Transportation Research Part C 86 (2018) 453–466

Contents lists available at ScienceDirect

Transportation Research Part C

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

Analysis of bus travel time distributions for varying horizons and real-time applications

Md. Matiur Rahman, S.C. Wirasinghe*, Lina Kattan

Department of Civil Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, T2N1N4 Calgary, Canada

ARTICLE INFO

Keywords:
Distribution
Interval
Horizon
Arrival
Travel time
Real-time

ABSTRACT

Given the increasing interest in real-time bus arrival information, producing reliable estimation is essential to maximize the benefits of real-time systems. The primary objectives of this paper are to analyze the changes of bus travel time characteristics as pseudo horizon varies and how such characteristics can be applied to real-time bus arrival estimation. In this study, "horizon" refers to the distance between a real-time bus location and a bus stop, whereas "pseudo horizon" refers to the distance from a GPS point to an upstream GPS point. In contrast to existing methods that provide point estimates of bus arrival times, this study provides interval estimates that take into account the uncertainty of future bus arrival times given that early and late buses have their own respective ramifications. A methodology is developed to analyze the bus travel time distribution systematically based on different pseudo horizons since such distributions are critical to producing reliable bus arrival information. The analysis of real transit GPS data shows a significant change in bus travel time characteristics around a pseudo horizon range of 8 km. The analysis of changes in probability densities with pseudo horizons shows that bus travel time distribution converges from a rightly skewed distribution to a more symmetrical distribution from a shorter to a longer pseudo horizon. Lognormal and normal distributions are found to be the best models for before and after a cut-off horizon of 7–8 km, respectively. Instead of a single distribution, the outcomes of this study suggest a combination of probability distributions based on the estimation horizon to be used to provide better bus arrival time estimations.

1. Introduction

Transit bus operations, in particular travel times, are affected by factors such as traffic flow, passenger boarding and alighting, traffic signals, weather, pedestrians, road maintenance, incidents and bus driver behavior (Abdelfattah and Khan, 1998; Shalaby and Farhan, 2004; Tao et al., 2018; Rahman et al., 2011). The resulting uncertainty in bus arrival times at stops creates a significant disutility for travelers. A reduction in travel time variation is valued by transit users (Bates et al., 2001; Lam and Small, 2001; Sun et al., 2003) who consider it more valuable than a decrease in average travel time (Bates et al., 2001). The focus of this study is on bus travel time variation and estimation for a regular route where buses run to a fixed schedule.

The provision of real-time information is a cost-effective way to reduce the disutility associated with uncertainty in bus arrival times (Dziekan and Vermeulen, 2006; Litman, 2008; Rahman et al., 2013; Zhou et al., 2017a, 2017b). Transit agencies and the relevant literature consider a bus information system as "real-time" when the arrival information is dynamically updated every few minutes based on tracking bus locations (Rahman et al., 2016; Sun et al., 2007; Park et al., 2007). The widespread adoption of

* Corresponding author.
E-mail addresses: matiurbuet02@gmail.com (Md. M. Rahman), wirasing@ucalgary.ca (S.C. Wirasinghe), lkattan@ucalgary.ca (L. Kattan).

https://doi.org/10.1016/j.trc.2017.11.023
Received 21 July 2016; Received in revised form 27 November 2017; Accepted 28 November 2017

smartphones by transit passengers (particularly with mobile apps on transit arrivals) has increased expectations for more reliable, real-time bus arrival information (Brakewood et al., 2015). Consequently, transit agencies are rapidly developing and implementing such information systems (Schweiger, 2011). We mainly find two types of bus arrival information systems for a scheduled bus system: (i) a published timetable, which does not change for a given season of the year, and (ii) a published timetable that is updated based on real-time bus arrival time information. Bus arrival time can be updated based on a number of inputs, and the minimal input information for a real-time bus arrival system is the GPS bus locations.

Producing reliable real-time information is essential to maximize the benefits of such systems. Estimated real-time bus arrival information is usually provided as point estimates; for example, "the bus will arrive in 5 min". However, since real-time estimations are subject to error, these point estimates do not reflect the associated uncertainty. Errors in the real-time estimation process can be more frustrating to passengers than deviations from a schedule with no real-time information (Rahman et al., 2013; Cats and Loutos, 2016). One solution to address this issue is to provide interval estimates; for example, "the bus will arrive in 4–6 min". Based on a recent transit user survey, Rahman et al. (2017) found that compared to point estimates, a majority of transit users preferred interval estimates. One of the main reasons for choosing interval estimates is that they reflect the inevitable variability of bus arrival times. This study attempts to contribute on two primary issues: (i) examining changes in bus travel time characteristics as the distance from a GPS point to an upstream GPS point varies; (ii) applying such travel time characteristics to produce interval estimates in real-time bus arrival estimation. Despite the importance, both issues are widely ignored in operational applications of bus arrival time estimations and are rarely mentioned in the literature.

The rest of the paper is organized as follows: Section 2 provides a review of the prior research. Section 3 presents the research methodology and describes the data. Section 4 discusses the results, and Section 5 presents the conclusions, policy implications, and opportunities for future studies.

2. Literature review

The factors contributing to bus travel time variation are typically grouped into either demand or capacity related factors (Bates et al., 2001; Mazloumi et al., 2009; Du et al., 2017). For instance, two main demand related factors are traffic flow and the number of passengers using a particular bus route (Li, 2002; Chien et al., 2002; Shalaby and Farhan, 2004). Weather conditions (Hofmann and O'Mahony, 2005; Zhou et al., 2017a, 2017b) and incidents, such as traffic collisions (Abdelfattah and Khan, 1998), are the usual capacity related factors. Human factors, such as driver characteristics (Strathman and Hopper, 1993), also contribute to travel time variation. Travel time variation for a given bus route can be analyzed from different viewpoints such as day-to-day variation within the same period or variations between different periods of a typical day (Mazloumi et al., 2009; Ma et al., 2016). The focus of this study is on variations within the same period for a regular bus route.

Several approaches, such as the use of only standard deviation, coefficient of variation, or skewness, and in some cases, the probability distribution, have been used to quantify bus travel time variability (Mazloumi et al., 2009). Several studies on road traffic in urban areas showed that the standard deviation of travel time increased linearly with the average travel time (Herman and Lam, 1974; Polus, 1979; May et al., 1989). A similar relationship was reported for bus travel times (Sterman and Schofer, 1976; Ng and Brah, 1998; Mazloumi et al., 2009). However, a travel time distribution illustrates the nature of the variability more comprehensively. Several studies have been conducted on fitting continuous distributions to travel time traffic data. A normal distribution was proposed to characterize vehicle travel time (Ma et al., 2016). However, several other studies identified distributions to be positively skewed (e.g. Richardson and Taylor, 1978). The lognormal model was the most commonly used distribution in traffic studies (Clark and Watling, 2005; Sumalee et al., 2006; El Faouzi and Maurin, 2007; Hollander and Liu, 2008). Other models included log-logistic (Chu, 2010), Weibull (Al-Deek and Emam, 2006), Gamma (Polus, 1979), and Burr (Taylor and Susilawati, 2012). However, compared to studies on the travel time distribution of traffic, there are a limited number of studies on bus travel time distributions mainly due to the unavailability of data (Ma et al., 2016).

Jordan and Turnquist (1979) reported that the Gamma distribution was a good fit for bus running times during the morning peak period. Taylor (1982) showed that a normal distribution was a good fit for the travel times of buses that started at a given time every day. Another study revealed that bus running times on arterial roadways were positively skewed and generally followed a lognormal distribution (Uno et al., 2009). Xue et al. (2011) showed that bus travel times during peak hours followed a log-logistic distribution. Kieu et al. (2014) suggested a lognormal distribution as the best fit of bus travel time on urban roads. Further, a bus travel time distribution tended towards a normal distribution for a short departure time window (Mazloumi et al., 2009). They also reported that while a normal distribution is a better fit for the peak period, a lognormal distribution is more suitable for the off-peak period for a longer starting time window.

Since bus travel times inevitably vary, the provision of real-time bus arrival information is one cost-effective way to reduce the disutility associated with the related uncertainty (Litman, 2008; Rahman et al., 2013). A number of studies examined the impact of real-time information in terms of improved perception regarding wait times and increased ridership (Ferris et al., 2010; Tang and Thakuriah, 2012; Gooze et al., 2013;Brakewood et al., 2014). For example, it was reported that perceived wait time at bus stops decreased by 20% due to the provision of real-time information (Dziekan and Vermeulen, 2006). Watkins et al. (2011) showed that riders who used real-time information waited at bus stops for about 2 min less than those of non-users. It was reported that perceived wait times of real-time information users were about 30% less than that of non-real-time information users. Tang and Thakuriah (2012) observed an increase of 1.8–2.2% in average weekday trips due to availability of real-time information. Brakewood et al. (2014) found a median increase of 1.7% for weekday route-level ridership. The impact of real-time information on ridership was summarized in detail by Brakewood et al. (2015).