Transportation Research Part C 90 (2018) 292–306

Contents lists available at ScienceDirect

Transportation Research Part C

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

Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: A smart card data analysis

Jesper Bláfoss Ingvardson[a,*], Otto Anker Nielsen[a], Sebastián Raveau[b], Bo Friis Nielsen[c]

[a] Department of Management Engineering, Technical University of Denmark, Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark
[b] Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile
[c] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Richard Petersens Plads, 2800 Kgs. Lyngby, Denmark

A R T I C L E I N F O

Keywords:
Automated fare collection data
Smart card
Waiting time
Public transport
Frequency-based timetables
Mixture distributions

A B S T R A C T

Waiting time at public transport stops is perceived by passengers to be more onerous than in-vehicle time, hence it strongly influences the attractiveness and use of public transport. Transport models traditionally assume that average waiting times are half the service headway by assuming random passenger arrivals. However, research agree that two distinct passenger behaviour types exist: one group arrives randomly, whereas another group actively tries to minimise their waiting time by arriving in a timely manner at the scheduled departure time. This study proposes a general framework for estimating passenger waiting times which incorporates the arrival patterns of these two groups explicitly, namely by using a mixture distribution consisting of a uniform and a beta distribution. The framework is empirically validated using a large-scale automatic fare collection system from the Greater Copenhagen Area covering metro, suburban, and regional rail stations thereby giving a range of service headways from 2 to 60 min. It was shown that the proposed mixture distribution is superior to other distributions proposed in the literature. This can improve waiting time estimations in public transport models. The results show that even at 5-min headways 43% of passengers arrive in a timely manner to stations when timetables are available. The results bear important policy implications in terms of providing actual timetables, even at high service frequencies, in order for passengers to be able to minimise their waiting times.

1. Introduction

Waiting time at public transport stops is perceived by passengers to be more onerous than in-vehicle time (Nielsen, 2000; Fan et al., 2016; Fosgerau et al., 2007). Reducing waiting time is therefore of great importance when designing public transport systems. For en-route transfers this can be achieved by optimising public transport timetables in order to ensure short transfer times (Parbo et al., 2014). However, this only affects transferring passengers which in the Greater Copenhagen Area are 45% of all passengers (Christiansen, 2015). For all passengers it is also important to consider the waiting time experienced at the origin stop so that passengers can actively reduce their waiting time.

In general, when timetables are available to passengers, two distinct types of travel behaviour are observed when arriving at a

* Corresponding author.
E-mail addresses: jbin@dtu.dk (J.B. Ingvardson), oani@dtu.dk (O.A. Nielsen), sraveau@ing.puc.cl (S. Raveau), bfni@dtu.dk (B.F. Nielsen).

https://doi.org/10.1016/j.trc.2018.03.006
Received 19 September 2017; Received in revised form 9 March 2018; Accepted 10 March 2018
0968-090X/

departure stop: (i) one group arrives randomly, and (ii) a second group will try to minimise the waiting time by arriving in a timely manner at the scheduled departure time (Csikos and Currie, 2008; Frumin and Zhao, 2012; Jolliffe and Hutchinson, 1975; Luethi et al., 2007). The shares of the two groups are influenced by service characteristics such as headway and reliability as well as other factors, such as time of day (Csikos and Currie, 2008; Frumin and Zhao, 2012; Luethi et al., 2007; Nygaard and Tørset, 2016). As headway or reliability decreases, the share of passengers arriving randomly increases as the potential benefit of reduced waiting time declines (Bowman and Turnquist, 1981). Capturing such behaviour accurately in transport models is important for estimating impacts of public transport investments. However, most traditional public transport assignment models assume all passengers to arrive randomly to the stop, hence assuming the average waiting time to be half the headway (Fu et al., 2012; Nielsen, 2000; Nökel and Wekeck, 2009; Schmöcker et al., 2011; Szeto et al., 2013, 2011). Therefore such models might overestimate waiting times as more passengers will time their arrival at the station leading to lower actual waiting times (Csikos and Currie, 2008).

This paper contributes to the existing literature by analysing a large-scale empirical data set of passenger arrivals and waiting times at rail stations in the Greater Copenhagen Area. The contribution is three-fold.

Firstly, this study proposes a general methodology to model passenger waiting times at public transport stations by explicitly taking into account passengers arriving randomly and non-randomly. The methodology is an extension of the approach proposed in Luethi et al. (2007) where the arrival patterns of passengers were modelled as a mixture of a uniform and a Johnson SB distribution, thus taking into account random and non-random passenger arrivals, respectively. In the present study the method is further developed by proposing a general mixture of a uniform and a beta distribution which can be fitted to specific service frequencies by adjusting the share of uniform passenger arrivals as well as the parameters for the beta distribution. The simple and general formulation makes it easily adoptable in public assignment models.

Secondly, data used for this study is based on a large-scale automated fare collection (AFC) system containing more than 1.5 million trips covering all modes of public transport in the Greater Copenhagen Area during September and October 2014. This allows for extending the work of Frumin and Zhao (2012), which deployed a large-scale AFC dataset from the London Overground to analyse passenger arrivals at boarding stations at headways in the range 7–30 min. This present study covers stations with headways ranging from 2 min (on the metro) to 60 min (on regional train lines). In addition, it includes modes with traditional published timetables (suburban and regional trains) and frequency-based timetables where actual departure times at stations are not published (metro). This makes it possible to compare results across a wide range of service frequencies and timetable types.

Thirdly, this study takes into account the effects of multiple station characteristics on the arrival patterns and waiting times of passengers. This data is joined onto databases including information on station layouts, station amenities and land use types surrounding the stations. This makes it possible to estimate the importance of such characteristics on passenger waiting times.

The paper is organised with a review of the existing literature on analysing passenger waiting times in public transport in Section 2. The methodology and data used in this study is described in Section 3 while the results are presented and discussed in Section 4. In Section 5 conclusions are drawn while policy implications are highlighted in Section 6.

2. Literature review

2.1. Estimation of waiting times

Some of the earliest studies focusing on passenger arrival patterns analysed the relationship between waiting time and headway by a simple linear relationship. O'Flaherty and Mangan (1970) found that average passenger waiting time, W, could be related to average bus headway, h (measured in minutes) by the simple linear relationship $W = 1.79 + 0.14h$ during an evening peak period in Leeds, UK. Seddon and Day (1974) improved the simple model by adding the influence of random bus arrivals. In their study the relationship was found to be $W = 2.34 + 0.26h$ (measured in minutes) for stops in Manchester during both peak and off-peak hours. Hence, evidence of non-random arrivals were found as random arrivals would have implied $W = 0.5h$. Jolliffe and Hutchinson (1975) further improved the estimation of passenger waiting times by analysing the influence of day-to-day variability of bus arrivals. The study also proposed a three-fold categorisation of passengers based on their behaviour: (i) those who arrive to minimise their waiting time, (ii) those who arrive randomly, and (iii) those whose arrival coincides with the bus, i.e. by running to catch it. Based on an analysis of ten bus stops with varying headways at 6–31 min in London they found that actual passenger waiting times were 30% less than if passengers arrived randomly. Bowman and Turnquist (1981) extended the estimation of passenger waiting times by modelling explicitly the arrival distribution of passengers timing their arrival at stops by using a decision model of arrival time choice. By using this model the study found that passengers are more sensitive to reliability than to scheduled headways when deciding between arriving randomly or timing their arrival.

Later studies investigated aggregate passenger arrival patterns with the purpose of estimating the threshold for when passengers arrive randomly. Fan and Machemehl (2002) found that a 10 min headway was the limit between random and non-random passenger arrivals based on 2491 observations of bus passengers in Austin, Texas. In a later study, the same authors found that 11 min was the transition point between practically random arrivals to less random arrivals, and that all passengers timed their arrival at headways of over 38 min (Fan and Machemehl, 2009). Similar results were found in a study on bus passenger arrivals in Trondheim, Norway, where passengers were found to time their arrival at headways of 10 min (Nygaard and Tørset, 2016). These studies only analysed arrival patterns of bus passengers. Luethi et al. (2007) analysed passengers across public transport modes, i.e. bus, tram, and commuter rail based on data from Zürich, Switzerland. The study estimated the threshold beyond which a significant group of passengers timed their arrival even lower at 5 min. This study also found that the share of passengers arriving non-randomly was higher at 5-min headways than at 6-min headways, hence highlighting the importance of timetables that are easy to remember.