Estimating Transit Travel Time Components
Based on Smart Card Data and Operational Records

Amr M. Wahaballa¹, Fumitaka Kurauchi², Toshiyuki Yamamoto³, Jan-Dirk Schmöcker⁴

¹Lecturer, Faculty of Engineering, Aswan University, Egypt, amr.whbala@aswu.edu.eg
²Professor, Department of Civil Engineering, Gifu University, Japan, kurauchi@gifu-u.ac.jp
³Professor, Institute of Materials and Systems for Sustainability, Nagoya University, Japan,
yamamoto@civil.nagoya-u.ac.jp
⁴Associate Professor, Department of Urban Management, Kyoto University, Japan,
schmoecker@trans.kuciv.kyoto-u.ac.jp

Due to the increased popularity of smart cards, a large amount of passengers travel data is available. Consequently, a wide range of travel choice prediction and estimation models utilising smart card data have been proposed during the last decades. This huge data is expected to reduce the need for conducting traditional manual surveys. But, most research concerns the estimation of travel time components and still use supplementary survey data in addition to the smart card data. More precisely, for transit rail passengers walking time between entry/exit platforms from/to transaction gates and transfer time among different lines are difficult to be approximated. In addition, transit routes that the passengers actually chose are in most cases not recorded, especially for journeys including transfers between lines. To supplement this information, there is a need for additional data sources and some theoretical and mathematical effort to deal with the problem.

The objective of this research is therefore to find a bridge between correlating a passenger’s route choice to an observation of his/her total travel time observed by smart card data. This may be achieved by estimating the distributions of travel time components for each single path line, then estimate transfer time distribution on transfer stations and finally using these distributions to estimate path choice proportions. This idea has been proposed by Zhou et al. (2015) assuming that all travel time components yield normal distributions except waiting time following uniform distribution. Sun and Xu (2012) assumed that waiting and walking time are both random. A Bayesian inference approach has been proposed by Fu et al. (2014) to analyse probabilistic route
choice of individual passenger. They used zone journey time and tested Gaussian and lognormal distributions. None of previous research has though considered actual service arrival times and all of them utilised surveying data for some of travel time components. This research will consider more realistic data sources toward finding more suitable distributions. To overcome previous behavioural assumptions and manual survey effort, this research uses smart card data supported with service operation records of different transit lines. This approach will be examined utilising Oyster card data for London underground in addition to working service statistics provided by Transport for London (TfL) for the same time period. In addition, the maximum walking time between gates and platforms of each station and the maximum transfer time are also provided by TfL. For London underground, Oyster card records the transactions at the entry and at the exit. Oyster card data for four weeks from Oct. 16th 2011 to Nov. 12th 2011 is available. The data contains entry/exit travel time and location and passenger ID in addition to the day of the transaction.

The total travel time can generally be divided into several components; entry walking time from the gate to the platform, the platform waiting time, on train travel time, transfer walking time, transfer platform waiting time, on train travel time (for the second line) and exit walking time from the platform at destination station to exit gate. On the basis of the independence of time components, a method is designed to infer the platform waiting time and on train travel time distribution parameters. The entry/exit walking time values are estimated based on the maximum values provided by TfL for each station with setting some reasonable assumptions for their minimum suitable values. The headway of trains are obtained from the working time table’s data and service performance records which are also used for approximating punctuality of train running times. At first, the single path station pairs are considered for stations on the same line where transfer is not required to estimate the platform waiting time and on train travel time distributions parameters for each station then the distribution parameters of the travel time components for each transit line are estimated. The future research is using these parameters to estimate the transfer walking time, transfer platform waiting time distribution parameters for transfer stations for all lines which may help to estimate the hyperpath choice behaviour for OD pairs containing more path options.
References

