Frequency-based Path Choice Models from Smart Card Data

Neema Nassir (MIT), neeman@mit.edu
Mark Hickman (University of Queensland), m.hickman1@uq.edu.au
Zhenliang Ma (University of Queensland), z.ma1@uq.edu.au

Smart card transaction data represents a considerable source of passively-collected information on passenger travel. With geographic coordinates and time stamps for these transactions, it is possible to identify a passenger’s origin, destination, time of travel, and path (collection of route segments) of a journey. When such data are available for a clear majority of passengers, this characterization of journeys allows a more intimate view at potential path choices made by these passengers.

In this study, we examine path choices in light of higher-frequency bus and rail services, where combined headways at a given stop are 15 minutes or less. In such circumstances, it is commonly assumed that passengers may arrive randomly to a stop, rather than timing their arrival with a specific scheduled vehicle departure. Further, it has been hypothesized that passengers arriving to stops with high frequency may actually behave according to a “strategy”, by choosing among different services (routes or route segments) that will take them toward their destination (Spiess and Florian, 1989). In a deterministic choice context, such a strategy is associated with a so-called “attractive” set of routes: a passenger selects a set of routes departing from the stop, and will board the next vehicle to depart from among that set of routes. Extensions have examined strategies under crowded or congested transit services (de Cea and Fernández, 1993).

Several researchers have observed such passenger strategies, or related path choice behaviour, by analysing smart card data. In large rail networks, passengers often have more than one path available between their origins and destinations. In such cases, researchers observe that passengers choose different paths, although the transfers in the rail network are not always explicitly observed. The challenge in this research is to identify the transfer locations, and then to describe passenger path choices in terms of level-of-service attributes, such as travel time, number of transfers, crowding, waiting time, and network topology. Examples of path choice research using smart card data in rail networks include: Kusakabe et al. (2010), Guo and Wilson (2011), Asakura et al. (2012), Zhou and Xu (2012), de Grange et al. (2012), Sun et al. (2012, 2015a, 2015b), Raveau et al. (2014), Raveau and Muñoz (2014), Fu et al. (2014), Zhu et al. (2014, 2015), Hong et al. (2015), and Lee and Sohn (2015).

Other recent research has used smart card data for path choice among bus passengers. More general research for bus path choice using smart card data was presented by Jánošíková et al. (2014), using a bus network in a small city. More explicit consideration of passenger strategies was discussed and modelled by Schmöcker et al. (2013), Viggiano (2013), and Kurauchi et al. (2014). These latter works have used selected bus corridors and origin-destination pairs, for the following reasons. Unlike rail networks, bus networks (or networks with multiple transit modes) have a very large number of origins and destinations, and a large number of possible paths for each origin-destination pair. The result is that it becomes computationally prohibitive to identify “reasonable” routes, and subsequently to estimate route choice models, across a large network. This computational challenge is only compounded by the very large number of smart card transactions.
We have recently proposed a framework to use network-wide smart card data for path choice modelling. This framework uses a link-based formulation, rather than a path-based formulation, to identify passenger decision-making. This approach was originally advocated by Nguyen et al. (1998), but without investigating methods for model calibration or empirical analysis. Recently, Fosgerau et al. (2013) and Mai et al. (2015) have presented methods to empirically estimate traffic route choice models using a link-based formulation. Our framework builds off these previous works to estimate transit passenger path choice models.

The data used in our models is taken from South East Queensland (SEQ), including the greater Brisbane metropolitan area, in Australia. We are fortunate to have tag-on, tag-off data from smart card transactions across a multi-modal network, including bus, rail, and ferry services. Using the individual transaction records, we are able to estimate path choice models that include in-vehicle travel time, transfers, transfer time, and related variables. Such a random utility-based model can improve upon the existing strategy approach that is currently used in SEQ.

References


