Real-time estimation of bus passenger OD patterns based on AVL data

Wenzhe Sun[1,*], Jan-Dirk Schmöcker[1], Koji Fukuda[2]

[1]Graduate School of Engineering, Kyoto University

[2]Hitachi Kyoto University Laboratory

*Corresponding author: wz.sun@trans.kuciv.kyoto-u.ac.jp

Introduction

Bus operators are confronting the problem of delay and in-vehicle overcrowding in many cities. Delay might be caused by unexpectedly long dwell time which could be further amplified by unplanned large demand and prolonged boarding or alighting time due to in-vehicle overcrowding. Being able to monitor the number of passenger on aboard as well as boarding and alighting passenger in real time potentially helps operators to deploy better and faster control strategies.

Several data sources give operators information on how many passengers are in the vehicle. However, APC (Automatic Passenger Count) or AFC (Automatic Fare Collection) data are still not available (or accessible) in many cities including in Kyoto City, Japan, where further few bus lines are equipped with cameras or other types of passenger counters. Besides, smart card data is operated by various companies who are not willing to share the data. This motivates this research to estimate passenger dynamics by bus AVL (Automatic Vehicle Location) data only.

Dwell time at each stop are important information that can be obtained from bus AVL data and that indicate passenger dynamics. In this study, a methodology to estimate bus passenger dynamics regarding boarding and alighting numbers at each stop along a bus route as well as the change in on-board passenger numbers by using bus AVL data is discussed. MCMC (Markov Chain Monte Carlo) modelling is used to sample the number of boarding and alighting passengers for each run at each stop given dwell time, headway and capacity constraints as well as a dwell time function.

We discuss that several boarding patterns can be observed and need to be considered: (1) passengers board and alight simultaneously; (2) passengers alight first because of in-vehicle overcrowding; (3) passengers alight first due to stop configuration reason. Results with data from two Japanese Cities will be presented.

Methodology

The notation used in this study is as follows.

$D_{i,k}$	dwell time of bus run *k* at stop *i*
$W_{i,k}$	passenger activity time of bus run *k* at stop *i*
$A_{i,k}$	number of alighting passenger of bus run *k* at stop *i*
$B_{i,k}$	number of boarding passenger of bus run *k* at stop *i*
$Q_{i,j,k}$	number of passenger boarding bus run *k* at stop *i* and alighting at stop *j*
$\Delta_{i,k}$	headway between front bus and bus run *k* at the arrival of stop *i*
$a_{i,j}$	arrival rate per minute of passengers boarding at stop *i* and alighting at stop *j*

For each bus run *k*, let arrival rate $a_{i,j}$ denote the arrival number per minute of passenger who board at stop *i* and alights at stop *j* and assume the actual arrival number follows a Poisson distribution, then the passengers travelling OD pair (*i*,*j*) accumulated since last bus departures can be obtained as (1)

$$Q_{i,j,k} \sim Poisson(a_{i,j}\Delta_{i,k}) \tag{1}$$

Let N denote last bus stop number, the number of passenger who board at stop *i* and alight at stop *j* can be derived in (2) and (3) accordingly

$$B_{i,k} = \sum_{j=i+1}^{N} Q_{i,j,k} \tag{2}$$

$$A_{i,k} = \sum_{j=1}^{i-1} Q_{j,i,k} \tag{3}$$

Let t^a and t^b denote average alighting and boarding time per passenger respectively and simply assume passenger boarding and alighting activity starts at the same time, activity time thus is the maximum of total boarding time and total alighting time, which can be obtained as (4)

$$W_{i,k} = \max(t^b B_{i,k}, t^a A_{i,k}) \tag{4}$$

Let constant c denote vehicle activity time which is the aggregation of time required for door opening, door closing, deceleration and acceleration time. Then the dwell time can be written as (5)

$$D_{i,k} = W_{i,k} + c \tag{5}$$

Dwell time and headway are the information derived from bus AVL data. Average boarding and alighting time and vehicle activity time can be calibrated by in-vehicle survey data in a small sample or according to experience. Markov Chain Monte Carlo (MCMC) method is used here to sample the arrival rate $a_{i,j}$ that fits observed dwell time of each bus run at each stop. As the results obtained from the MCMC method is the distribution of parameter, a range is given for the arrival rate of each OD pair. To decrease the sampling dimensionality, the gravity model and matrix factorization method are used to approximate the arrival rate matrix with less parameters.

Experiment and validation

In order to validate the methodology, two experiments are conducted in this study. For Experiment 1, smartcard data from a Japanese city is used. Dwell time and headway parts in bus AVL data are replicated by first and last user entries at each bus stop which are considered as arrival time and departure time respectively. Average boarding and alighting time is calibrated by smartcard data. The MCMC model output is compared with the real OD matrix derived from the same data source. Experiment 2 applies the methodology to AVL data from Kyoto. An in-vehicle survey is conducted for validation in terms of number of boarding and alighting passengers. In Kyoto, severe in-vehicle overcrowding and bus bunching can be observed, resulting in boarding and alighting activity starting often sequentially. Therefore, an adjusted dwell time model is implemented. Three bus lines with different OD patterns are investigated in this survey, and the robustness of the proposed methodology under crowding and stochastic circumstances is discussed.

Keywords: Boarding/alighting dynamics; OD estimation; Dwell time modelling; Markov Chain Monte Carlo; Bus AVL data