Transportation Research Part C 43 (2014) 33–49

Contents lists available at ScienceDirect

Transportation Research Part C

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

A space–time diurnal method for short-term freeway travel time prediction

Yajie Zou [a,*], Xinxin Zhu [b,1], Yunlong Zhang [a,2], Xiaosi Zeng [a,3]

[a] Zachry Department of Civil Engineering, Texas A&M University, 3136 TAMU, College Station, TX 77843-3136, United States
[b] Blocker Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143, United States

ARTICLE INFO

Article history:
Received 28 February 2013
Received in revised form 24 October 2013
Accepted 25 October 2013

Keywords:
Spatial and temporal correlation
Space–time
Diurnal
Freeway link
Probabilistic
Travel time prediction

ABSTRACT

A number of short-term travel time prediction approaches have been developed in the past decade. However, few studies take into account spatial and temporal travel time information simultaneously in the prediction approach. In this study, we proposed a space–time diurnal (ST-D) method, which merges the spatial and temporal travel time information to obtain accurate short-term travel time prediction for freeway corridors under different traffic conditions. The proposed approach can take into account important characteristics of travel times: spatial and temporal correlation, diurnal pattern, and the nonnegativity of the travel time. We use two distributions to model the 5-min average travel time: a truncated normal distribution and a lognormal distribution. Contrary to the most existing methods that yield a point prediction of short-term travel time, this probabilistic modeling approach can overcome the drawbacks of the point prediction by fitting a probability distribution to describe the uncertainty of the future travel times, and from which prediction intervals can be calculated. We use minimum continuous ranked probability score (CRPS) estimation to numerically estimate the parameters in the prediction models. The ST-D method is examined using the travel time data collected on a segment along the US-290 in Houston, Texas. The proposed method provides prediction of travel time over 5-min intervals for up to 1 h in advance. It was found that travel time data from neighboring links along the freeway corridors can be efficiently used to obtain reliable short-term prediction of travel time. The study results suggest the ST-D method is more robust than the traditional vector autoregressive models.

1. Introduction

Travel time information is an important measure of roadway traffic conditions. It is necessary to obtain accurate travel time prediction since the travel time information is an essential input to Intelligent Transportation Systems (ITS) applications, particularly to the Advanced Traveler Information Systems (ATIS). Because of the highly dynamic and nonlinear nature of traffic condition over time and space, travel time prediction remains a difficult yet important challenge for transportation engineers (Xia et al., 2011). There is a rich body of literature on the development of short-term travel time prediction approaches. Previously, van Lint et al. (2005) categorized the explored techniques into three major strands: model-based approaches (e.g., DynaMIT (Ben-Akiva et al., 2002), DynaSMART (Hu, 2001)), instantaneous approaches (e.g., Zhang and Rice,

* Corresponding author. Tel.: +1 936/245 5628; fax: +1 979/845 6481.
E-mail addresses: zouyajie@gmail.com (Y. Zou), xzhu@stat.tamu.edu (X. Zhu), yzhang@civil.tamu.edu (Y. Zhang), david_zeng@neo.tamu.edu (X. Zeng).
[1] Tel.: +1 979/845 3143; fax: +1 979/845 3144.
[2] Tel.: +1 979/845 9902; fax: +1 979/845 6481.
[3] Tel.: +1 979 862 3559; fax: +1 979 845 9873.

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.10.007

2003), and data-driven approaches. So far, many techniques and models developed for short-term travel time prediction belong to the category of data-driven approaches. Examples include generalized linear regression (Zhang and Rice, 2003; Sun et al., 2003), nonlinear time series (Ishak and Al-Deek, 2002), Kalman filters (Chien and Kuchipudi, 2003; Nanthawichit et al., 2003; Chu et al., 2005; van Lint, 2008; Xia et al., 2011), support vector regression (Lam and Toan, 2008), and various neural network models (Park et al., 1999; Rilett and Park, 2001; van Lint et al., 2002; van Lint et al., 2005; van Lint, 2006; Wei and Lee, 2007; van Hinsbergen et al., 2009; Zeng and Zhang, in press).

Despite a large number of short-term travel time prediction approaches have been developed in the past decade, a few studies take into account spatial and temporal travel time information simultaneously in the prediction model. Such techniques are particularly useful in predicting the freeway link travel times since the traffic condition on the neighboring links can help identify the traffic condition on the target link. Previously, Park and Rilett (1999) examined the artificial neural networks, which have the ability to include the spatial and temporal travel time information in predicting multiple-periods link travel times. Upstream and downstream traffic data have also been considered by van Lint (2006) as spatial inputs in the neural network modeling framework for route travel time prediction. Up to now, some researchers consider the traffic data from several locations in the traffic flow prediction models. Stathopoulos and Karlaftis (2001) examined temporal and spatial variations of traffic flow in urban areas. Williams (2001) proposed a multivariate ARIMA approach that includes upstream sensor data to model traffic flow. Stathopoulos and Karlaftis (2003) developed a state-space approach that uses upstream detector data to improve the predictions of traffic flow at downstream locations. Kamarianakis and Prastacos (2003, 2005) introduced the space–time autoregressive integrated moving average (STARIMA) models for predicting traffic flow conditions in an urban network. Sun et al. (2006) and Sun and Xu (2011) proposed different Bayesian network approaches for traffic flow forecasting, which includes spatial information from adjacent road links. Vlahogianni et al. (2007) used genetically optimized modular networks to obtain spatio-temporal prediction of urban traffic volume. A vector auto regressive model was proposed by Chandra and Al-Deek (2009) to predict freeway speeds and volumes using the spatial information from neighboring stations. Min and Wynter (2011) developed a spatio-temporal method for real-time speed and volume prediction by including the spatial characteristics of a road network. Most recently, Zeng and Zhang (in press) examined the importance of considering spatial and temporal input interactions on improving prediction accuracies. Li et al. (2013) extended the probabilistic principle component analysis based imputing method to utilize the information of multiple points by considering temporal and spatial dependence. Pan et al. (2013) proposed a stochastic cell transmission model framework to consider the spatial–temporal correlation of traffic flow and to support short-term traffic state prediction. In this study, we introduced a space–time diurnal (ST-D) method, which merges the spatial and temporal travel time information to obtain accurate short-term travel time prediction of freeway corridors under different traffic conditions. The proposed approach can take into account important characteristics of travel times: spatial and temporal correlation, diurnal pattern, and the nonnegativity of the travel time. Unlike the neural networks and fuzzy logic methods (van Lint et al., 2002; Zhang and Ye, 2008) which use a "black box" approach to predict traffic conditions and often lack a good interpretation of the model, our method makes use of geographically dispersed travel time observations as predictors to obtain short-term prediction and can yield theoretically interpretable prediction models. In this study, the ST-D method is examined using the travel time data collected on a segment along the US-290 in the Houston area. In addition, previous literature usually does not predict more than a single time point into the future. However, prediction on multiple time periods into the future allows for a wider range of applications to make use of the predictions (Min and Wynter, 2011). Thus, this study also seeks to predict travel time for up to 1 h ahead.

The remainder of the paper is organized as follows. In Section 2, we introduce the travel time data in the study. We describe the data collection site and analyze the temporal and spatial correlation as well as the diurnal pattern observed in the data. In Section 3, we provide two versions of the ST-D model for up to 1-h ahead probabilistic prediction of travel time. We also discuss the strategies for selecting predictors, estimating prediction models and choosing appropriate training periods. In Section 4, we evaluate the prediction performance of the ST-D method and compare the proposed models with two conventional prediction models. In Section 5, we provide the conclusions and some future works.

2. Travel time data

2.1. Site information and data collection

The travel time data used in this study is collected on a westbound segment of the US-290 freeway stretch between I-610 and FM-1960 in Houston, Texas. This segment has an imbalanced traffic flow pattern, where the majority of the commuting traffic in the afternoon peak goes westbound. The westbound traffic during the morning peak is less heavy compared with the eastbound traffic due to the directional traffic pattern. The length of the freeway stretch of interest is approximately 12 miles and the free-flow travel time along this stretch is about 13 min. There are 6 automatic vehicle identification (AVI) readers along this stretch, dividing the segment into five links. Vehicles with toll tags passed through the AVI readers will be recorded with their ID and timestamps. By matching their IDs from a pair of AVI readers, the travel time of the link enclosed by this pair of AVI readers is the difference in the timestamps (Songchitruksa et al., 2009). Travel time data were collected 24 h a day for 8 months from January 2008 to August 2008. Individual travel times were aggregated into 5-min intervals for each link. Although the stretch includes two different lane types: one high occupancy vehicle (HOV) lane