Expert Systems With Applications 54 (2016) 105–120

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Application of hybrid computational intelligence models in short-term bus load forecasting

Ioannis P. Panapakidis*

Department of Electrical Engineering, Technological Education Institute of Thessaly, 41110 Larisa, Greece

ARTICLE INFO

Keywords:
Artificial neural networks
Bus load forecasting
Load modeling
Time-series clustering

ABSTRACT

Artificial neural networks (ANNs) are a favorable scheme in load forecasting applications mainly due to their endogenous capacity of robust modeling of data sets with highly non-linear relationship between inputs and outputs. Usually, the inputs correspond to historical load values, exogenous variables like temperature, day type identification codes and others. The outputs refer to the load values under examination. The majority of the load forecasting related literature focuses in aggregated load system level. While contemporary research efforts focus in smart grid technologies, there is need to study the characteristics of small scaled loads. Bus load forecasting refers to prediction of the demand patterns in buses of the transmission and distribution systems. Bus load exhibits low correlation with the aggregated system load, since it is characterized by a high level of stochasticity. Hence, a proper selection and formulation of the forecasting model is essential in order to keep the prediction accuracy within acceptable ranges. The treatment of bus load characteristics is held with computational intelligence techniques such as clustering and ANN. Neural network based systems are a favorable scheme in recent years in price and load predictions over traditional time series models. ANN can fully adapt expert knowledge and modify their parameters accordingly to simulate the problem‘s attributions through training paradigms. Thus, ANN based systems are an essential choice, justified by the paper‘s findings, for highly volatile time series. This work focuses on the short-term load forecasting (STLF) of a number of buses within the Greek interconnected system. Firstly, a modified version of the ANN already proposed for the aggregated load of the interconnected system is employed. To enhance the forecasting accuracy of the ANN, the load profiling methodology is used resulting to the formulation of two novel hybrid forecasting models. These models refer to the combination of the ANN with a clustering algorithm, resulting to superior performance. Simulation results indicate that the combination captures and successfully treats the special characteristics of the bus load patterns. The scope of the present paper is to develop efficient forecasting systems for short-term bus load predictions. This is a current research challenge due to the high interest for smart grids and demand side management applications by utilities, regulators, retailer and energy service companies. Bus load forecasting appears to be a more difficult engineering problem compared to forecasting of the total load of a country. No hybrid models for bus load predictions have been presented so far in the literature. Two novel clustering based tools are developed and successfully tested in a number of loads covering different types of electricity consumers and demand levels.

1. Introduction

1.1. Motivation

Noticeable research interest by the power systems community emphasizes in smart grids and micro-grids (Fadaeenejad et al., 2014; Luthra, Kumar, Kharb, Ansari, & Shimmi, 2014). These grids include small scale generation and storage technologies and involve small scale loads. In order to fully assess the potential of these technologies, there is need to accurately predict the energy patterns of the small scaled loads such as bus loads. While load forecasting in aggregated system level counts decades of research efforts and applications, the literature focusing in small scale load forecasting is relatively limited. Bus loads are characterized by high degree of complexity due to the presence of more atypical loads and by volatility. Also, due to possible network configuration changes, current low values may present low correlation with historical ones. For these reasons, reported forecasting accuracy is considerably higher regarding to the one of the system load (Amjady, 2007). This means that a robust bus forecasting model should be able to capture the complexity of bus load patterns,

* Tel.: +30 2410 684303; fax: +30 2410 684325.
E-mail address: jpanap@gmail.com

http://dx.doi.org/10.1016/j.eswa.2016.01.034

should be easy to implement and characterized by high level of parameterization. The analysis of the paper is centered towards these goals.

1.2. Solution approach

The aim of this work is to explore the potential of the implementation of the load profiling methodology for the purpose of increasing the prediction accuracy of feed-forward neural network trained by the resilient back-propagation algorithm, leading to the construction of hybrid forecasting models (Riedmiller & Braun, 1992). Load profiling refers to a procedure that groups together load patterns with similar curve shapes. There are numerous studies proposed to classify different data sets of electricity consumers (Panapakidis, Alexiadis, & Papagiannis, 2012). The classification is held with clustering algorithms. This study investigates a load profiling application in bus STLF. The data set under study includes the hourly load values of 10 buses located in the area of Thessaloniki, North Greece. Generally, bus loads can refer to the load supplied by transmission and distribution systems transformers. The available data set covers the period between 01/01/2006 and 31/12/2010. Among them 80% of the data are used for training set and the rest 20% is used as test set. The training set refers to 4 years period and is used to define the optimal ANN configuration. A parametric analysis takes place for the purpose of determining the optimal selection of the type of neurons activation function, number of hidden layers, number of neurons in the hidden layer(s) and maximum number of training epochs. The test set refers to 1 year period and is accounted for the models comparison.

As a benchmark for the hybrid models performance, a slight variation of the ANN proposed for the Greek interconnected system, hereafter referred to as Model A, is utilized (Bakirtzis, Petridis, Kiartzis, Alexiadis, & Maissis, 1996; Kiartzis, Zournas, Theocharis, Bakirtzis, & Petridis, 1997). The combination of the load profiling tool with a conventional forecaster (i.e. Model A) leads to the construction of hybrid models that utilize unsupervised machine learning stage via the load curves clustering and supervised machine learning stage via the training and application of the ANN. More specifically, two hybrid models are regarded, namely Model B and Model C. Model B aims on the management of the holidays load by entering coded information in the ANN. The operation of Model C includes the entrance of normalized load curves that are correlated with the target load.

The selection of feed-forward neural networks (FFNNs) in the present study among other ANNs such as support vector machines, Elman networks or Radial Basis Networks is support by the following factors: FFNNs can be trained by several algorithms (i.e. contrary to SVMs) and the training phase is fast (i.e. contrary to RBFs). Also, to further increase the effectiveness of a FFNN forecasting, the clustering tool is used. While the majority of the literature uses the clustering to create subsets of the training set and therefore requiring a number of forecasters equal to the number of subsets, this study uses the clustering under two novel approaches. Only a single forecaster is used into the two developed hybrid systems leading to lower computational complexity.

As for the limitations of FFNNs, a proper network training demands enough data. FFFNs cannot satisfactory simulate small amounts of data. Another crucial limitation is the careful selection of the number and types of inputs, especially in the cases where the external expert knowledge is absent. Regarding the clustering tool, the sophisticated selection of the algorithm increases the forecasting problem's concerns. Algorithms with many input requirements require special care. The minCEntropy algorithm adopted for the needs of this study belongs to the family of partitional algorithms. These algorithms tend to form equally sized clusters and thus atypical data cannot easily be tracked.

1.3. Literature survey and contributions

In the majority of the researches, load profiling based load curve characterization is held through the utilization of clustering algorithms. The load curves are grouped in the optimal number of clusters. Each cluster is identified by its population and its representative pattern, i.e. the centroid. According to the literature, the load profiling concept is also manifested in load forecasting problems with clustering algorithms. The objective is to increase forecasting robustness exploiting the information extracted by the load profiling procedure. The published works can be categorized into three major classes:

- Approaches that involve the clustering algorithm solely. There is an absence of the supervised learning stage of the conventional forecaster. The clustering is used both for grouping together similar load patterns and performing the forecasting. In López, Valero, Senabre, Aparicio, and Gabaldon (2012) the historical load data are grouped into several clusters through the utilization of a Self-Organizing Map (SOM). After the clustering, the association stage follows. The known load data of the target day are associated with the most similar clustered data. The remaining load data within the 24-h period of the most similar pattern is regarded as the forecasted data of the target day. Martínez-Álvarez, Troncoso, Riquelme, and Aguilar-Ruiz (2011) involve a pattern sequence similarity approach. The K-means algorithm is employed for the purpose of grouping and labeling the samples from the dataset. Next, the pattern sequence prior to the day to be predicted is obtained. This sequence is searched in the historical data and the prediction is calculated by averaging all the samples immediately after the matched sequence. The study of Carpinteiro, Reis, and da Silva (2004) refers to hour ahead forecasting using two SOMs in serial connection. The 1st SOM forecasts the loads of hours 1–6 of the daily load. The remaining hours are treated by the 2nd SOM.
- Approaches in which the clustering algorithm and the conventional forecaster are hybridized to combined models. This category of researches is the most common in the literature. Employing the clustering algorithm, the historical data, which may refer exclusively to load records or load records with weather data and day type identification codes, are grouped together. Next, for each cluster a dedicated forecasting model is trained and applied. This approach corresponds to increased computational complexity; the number of forecasting models equals the number of clusters. Note that since the conventional forecaster is trained with a different sub-set of the historical data, the optimal parameters of each forecaster obtained by the training phase may differ. The conventional forecaster involves either time series model or a computational intelligence related algorithm such as ANN or Support Vector Machine (SVM). Nagi, Yap, Nagi, Tiong, and Ahmed (2011) propose the combined implementation of SOM and SVM for long-term peak load forecasting, while in Mori and Yuihara (2001) the hybrid model includes deterministic annealing clustering together with ANNs. In Sperandio, Bernardon, and Garcia (2013), for each generated cluster, a discrete probability model is employed formulated. The prediction is accomplished through the estimation of the probability of a certain demand level to happen given a climatic condition. The wavelet transform is utilized in Kim, Yu, and Song (2002) in order to decompose the clustered load data in four signal components. The forecasting is made for each component separately and the predicted load is obtained by the reconstruction of the components. Yadav and Srinivasan (2011) involve the SOM with a time series model applied in hour ahead forecasting. Again the SOM is used to categorize the training data in Fan, Mao, and Chena (2006). For each