Transportation Research Part C 74 (2017) 366–382

Contents lists available at ScienceDirect

Transportation Research Part C

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

Using survival models to estimate bus travel times and associated uncertainties

Zhengyao Yu [a], Jonathan S. Wood [b], Vikash V. Gayah [a,*]

[a] Department of Civil and Environmental Engineering, The Pennsylvania State University, University Park, PA 16802, United States
[b] Department of Civil and Environmental Engineering, South Dakota State University, Brookings, SD 57007, United States

ARTICLE INFO

Article history:
Received 28 February 2016
Received in revised form 19 October 2016
Accepted 12 November 2016
Available online 7 December 2016

Keywords:
Bus travel time prediction
Accelerated failure time survival models
Bus travel time reliability
Real-time transit information systems

ABSTRACT

Transit agencies often provide travelers with point estimates of bus travel times to downstream stops to improve the perceived reliability of bus transit systems. Prediction models that can estimate both point estimates and the level of uncertainty associated with these estimates (e.g., travel time variance) might help to further improve reliability by tempering user expectations. In this paper, accelerated failure time survival models are proposed to provide such simultaneous predictions. Data from a headway-based bus route serving the Pennsylvania State University-University Park campus were used to estimate bus travel times using the proposed survival model and traditional linear regression frameworks for comparison. Overall, the accuracy of point estimates from the two approaches, measured using the root-mean-squared errors (RMSEs) and mean absolute errors (MAEs), was similar. This suggests that both methods predict travel times equally well. However, the survival models were found to more accurately describe the uncertainty associated with the predictions. Furthermore, survival model estimates were found to have smaller uncertainties on average, especially when predicted travel times were small. Tests for transferability over time suggested that the models did not over-fit the dataset and validated the predictive ability of models established with historical data. Overall, the survival model approach appears to be a promising method to predict both expected bus travel times and the uncertainty associated with these travel times.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction and background

Travel time reliability is a key indicator of transit service quality and has a strong impact on transit ridership (Paine et al., 1967; Golob et al., 1972; Prashker, 1979). Previous research suggests that travel time reliability may be valued more highly than travel time itself among transit users (Bates et al., 2001; Brownstone and Small, 2005). Furthermore, negative experiences related to unreliable transit service discourage users from continuing to use public transportation (Carrel et al., 2013). Thus, maintaining travel time reliability is extremely important for improving transit competitiveness.

Unfortunately, transit agencies have a difficult time maintaining reliable travel times as bus transit systems are inherently unstable (Newell and Potts, 1964; Newell, 1974). The mechanism that causes this instability is the passenger arrival and service process: the time that a bus spends serving passengers at a stop generally increases with the time between the current and preceding bus arrivals to that stop. For this reason, a bus arriving late to a stop spends more time serving

* Corresponding author.
E-mail addresses: zuy107@psu.edu (Z. Yu), jonathan.wood@sdstate.edu (J.S. Wood), gayah@engr.psu.edu (V.V. Gayah).

http://dx.doi.org/10.1016/j.trc.2016.11.013
0968-090X/© 2016 Elsevier Ltd. All rights reserved.

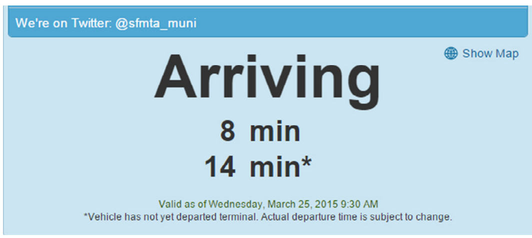
passengers, causing it to arrive even later to the next stop. The reverse is true for a bus arriving early to a stop. This positive feedback loop causes buses to eventually become paired together. Various control strategies have been proposed to combat this instability. These include operational strategies designed to directly prevent bunching (Daganzo, 2009; Delgado et al., 2009, 2012; Xuan et al., 2011; Bartholdi and Eisenstein, 2012) and transit preferential treatment strategies at individual intersections (Stevanovic et al., 2008; Xuan et al., 2009, 2012; Christofa and Skabardonis, 2011; Guler and Menendez, 2014; He et al., 2014; Ma et al., 2014; Ahmed and Hawas, 2015) or along corridors (Viegas and Lu, 2001, 2004; Eichler and Daganzo, 2006; Viegas et al., 2007; Guler and Cassidy, 2012; Guler et al., 2016).

Transit agencies also address bus transit reliability through the provision of high-quality information on real-time bus operations. This helps to improve the *perceived reliability* of the bus system from the user perspective. This information, often obtained using Automatic Vehicle Location (AVL) and Automated Passenger Counter (APC) systems, provides transit users with a snapshot of the most up-to-date operating conditions. Some services also provide real-time bus arrival time predictions based on AVL and APC data; see Fig. 1 for pertinent examples. This type of real-time transit information has been shown to improve perceived reliability and transit ridership (Ettema and Timmermans, 2006; Watkins et al., 2011; Tang and Thakuriah, 2012).

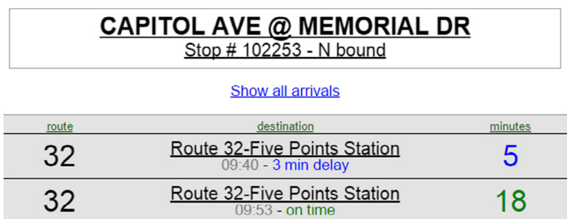
Various methods have been used to develop bus travel time estimates. Theoretical queuing models have been used to describe bus travel through consecutive stops (Vuchic, 2007). This approach has been used to study bus transit system reliability (Islam et al., 2015) and automobile vehicle travel times (Helbing, 2003). Insightful as they are, these queuing models are established upon strong theoretical assumptions and usually require information not available in real time, which makes real-world application and prediction difficult. Instead, this paper focuses on statistical methods applied to empirical data that are easy to reproduce and require minimal modeling assumptions.

In general, empirical efforts to model bus travel times can be categorized into three groups: speed-based models (Sun et al., 2007; Chen et al., 2011), classic regression models (Frechette and Khan, 1998; Patnaik et al., 2004), and artificial neural network (ANN) models (Chien et al., 2002; Chen et al., 2007). *Speed-based models* split the bus route into segments and estimate the average speed on each segment separately. Historical and real-time speed data are weighted according to the vehicle's location in the segment: when the vehicle is close to the beginning of a segment, the historical speed data receive a larger weight and vice versa. Kalman Filters are typically used to provide the appropriate weights between real-time and historical information. *Classic regression models* build clear relationships between travel times and related factors. However, selecting independent variables and finding suitable transformations for the variables in the models is often difficult. Previous studies created regression models using average flow, average transit bus flow, heavy vehicles proportion, travel distance, average dwell time, number of stops, and time period of the day as independent variables, resulting in promising predictions. Linear regression models appear to be the most common method to estimate bus travel times in the literature. Finally, *Artificial Neural Networks* are a powerful tool for modeling complicated problems where the relationship between input and output is not clear (i.e., there is not a clear theory specifying how the relationships between the independent and dependent variables should be related). Some ANN models with dynamic algorithms were developed and tested for bus travel time predictions. Although ANN models performed better than linear regression models in terms of prediction power in the literature (Jeong and Rilett, 2004; Yu et al., 2011), the model form of ANN models makes it difficult to estimate and interpret.

Unfortunately, previous empirical studies generally used very small samples (a month in Chen et al., 2011; 45 h in Frechette and Khan, 1998; two weeks in Sun et al., 2007) or simulated data (Chien et al., 2002), which limits the applicability of the models for predictive purposes. Furthermore, these existing empirical methods generally only provide point estimates for expected bus travel times (as shown in Fig. 1) without any indication of the variability associated with the predictions. However, bus travel times along a link might vary significantly due to various factors, such as traffic control at intersections, passenger demands at intermediate stops, and interruptions from other modes (Mazloumi et al., 2009). Users relying on point estimates of bus travel time are liable to experience long waits (for buses that are late relative to the predictions) or perhaps miss the bus altogether (for buses that arrive early relative to the predictions), both of which could generate negative feelings towards bus transit systems. Providing an indication of the uncertainty associated with predictions could help mitigate these negative outcomes and improve perceived reliability of these estimates. This could be as simple as an



(a) NextBus (NextBus, 2015)



(b) OneBusAway (OneBusAway, 2015)

Fig. 1. Examples of travel time estimates available in real-time transit information systems NextBus (2015) and OneBusAway (2015).

Download English Version:

https://daneshyari.com/en/article/4968596

Download Persian Version:

https://daneshyari.com/article/4968596

Daneshyari.com