Transportation Research Part C 85 (2017) 292–311

Contents lists available at ScienceDirect

Transportation Research Part C

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

Forecasting journey time distribution with consideration to abnormal traffic conditions

R.X. Zhong [a], J.C. Luo [a], H.X. Cai [a], A. Sumalee [b,c], F.F. Yuan [e], Andy H.F. Chow [d,*]

[a] Guangdong Key Laboratory of Intelligent Transportation Systems, School of Engineering, Sun Yat-Sen University, Guangzhou, China
[b] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region
[c] Department of Civil Engineering, King Mongkuts Institute of Technology, Ladkrabang, Bangkok, Thailand
[d] Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong
[e] China Mobile Limited Group Guangdong, Guangzhou, China

ARTICLE INFO

Article history:
Received 4 November 2016
Received in revised form 22 June 2017
Accepted 25 August 2017

Keywords:
Travel time prediction
Functional principal component analysis
Probabilistic nested delay operator
Parallel computing
Adaptiveness to traffic incident

ABSTRACT

Travel time is an important index for managers to evaluate the performance of transportation systems and an intuitive measure for travelers to choose routes and departure times. An important part of the literature focuses on predicting instantaneous travel time under recurrent traffic conditions to disseminate traffic information. However, accurate travel time prediction is important for assessing the effects of abnormal traffic conditions and helping travelers make reliable travel decisions under such conditions. This study proposes an online travel time prediction model with emphasis on capturing the effects of anomalies. The model divides a path into short links. A Functional Principal Component Analysis (FPCA) framework is adopted to forecast link travel times based on historical data and real-time measurements. Furthermore, a probabilistic nested delay operator is used to calculate path travel time distributions. To ensure that the algorithm is fast enough for online applications, parallel computation architecture is introduced to overcome the computational burden of the FPCA. Finally, a rolling horizon structure is applied to online travel time prediction. Empirical results for Guangzhou Airport Expressway indicate that the proposed method can capture an abrupt change in traffic state and provide a promising and reliable travel time prediction at both the link and path levels. In the case where the original FPCA is modified for parallelization, accuracy and computational effort are evaluated and compared with those of the sequential algorithm. The proposed algorithm is found to require only a piece rather than a large set of traffic incident records.

1. Introduction

Travel time is a key performance index of transport systems and an intuitive measure that can be perceived by road users. Travel time can be used as an essential evaluation criterion for traffic managers to devise efficient traffic control schemes to alleviate traffic congestion. However, the provision of travel time information through advanced traveler information systems (ATISs) may have a great impact on drivers decisions, such as route choices and departure times, and thus improve network performance. As traffic networks are subject to both demand and supply uncertainties, especially incident scenarios and adverse weather conditions, travel time indicates stochastic fluctuations (Lam et al., 2008). As unexpected traffic inci-

* Corresponding author.
E-mail addresses: zhrenxin@mail.sysu.edu.cn (R.X. Zhong), agachai.sumalee@polyu.edu.hk (A. Sumalee), andy.chow@ucl.ac.uk (A.H.F. Chow).

https://doi.org/10.1016/j.trc.2017.08.021
0968-090X/

dents account for major delays on freeway systems (Ozbay and Kachroo, 1999; Stathopoulos and Karlaftis, 2002; Du et al., 2012), the variability of travel time is a growing concern for individual travelers and other participants, especially those with tight time constraints. Under such circumstances, estimating or predicting the average travel time may not be sufficient. For these reasons, travel time reliability (TTR) is recognized as a critical factor contributing to the efficiency and service quality of a transportation system. To evaluate TTR, it is necessary to obtain travel time distributions that reflect real-world variability and uncertainty. However, stochastic traffic phenomena render the estimation or prediction of travel time distributions a challenging task that requires a combination of correct physics and strong statistical tools (Yildirimoglu and Geroliminis, 2013). In addition, the dissemination of travel time information requires on-line deployment. Thus, there is a strong need to devise an online prediction methodology that allows travel time distributions that can adapt to anomalies such as traffic accidents and adverse weather conditions.

Major travel time estimation methods can be categorized into two groups: data-driven methods and hybrid data-driven/model-driven methods. Data-driven methods are the most commonly used approaches for travel time estimation or prediction using various traffic data, e.g. loop detector data (Kwon et al., 2001; Coifman, 2002; Coifman and Krishnamurthy, 2007), Bluetooth (Nantes et al., 2015), probe vehicles (Herrera and Bayen, 2010; Jenelius and Koutsopoulos, 2013, 2015), cellular phones (Rose, 2006), vehicle re-identification (VRI) data (Wang et al., 2014) and multiple data sources fusion (Patire et al., 2015). There are two major categories of methods applied to data-driven travel time estimation in the transportation literature: parametric methods (e.g., linear regression, time series models) and non-parametric methods (e.g., neural network models). Interested readers are referred to several state-of-the-art review papers in this stream of the literature, i.e. Karlaftis and Vlahogianni (2011), Vlahogianni et al. (2014), Oh et al. (2015), Mori et al. (2015). Roughly speaking, parametric methods have more solid and widely accepted mathematical foundations than non-parametric methods. However, parametric statistical methods are more computationally extensive, suffer from the curse of dimensionality, and frequently fail when dealing with complex and highly nonlinear data. Non-parametric methods may outperform parametric methods in these respects, yet suffer from complex training with site-specific limitations and black-box procedures. Non-parametric methods are not explanatory for field applications in spite of their prediction performance.

In particular, few studies have tackled the problem of forecasting travel time under anomalies. A pattern recognition method (Bajwa et al., 2005), an online support vector machine for regression (Castro-Neto et al., 2009) and an adaptive dynamic linear model (Fei et al., 2011) were developed for travel time prediction under both recurrent and non-recurrent traffic conditions. Cheng et al. (2014), Fusco et al. (2016) assessed the performance of several state-of-art methods for short-term prediction such as the autoregressive integrated moving average (ARIMA) and neural network (NN). Nevertheless, these studies drew the same conclusion: the performance of those state-of-the-art methods deteriorated tremendously under abnormal traffic patterns caused by demand and supply uncertainties including traffic incidents. In contrast, Wu et al. (2012) proposed a gradient boosting approach to predict traffic conditions under anomalous conditions while disabling the approach when the traffic state returned to normal. Although data-driven methods show good adaptiveness to measurements, they are fragile under demand and supply uncertainties (or lack of robustness) and unable to describe the traffic dynamics, e.g., congestion onset and dissolution (Sumalee et al., 2013; Pan et al., 2013). Moreover, the essential problem facing by the dissemination of travel time information is the prediction of "short-term" future traffic conditions along the route. An appropriate traffic model would be helpful in this area, and even better if the underlying model is robust to demand and supply uncertainties (Zhong et al., 2014).

Combining the data-driven approach with a certain traffic model would achieve both adaptiveness to the actual measurements and robustness to uncertainties. Xia et al. (2010) used physical queue length under incident conditions to enhance the prediction accuracy of travel time. Domenichini et al. (2012) developed a travel time prediction model for both normal and accident conditions requiring the knowledge of accident characteristics, traffic flow during the accident and the capacity drop. Yildirimoglu and Geroliminis (2013) proposed an automatic bottleneck identification algorithm and congestion search algorithm to predict travel times along with the traffic flow fundamentals. Nantes et al. (2016) developed a robust estimation of traffic state using a certain fundamental diagram and fusing heterogeneous sources of synthetic data. In addition, some studies were proposed to evaluate link travel time distributions. Kharoufeh and Gautam (2004) captured stochastic link travel time distributions analytically, which implicitly described the time dependence of vehicle speed in terms of partial differential equations. Kachroo and Sastry (2016) analyzed the limitation of the flow-based travel time functions and then proposed a density-based (link) travel time function and further developed its dynamics in terms of hyperbolic partial differential equations from a given fundamental traffic relationship and vehicle characteristics. These analytical models cannot deduce general explicit expressions for high-order moments of travel time while it is computationally demanding and not ready to be extended to capture route travel time distributions.

There are two ways for evaluating journey time: instantaneous and experienced (also known as the nested delay operator in the dynamic traffic assignment literature). Instantaneous journey time is calculated by summing the link travel times along the path at the departure time of the trip. An important part of the literature provides instantaneous journey time estimation/prediction (Yildirimoglu and Geroliminis, 2013). Experienced journey time is evaluated by tracing a trajectory through the velocity field (Yildirimoglu and Geroliminis, 2013; Chow et al., 2015, 2017). Sumalee et al. (2013), Yildirimoglu and Geroliminis (2013) analyzed the superiority of experienced journey time over instantaneous journey time and proposed a methodological framework that considered the traffic flow essentials (e.g., shockwaves and bottlenecks) for experienced journey time prediction using historical and real-time traffic data. They found that the accuracy of such an approach depended on the way the data were clustered to generalize the stochastic congestion map and the definition of