ARTICLE IN PRESS

J. Parallel Distrib. Comput. ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

J. Parallel Distrib. Comput.

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

An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders

Chao Tong [a], Jun Li [a], Chao Lang [a], Fanxin Kong [b], Jianwei Niu [a], Joel J.P.C. Rodrigues [c,d,e,f,*]

[a] School of Computer Science and Engineering, Beihang University, Beijing, 100191, China
[b] School of Computer Science, McGill University, Montreal, Canada, H3A2T5
[c] National Institute of Telecommunications (Inatel), 37540-000 Santa Rita do Sapucaí - MG, Brazil
[d] Instituto de Telecomunicações, 6201-001 Covilhã, Portugal
[e] University of Fortaleza (UNIFOR), 60811-905 Fortaleza-CE, Brazil
[f] ITMO University, 191002 St. Petersburg, Russia

HIGHLIGHTS

- We propose a deep learning based model for forecasting day-ahead electricity load.
- It uses history load data, weather and season parameters.
- It uses multiple stacked denoising auto-encoders to extract features.
- The refined features and a season parameter are fed into a SVR model for training.

ARTICLE INFO

Article history:
Received 2 March 2017
Received in revised form 5 May 2017
Accepted 9 June 2017
Available online xxxx

Keywords:
Deep learning
Multi-modal
Stacked denoising auto-encoders
Feature extraction
Support vector regression

ABSTRACT

In real word it is quite meaningful to forecast the day-ahead electricity load for an area, which is beneficial to reduction of electricity waste and rational arrangement of electric generator units. The deployment of various sensors strongly pushes this forecasting research into a "big data" era for a huge amount of information has been accumulated. Meanwhile the prosperous development of deep learning (DL) theory provides powerful tools to handle massive data and often outperforms conventional machine learning methods in many traditional fields. Inspired by these, we propose a deep learning based model which firstly refines features by stacked denoising auto-encoders (SDAs) from history electricity load data and related temperature parameters, subsequently trains a support vector regression (SVR) model to forecast the day-ahead total electricity load. The most significant contribution of this heterogeneous deep model is that the abstract features extracted by SADs from original electricity load data are proven to describe and forecast the load tendency more accurately with lower errors. We evaluate this proposed model by comparing with plain SVR and artificial neural networks (ANNs) models, and the experimental results validate its performance improvements.

1. Introduction

1.1. Background and motivation

Smart Grids (SGs) [17,18,29] (as shown in Fig. 1) are new-type electrical grids which provide a promising scheme for more utility electricity delivery, aiming at reaching higher reliability and using electrical resources more economically and rationally, based on computer remote control and automation with advanced techniques of communication and sensing. These systems are made possible by two-way communication technology which has been used for decades in other industries. Recent years, both theory development and practical applications of SGs have gradually made major strides in the real life. Sensor technique [15,23] is one of the applications which make grids "smart". Massive sensors with different functions are deployed and return abundant heterogeneous data describing the actual situation of grids in detail, and then some controls and adjustments can be done to fit the dynamic changes in the real environment. Thus the occurrence of SGs strongly pushes electrical grids research into a "big data" era due to a huge amount of information collected by sensors.

One of the most common tasks in electrical grids is forecasting the day-ahead electricity load for an area, e.g., a city or a state.

* Corresponding author at: National Institute of Telecommunications (Inatel), 37540-000 Santa Rita do Sapucaí - MG, Brazil.
E-mail address: joeljr@ieee.org (J.J.P.C. Rodrigues).

http://dx.doi.org/10.1016/j.jpdc.2017.06.007
0743-7315/© 2017 Elsevier Inc. All rights reserved.

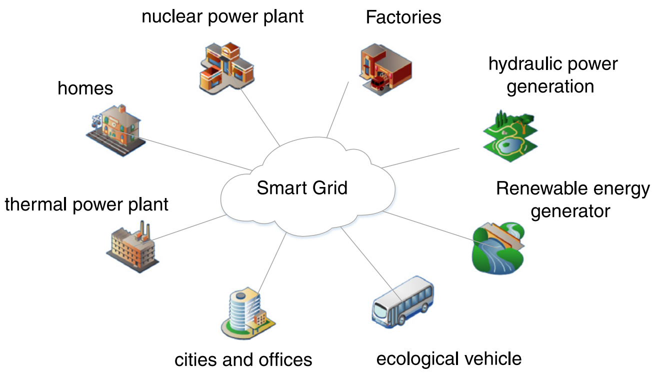

Fig. 1. Illustration of a simple smart grid.

Accurate and proper forecasting helps rational scheduling of electricity generator units, which is beneficial for saving electric power and reducing production cost. Some machine learning-based methods like support vector regression (SVR) [2,21] and artificial neural networks (ANNs) [3] have been proposed for this kind of load forecasting problems, and the accuracy has been improved by utilizing more and different features like holiday information and weather parameters. However, these features are often directly fed into a model together for training with no pre-processing separately. We believe this kind of data processing modes are not elaborate and not effective enough when the data scale increases, and the original data should be further handled to refine the potential information. Thus we expect to build a multi-modal architecture [26,33,34] and separately pre-process the low-level data with several different models. Then we can extract some higher-level features to form more abstract descriptions, and use them to train a composite forecasting model.

The motivation of this paper is to create an efficient model to process massive heterogeneous data (history electricity load distributions, weather parameters and season parameters) based on deep learning networks to reach a better performance and higher forecasting accuracy.

1.2. Literature survey

So far, many efforts have been devoted in electricity load forecasting problems. Hooshmand et al. have proposed a new two-step algorithm to forecast short-term electricity load [24], which firstly uses a wavelet transformation (WT) and an artificial neural network for preliminary forecasting with weather parameters (temperature, humidity, and wind speed) and the previous-day load data, and secondly uses a WT and the adaptive neural fuzzy inference system in order to improve the results. Combined with the knowledge of the computational intelligence, Ghanbari et al. have proposed a hybrid computational intelligence methodology [11], called ACO-GA, by integration of ant colony optimization (ACO), genetic algorithm (GA) and fuzzy logic, claiming a better result than the adaptive neuro-fuzzy inference system. Besides, Paparoditis et al. have introduced a novel functional time-series methodology which forecasts electric energy load by solving the weighted average of some elaborately chosen load segments [22], for increasing the forecasting accuracy, working days, weekends, and special days are also studied separately. Guan et al. have proposed a method of wavelet neural networks with data pre-filtering by spike filtering technique [12]. The filtered loads are decomposed into multiple components at different frequencies by wavelet, and then the features of individual components are processed by separate neural networks. Coelho et al. have created a hybrid evolutionary fuzzy model with parameter optimization [8], which is tackled by a bio-inspired optimizer stemming from a combination between two heuristic approaches, and this model is claimed to perform greatly.

In order to achieve better forecasting effectiveness, more variations are considered. Considering the changes of temperature, Selakov et al. have proposed a practical hybrid model for short term load forecasting based on particle swam optimization (PSO) and support vector machines (SVM) [25], primarily targeting to forecast load when there are significant temperature variations. Specifically, it processes the inputs with PSO and trains the model with SVM. Wu et al. have proposed a hybrid model which combines seasonal exponential adjustment method (SEAM) with the regression methods [31], he believes that the separated forecasting to the seasonal item and the trend item would improve the forecasting accuracy. The Kendall correlation testing method is used to verify the seasonal item in their data series. Wi et al. focus on the holidays' load forecasting [30], he presents a fuzzy polynomial regression method with data selection based on Mahalanobis distance for load forecasting on holidays, which has an extra difficulty due to the lack of load data of holidays, compared to normal weekdays and weekends.

Customary methods based on time series modeling [14] such as autoregressive (AR) and autoregressive moving average (ARMA) have been applied as reputable tools for load forecasting in recent years [16,27]. In [7], Minimax Probability Machine Regression (MPMR) is introduced to forecast chaotic load variations and it is guaranteed that the estimated load is in a domain of ϵ from the real load value. In [20], a forecasting algorithm based on the largest Lyapunov exponent estimation is presented.

In recent years, as the consequence of the predominant capability in defining the nonlinear topography between different nonlinear variables, ANNs have become an approved procedure in load forecasting area [5]. In [28], the relation between the load and weather is studied and an ANN is utilized for this purpose. A resembling ANN-based method is proposed in [6] and wavelet is used to model the complex features of the load variations. Here firstly, load data is separated into its low and high frequency components and then multiple neural networks are trained to foretaste each load frequency component exclusively. In [1], Recurrent Wavelet Network (RWN) has been employed for load forecasting in which Orthogonal Least Square (OLS) is applied to solve the problem of RWN initialization at the same time.

The forecasting precisions of these above-mentioned methods are not satisfying enough in real applications, and some of them are especially complicated for implementation in industry.

Deep learning (DL) has attracted huge attention recent years for its remarkable performance in many conventional fields like image classification [10,32] and speech recognition [4,19]. It is born to processing "big data" and naturally we are inspired to apply some proper DL techniques to process the massive data in SGs for electricity load forecasting.

1.3. Our contribution

Utilizing three kinds of significant data: history electricity load data, the date-corresponding weather parameters and the season parameter. we propose a deep learning-based architecture which consists of two modules: a deep feature extraction module and an electricity load forecasting module. The feature extraction module is made up of three deep sub-networks, and we choose stacked denoising auto-encoders (SDAs), which are variants of auto-encoders, as the basic sub-networks because of their adaptability for numerical inputs and impressive abilities of feature extraction. The first sub-network receives the previous days' electricity load vector containing 24 components denoting the hourly electricity loads. The input of the second deep sub-network is an integrated electricity load vector, which is concatenated by the latest 3 days'