithms have been introduced to seek approximate solutions. The k-means algorithm is the most widely used squared-error based method [20].

In case of *Mixture densities-based clustering* the data which have to be clustered are assumed to be generated by several probability distributions. If these distributions can be determined then the

[1] Information of location, consumer behavior and the changes of them are unknown.