Transportation Research Part C 43 (2014) 158–173

Contents lists available at ScienceDirect

Transportation Research Part C

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

On-line prediction of border crossing traffic using an enhanced Spinning Network method

Lei Lin [a], Qian Wang [a], Shan Huang [b], Adel W. Sadek [a,*]

[a] *Department of Civil, Structural and Environmental Engineering University at Buffalo, the State University of New York, Buffalo, NY 14260, USA*
[b] *Terra Technology, Inc., 1515 E Woodfield Rd., Ste 104, Schaumburg, IL 60173, USA*

ARTICLE INFO

Article history:
Received 28 February 2013
Received in revised form 13 November 2013
Accepted 13 November 2013

Keywords:
Spinning Network (SPN)
Dynamic Time Warping (DTW)
Seasonal Autoregressive Integrated Moving Average (SARIMA)
Support Vector Regression (SVR)
Short-term traffic volume prediction
Data classification
Border crossing

ABSTRACT

This paper improves on the Spinning Network (SPN) method, a novel forecasting technique, inspired by human memory which was recently developed by Huang and Sadek (2009). The improvement centers on the use of the Dynamic Time Warping (DTW) algorithm to assess the similarity between two given time series, instead of using the Euclidean Distance as was the case with the original SPN. Following this, the enhanced method (i.e., hereafter referred to as the DTW–SPN) is used to predict hourly traffic volumes at the Peace Bridge, an international border crossing connecting Western New York State in the U.S. and Southern Ontario in Canada. The performance of the DTW–SPN is then compared to that of three other forecasting methods, namely: (1) the original SPN (referred to as the Euclidean–SPN); (2) the Seasonal Autoregressive Integrated Moving Average (SARIMA) method; and (3) Support Vector Regression (SVR). Both classified as well as non-classified datasets are utilized, with the classification made on the basis of the type of the day to which the data items belong (i.e. Mondays through Thursdays, Fridays, weekends, holidays, and game days). The results indicate that, in terms of the Mean Absolute Percent Error, the DTW–SPN performed the best for all data groups with the exception of the "game day" group, where SVR performed slightly better. From a computational efficiency standpoint, the SPN-type algorithms require runtime significantly lower than that for either SARIMA or SVR. The performance of the DTW–SPN was also quite acceptable even when the data was not classified, indicating the robustness of the proposed forecasting method in dealing with heterogeneous data.

1. Introduction

With increasing concerns about traffic congestion and interest in active traffic management, as a means to address that problem, approaches to providing accurate on-line and short-term forecasts of traffic volume have recently been attracting the attention of traffic researchers and operators alike (see Van Lint and Van Hinsbergen (2012) for a comprehensive review). This is because on-line short-term prediction, which focuses on road traffic condition changes in the near future (ranging from 5 min to about 1 h), is key to monitoring system performance and optimizing traffic operations decisions. While different classification schemes of traffic volume forecasting methods are possible, in this paper we broadly classify them into the following two groups: (1) statistical or time-series approaches; and (2) Artificial Intelligence (AI) based approaches.

With respect to the first group, the Box and Jenkins techniques (e.g., Autoregressive Integrated Moving Average (ARIMA) models) were firstly applied to the field of traffic forecasting by Ahmed and Cook (1979). Since then more and more

* Corresponding author. Tel.: +1 716 645 4367x2421; fax: +1 716 645 4367.
E-mail address: asadek@buffalo.edu (A.W. Sadek).

0968-090X/$ - see front matter
http://dx.doi.org/10.1016/j.trc.2013.11.018

advanced techniques from that family have been applied to traffic volume prediction, such as the seasonal ARIMA models (SARIMA) (Williams and Hoel, 2003; Smith et al., 2002), the ARIMA models with intervention *x*-variables (ARIMAX) (Williams, 2001; Cools et al., 2009), and the combination of Kohonen self-organizing map with ARIMA models (Van Der Voort et al., 1996). In addition to the Box and Jenkins models, other multivariate time series techniques were exploited to increase prediction accuracy, including the state space model (Stathopoulos and Karlaftis, 2003) and the multivariate structural time series model (MST) (Ghosh et al., 2009). Kalman filtering theory was also utilized for short-term traffic forecasting. Examples include its initial examination by Okutani and Stephanedes (1984), the state space based of Stathopoulos and Karlaftis (2003), and the work by Xie et al. (2007) which used a Kalman filter with discrete wavelet decomposition for short term traffic prediction. Most recently, Min and Wynter (2011) adopted a multivariate spatial–temporal autoregressive model (MSTAR) to predict the network-wide speed and volume in real time.

On the AI side, among the most widely used methods are Neural Networks (NNs). Several NN topologies have been utilized in previous studies including the multilayer perceptron networks (MLP) (Smith and Demetsky, 1994; Chang and Su, 1995; Ledoux, 1997), radial basis function networks (Park et al., 1998), resource allocating networks (Chen and Grant-Muller, 2001), and wavelet networks (Chen et al., 2006; Xie and Zhang, 2006). NNs were also sometimes combined with other methods, such as fuzzy sets and genetic algorithms, to develop hybrid and more powerful predictions methods (e.g., Yin et al., 2002; Vlahogianni et al., 2005, and Wei and Chen, 2012). Besides NNs, other AI methods were recently proposed for short-term traffic prediction. Dimitriou et al. (2008), for example, proposed an adaptive hybrid fuzzy rule-based system approach to predict traffic flow in urban arterial networks. Support Vector Regression (SVR) has also recently been exploited for short-term traffic flow prediction. Specifically, Zhang and Xie (2008) compared a v-support vector machine (v-SVM) model with a NN model and concluded that the former performed better. Other examples of applying SVR to traffic prediction include Castro-Neto et al. (2009) and Hong et al. (2011).

Besides the time series models and AI models surveyed above, Huang and Sadek (2009) recently developed a novel forecasting method that attempts to mimic some of the key features of human memory. The method is called the Spinning Network (SPN) method because, as will be explained later in the paper, the method is based on a set of consecutive rings which hold the data items and continuously spin as they receive new data. Huang and Sadek (2009) also tested the SPN on a 5-min traffic volume dataset from the Hampton Roads area Virginia, and showed that the method yielded superior predictive accuracy in comparison to the back propagation NN and the *k*-nearest neighbor algorithm. More importantly, the SPN only consumed a fraction of the time required by either the nearest neighbor algorithm or the back propagation NN.

In this paper, the focus is on additional testing and improvement of the SPN algorithm on a more challenging short-term traffic forecast dataset than the test dataset utilized in our original paper (Huang and Sadek, 2009). Specifically, the dataset utilized herein is an hourly traffic volume dataset that comes from one of Northern America's busiest border crossings, namely the Peace Bridge, connecting Western New York in the US and Southern Ontario in Canada. Compared to traffic volume datasets from typical locations, such as the Virginia dataset used in the original study (Huang and Sadek, 2009), border crossing traffic has several unique features (e.g., significant differences between weekday and weekend traffic, sensitivity to special events such as sporting events and national and religious holidays, etc.) and is thus much harder to predict. In fact, the highly non-linear trends in the Peace Bridge hourly volume dataset required introducing new features to the original SPN algorithm, as well as considering classifying the dataset into more homogeneous groups, as will be described in more detail later. The performance of the SPN algorithm was also compared to both the SARIMA model (as a representative of the statistical time series forecasting approach) and SVR (as a representative of the AI-based approach).

The remainder of the paper is organized as follows. Section 2 provides background information about: (1) the original SPN method developed by Huang and Sadek (2009); (2) SARIMA and SVR, the two models used in this study as the benchmarks; and (3) a handful of previous studies that focused on predicting border crossing volume and delays. Section 3 describes the modeling dataset and the statistical tests performed to characterize the dataset and to assess the difficulty level of predicting the time series in comparison to the Virginia dataset utilized in our previous study (Huang and Sadek, 2009). Section 3 also discusses how the dataset was classified into groups to facilitate the forecasting process. Section 4 describes the details of the methodology followed in this study, including a discussion of how the original SPN was enhanced to improve its accuracy in predicting the border crossing traffic volumes. The results from comparing the performance of the both the enhanced and the original SPN to the performance of SARIMA and SVR are presented in Section 5. Finally, the main conclusions from the study are summarized in Section 6, along with a few recommendations for future research.

2. Background

2.1. SPN method

The SPN method is a novel forecasting algorithm originally proposed by Huang and Sadek (2009). The method is inspired by the functionality of human memory in sensing, processing, and predicting the states of the surrounding environment, and attempts to mimic some aspects of human memory including: (1) the fuzzy nature of the information retrieved; (2) the instinctive association of ideas; and (3) the fact that the quality of the information retrieved is a function of the time and effort invested. While the method shares some features with the nearest neighbor approach, one of its key advantages is