Neural Networks 71 (2015) 204–213

Contents lists available at ScienceDirect

Neural Networks

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

Prediction of telephone calls load using Echo State Network with exogenous variables

Filippo Maria Bianchi [a,*], Simone Scardapane [a], Aurelio Uncini [a], Antonello Rizzi [a], Alireza Sadeghian [b]

[a] Department of Information Engineering, Electronics and Telecommunications (DIET), "Sapienza" University of Rome, Via Eudossiana 18, 00184 Rome, Italy

[b] Department of Computer Science, Ryerson University, 350 Victoria Street, Toronto, ON M5B 2K3, Canada

ARTICLE INFO

Article history:
Received 5 June 2015
Received in revised form 23 July 2015
Accepted 28 August 2015
Available online 7 September 2015

Keywords:
Time-series
Forecasting
Echo State Networks
Exogenous variables
Genetic algorithm
Call data records

ABSTRACT

We approach the problem of forecasting the load of incoming calls in a cell of a mobile network using Echo State Networks. With respect to previous approaches to the problem, we consider the inclusion of additional telephone records regarding the activity registered in the cell as exogenous variables, by investigating their usefulness in the forecasting task. Additionally, we analyze different methodologies for training the readout of the network, including two novel variants, namely ν-SVR and an elastic net penalty. Finally, we employ a genetic algorithm for both the tasks of tuning the parameters of the system and for selecting the optimal subset of most informative additional time-series to be considered as external inputs in the forecasting problem. We compare the performances with standard prediction models and we evaluate the results according to the specific properties of the considered time-series.

1. Introduction

Time-Series Forecasting (TSF) refers to the problem of predicting future values of a time-series (TS), starting from a previously observed history (De Gooijer & Hyndman, 2006). In this paper, we are concerned specifically with the TSF problem of telephone activity loads. This is closely related to the forecasting of workload in call centers (Aksin, Armony, & Mehrotra, 2007) where, usually, only the TS containing the load of incoming calls is taken into account and the other external variables considered for the prediction usually possess a very different nature (e.g. advertisement, catalogs, calendar effects Andrews & Cunningham, 1995; Antipov & Meade, 2002; Soyer & Tarimcilar, 2008). An accurate Short-Term Load Forecast (STLF) method would save operating costs, keep power markets efficient and provide a better understanding of the dynamics of the observed system. On the other hand, a wrong prediction could cause either a load overestimation, which leads to the excess of reserving resources and consequently more costs and contract curtailments for market participants, or a load underestimation resulting in failures in providing enough reserves, thereby more costly ancillary services (Bunn, 2000; Ruiz & Gross, 2008).

Specifically, in this work we treat the problem of STFL relative to the telephonic activities registered on a cell covered by an antenna of a mobile phone network. Relatively to each cell there are different kinds of data that describes the volume and the number of both outgoing and incoming calls, from which we generate different TSs. Our work is focused on forecasting the values of a specific TS using past measurements and leveraging on the information contained in the remaining TSs, considered as exogenous variables which are presented as input to the system along with the TS that must be predicted. In particular, in this work we consider call records collected in the Orange telephone dataset published for the "Data for Development" (D4D) challenge (Blondel et al., 2012). More information on the TSs and how they are generated in a pre-processing phase is provided in Section 3.

As forecast method we use a standard Echo State Network (ESN) (Butcher, Verstraeten, Schrauwen, Day, & Haycock, 2013; Jaeger & Haas, 2004; Lukoševičius & Jaeger, 2009; Verstraeten, Schrauwen, d'Haene, & Stroobandt, 2007), which is a particular class of Recurrent Neural Network (RNN). The main peculiarity of ESNs is that the recurrent part of the network (the *reservoir*) is considered fixed, and only a non-recurrent part (termed *readout*) is

* Corresponding author. Tel.: +39 06 44585495; fax: +39 06 4873300.
E-mail addresses: filippomaria.bianchi@uniroma1.it (F.M. Bianchi), simone.scardapane@uniroma1.it (S. Scardapane), aurelio.uncini@uniroma1.it (A. Uncini), antonello.rizzi@uniroma1.it (A. Rizzi), asadeghi@ryerson.ca (A. Sadeghian).

http://dx.doi.org/10.1016/j.neunet.2015.08.010
0893-6080/

effectively trained. In this way, it is possible to use standard linear regression routines for solving the overall optimization problem, without the need of complex backpropagation of the errors over the dynamic portion of the network (Pearlmutter, 1995).

The ESN has proven to be a very effective methodology for the STLF problem in a wide range of applicative domains, including the forecasting of the water inflow to an hydropower plant reservoir (Sacchi, Ozturk, Principe, Carneiro, & da Silva, 2007), electric load in power systems (Deihimi & Showkati, 2012; Qingsong, Xiangmo, Zuren, Yisheng, & Baohua, 2011; Showkati, Hejazi, & Elyasi, 2010), price prediction in economic markets (Lin, Yang, & Song, 2009) and hourly wind speeds (Ferreira, Ludermir, de Aquino, Lira, & Neto, 2008). On the other hand, because of the property of 'short-term memory', ESN is not suitable for long-term predictions (Peng, Lei, Li, & Peng, 2014); for longer forecasting horizons, very basic averaging models outperforms many more complex alternatives (Taylor, 2008). The traditional methods often encounter difficulty in properly modeling the complex nonlinear relationships between load and various exogenous factors, in particular the stochastic ones that influence the load. Conversely, methods based on an ESN model have the ability to learn or map those nonlinear relationships so that, if they are properly constructed and they suitably leverage on historical data, the accuracy of the predicted results can be very high.

In this work we approach the TSF problem with an ESN, facing a novel application concerning the call loads registered in a cell of a mobile telephone network. We use an automatic procedure for identifying the set of exogenous variables during the optimization phase, used as a support information in the prediction. In fact, unlike external variables commonly used in forecasting problems (such as the ambient temperature in the prediction of the electric load in a distribution network), the correlations between the different TSs are not obvious. For training the ESN we consider several approaches in our application, including least-square regression, an elastic net penalty and two versions of the ν-SVR (Shi & Han, 2007). To the best of our knowledge, the last two methods were never investigated for the task of ESN training. We evaluate their performances in the experimental section, where the prediction accuracies are compared with the performances on the same problem by standard forecasting methods, namely the classic Box–Jenkins Auto-Regressive Integrated Moving Average (ARIMA) (Anderson, 1976), Auto-Regressive Integrated Moving Average with Exogenous inputs (ARIMAX) (Huang, Huang, & Wang, 2005) and Triple Exponential Smoothing (TES) (Kalekar, 2004).
It should be pointed out that the ESN depends on several parameters and a correct tuning is essential in order to obtain good performances from the system. An innovative aspect of this paper is the use of a genetic algorithm for simultaneously identifying the optimal settings of the network, and for selecting the most informative subset of TSs to be considered as exogenous variables in the TSF problem.

The remainder of the paper is organized as follows: in Section 2 we review some related works on the STLF using different methods, among which the ESN. In Section 3 we describe the considered TS, in Section 4 we present the proposed ESN-based forecasting system and in Section 5 we report the results obtained and we compare them with the other systems. Finally, in Section 6 we discuss the conclusions and the future works.

Notation

In this paper, vectors are denoted by boldface lowercase letters, such as $\mathbf{a}$, while matrices are denoted by boldface uppercase letters, such as $\mathbf{A}$. All vectors are assumed column vectors, with $\mathbf{a}^T$ denoting the transpose of $\mathbf{a}$. The ith element of a vector is denoted as a_i, while $\mathbf{a}[n]$ is used for time dependence on the index n. $\|a\|_p$ is used for the L_p-norm of a vector. For $p = 2$ this is the standard Euclidean norm, while for $p = 1$ we have $\|\mathbf{a}\|_1 = \sum_i a_i$. Finally, the spectral radius of a generic matrix $\mathbf{A}$ is $\rho(\mathbf{A}) = \max_i \{|\lambda_i(\mathbf{A})|\}$, where $\lambda_i(\mathbf{A})$ is the ith eigenvector of $\mathbf{A}$.

2. Related works

In this section we report different state-of-the-art works in the literature which are related to the problem and to the methodologies considered in this paper. In particular, we firstly review the works which dealt with the STFL problem considering exogenous variables in Section 2.1. Then, we consider the TSF problem approached with methods based on recurrent neural architectures, with particular emphasis on ESN models, in Section 2.2.

2.1. Load prediction using exogenous variables

Load forecasting is vitally important in the deregulated economy and it has many applications including load switching, contract evaluation, and infrastructure development. A large variety of mathematical methods have been developed for load forecasting and some of them use external variables as a support in the prediction task, for increasing the forecast accuracy. Early forecasting studies relative to STLF of call volumes (Andrews & Cunningham, 1995; Antipov & Meade, 2002) applied Auto-Regressive Moving Average (ARMA) models incorporating exogenous inputs along-side MA and AR variables, using transfer functions to help predict outliers, such as special sales promotion periods, and adding exogenous variables in a multiplicative manner for including advertising response and special calendar effects. The underlying intuition is that the predictable behavior of advertising is embedded in the structure of the applications series, thus it is the unpredictable behavior of advertising that will be responsible for innovations in application. In Ibrahim and L'Ecuyer (2013) the authors propose a stochastic model for predicting the call load, which estimates the load in a given period of the day using the correlations between the load observed in the previous hours, the load observed in the day during the previous weeks and the load observed on a second exogenous TS. They stated that modeling correlation structures in the data is not necessary for long-term forecasts. Thus, it is sufficient for real-life call center managers to base their long-term managerial decisions on historical averages, e.g., fixed-effects models. A similar work on TSF (Aldor-Noiman, Feigin, & Mandelbaum, 2009) consists in using a statistical model for predicting the work load in a call center, which is based on a mixed Poisson process approach that considers the effect of events such as billing on the arrival process. The authors demonstrate how to incorporate them as exogenous variables in the model. Exogenous variables are considered also in Niu, Ji, Xing, and Wang (2012), where the prediction of the daily electricity load is performed including the registered temperature in the model as an external variable. They used a special ESN with a different reservoir for each TS: this allow to better catch the dynamic of each single variable. Authors also propose a method for pruning the output weights which, due to the multiplicity of reservoirs, will be huge and could lead to an over-fitting during the training.

2.2. Forecasting with ESNs

Artificial Neural Networks (ANN) have been intensively applied in time-series analysis for several years and many different methodologies and approaches have been explored in the last decade, relatively to the TSF problem (De Gooijer & Hyndman, 2006; Ghiassi, Saidane, & Zimbra, 2005; Zhang, 2001; Zhang, Patuwo, & Hu, 1998). Recently, the ESN has proven to be one of the most effective typology of ANN to be used in TSF, especially