Investigating urban bus travel time reliability patterns in London using iBus Automatic Vehicle Locating and Live Bus Arrivals data

Selini N. Hadjidimitriou [1], Ioannis Kaparias [2], Mauro Dell'Amico [1]

[1] University of Modena and Reggio Emilia, Italy
[2] University of Southampton, UK

Introduction

The importance of the variability of travel time in road networks, which results in uncertainty when attempting to predict it, has long been recognised in the literature. Several studies have been carried out in that field and have almost unanimously reached the same conclusion: that although travel time is an important factor affecting the traveller's route choice behaviour, low travel time variability (or high reliability) can be even more important. As opposed to road networks, however, where traffic congestion can be fairly easily identified as the sole source of uncertainty, passengers in public transport networks may be exposed to delays arising from a number of sources, such as service reliability and overcrowding. Due to the prevailing data scarcity in the public transport field up until recently, nevertheless, the causes and effects of such phenomena have been mostly analysed at a theoretical level, with practical aspects having received relatively little attention in the literature.

Recent advances in the field of Intelligent Transport Systems (ITS), however, have changed the scenery as concerns data availability. The introduction of innovative public transport information services now provides new forms of data in vast quantities, which can be used to understand the behaviour of travellers on one hand, but also to act as a proxy for service characteristics on the other. In London, for instance, a valuable data source can be identified in the information requests made to Transport for London's (TfL) "Live Bus Arrivals" (LBA) service [1], which offers travellers real-time countdown information to the arrival of specific bus lines at specific stops upon request through a web or mobile app interface. The service is empowered by data from TfL's "iBus" Automatic Vehicle Locating (AVL) system.

The aim of the present study is to investigate the link between the usage of the LBA service (as exhibited by the number of requests for real-time bus arrivals to the service) and the reliability of bus travel times. Such analysis can be of value to public transport operators, as it could help identify locations around the network where the travel time is frequently longer than planned for, and where inefficiencies potentially affecting customer satisfaction may be present (e.g. overcrowded busses at specific times). More importantly, the analysis of the interrelations between these two data sources can enable the development of advanced models, which are able to predict the travel time reliability at each stop and bus line on the basis of historical data on the number of requests for real time arrivals alongside other characteristics.

Study area and dataset

The study uses the bus network of the London Borough of Harrow (Figure 1) as a case study, one of the 32 boroughs of Greater London, with an area of 50 km^2 and a population of 250,000 [2]. Data from two sources are obtained for each of the 13 lines serving the Harrow bus network for a 28-day period, and specifically between 16 July and 12 August 2012:

- A log of all requests/queries made by bus users to the LBA service, containing information on date, time, bus stop, bus line and type of device used.
- A record of timestamps from the iBus AVL system, providing information on the time of arrival at and departure from each stop of each vehicle.

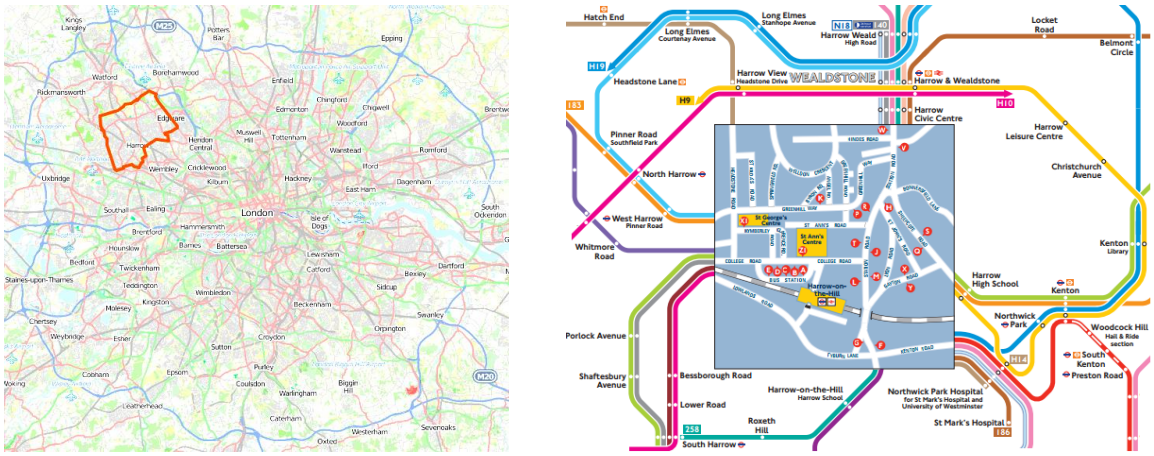

Figure 1: Left: Location of Harrow [3]; Right: Harrow bus network [4]

Classification of LBA users

Previous work by the authors [5] analysed an LBA dataset for the same period but for the entire London bus network and classified user requests to the service using a Support Vector Machine by Sequential Minimal Optimisation (SVM-SMO) approach. Six main user behaviours were identified, thus providing an insight into the usage of the service by different categories of travellers, namely:

- waiting users (querying live bus arrivals at one particular bus stop);
- stop explorers (querying bus arrival times at a number of nearby-located stops);
- route explorers (querying one particular line without requesting bus arrival times);
- deep explorers (querying information and arrival times for a particular line and stop);
- interchange users (querying different connected lines and stops); and
- outliers, i.e. users that cannot be allocated to one of the defined categories.

The present study considers the LBA requests by waiting users, route explorers, deep explorers and interchange users only. Given that the users of these four categories are most likely querying bus arrivals whilst already en route, high numbers of requests from them may suggest user restlessness and abnormally long waiting times, and may hence be potentially associated with travel time unreliability.

Calculation of reliability

The reliability of individual bus lines is calculated using the method of Kaparias et al [6], which consists of two reliability indices for earliness and lateness based on central tendency and dispersion measures of the bus travel time distribution. Specifically, the lateness reliability index is used here, which for a link l is defined as:

$$r_L(l) = exp[½ \cdot T_{log}(l) - z_{0.05} \cdot \sqrt{T_{log}(l)}] \quad (1)$$

with

$$T_{log}(l) = ln\left(1 + \frac{var[t(l)]}{[\bar{t}(l)]^2}\right) \quad (2)$$

where $t(l)$ is the travel time on link l, following a log-normal distribution with mean $\bar{t}(l)$ and variance $var[t(l)]$, and $z_{0.05}$ = 1.65 is the tail probability of the 90%-confidence level employed.

The focus is on the calculation of the reliability of the in-vehicle travel time for each link (line segment between two stops) and each line, aggregated by time of the day and day of the week

(weekday or weekend). In-vehicle time is obtained from measurements of the difference in the departure timestamps of the bus at subsequent stops, as provided by the iBus system. As such, the lateness reliability index is computed for each hourly time interval over the period covered by the dataset at each stop of the 13 bus lines, based on the travel time mean and variance on the link immediately preceding the stop (i.e. the bus line segment connecting the stop immediately upstream with the stop in question). Figure 2 shows an example of the computed reliability index values at each stop of bus line 114 between 8.00 and 9.00 on 18 July 2012.

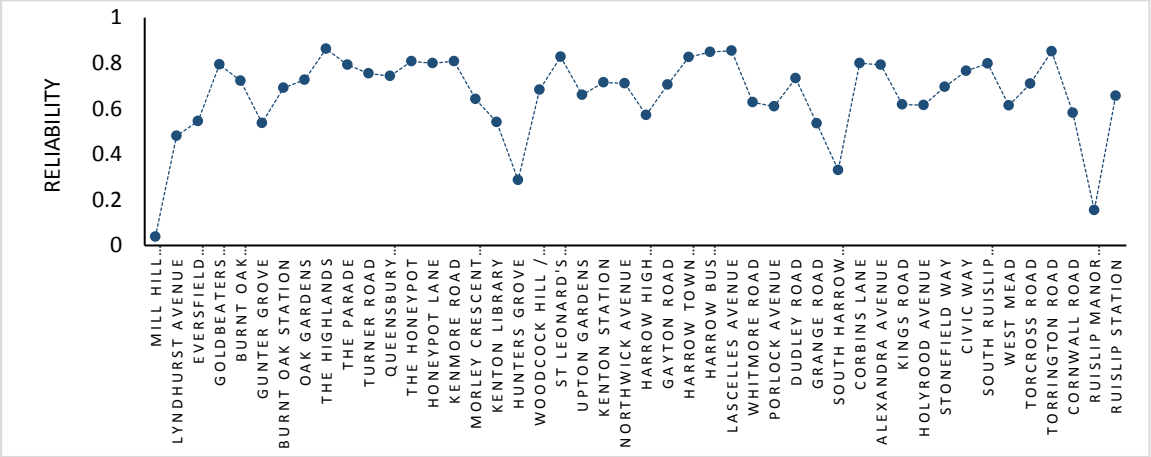

Figure 2: Lateness reliability index of bus line no. 114 between 8.00 and 9.00 on 18 July 2012

Analysis and results

The analysis first concentrates on filtering the relevant LBA request records from the complete dataset and assigning them to their corresponding bus stops and time intervals. This results in a certain number of LBA requests for any bus stop and time interval, which can be associated with a corresponding value of the lateness reliability index.

Correlation analysis is, hence, carried out for each time interval and day to investigate the link between travel time reliability and LBA request patterns. Example calculations for two time intervals are shown in Figure 3. Results overall indicate that, as conjectured, high numbers of LBA requests by users are generally correlated with lower values of travel time reliability for the bus lines concerned, which suggests that LBA system usage could potentially be used as a proxy for bus travel time reliability.

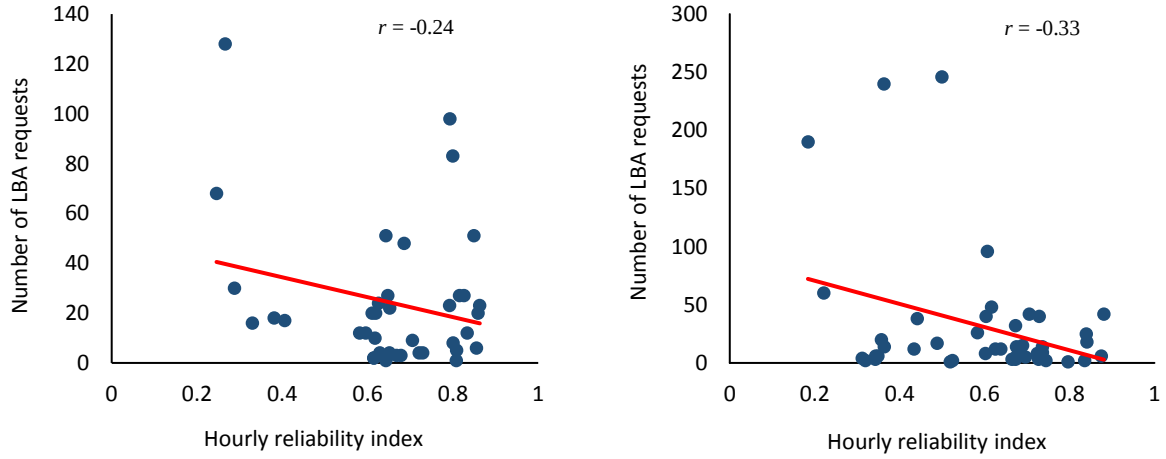

Figure 3: Correlation between LBA requests and lateness reliability index in Harrow on Wed. 18 July 2012: left: morning peak (8.00-9.00); right: evening peak (18.00-19.00)

Concluding remarks

The initial results from this study provide an interesting insight into the potential of using the LBA system to predict bus travel time reliability. However, LBA system usage is likely to be influenced by numerous other factors that have not been considered here, such as unique line and bus stop characteristics, as well as demographic factors, and it is the objective of future work to explore them. Also, a limitation of the study is that, due to the nature of the iBus dataset, the in-vehicle time currently incorporates the dwell time of vehicles at bus stops, which is heavily dependent on the boarding and alighting passenger volumes; it is planned to separate this in future steps of the research through the usage of additional datasets, such as on-board cameras and Automatic Passenger Counting (APC) systems. Finally, further research will focus on deepening the analysis further, and in particular on the derivation of predictive models and algorithms for bus travel time reliability.

Acknowledgements

The authors would like to thank TfL Buses for providing the iBus and LBA data used in the study.

References

[1] Hardy, N. (2012), "Provision of bus real time information to all bus stops in London", *19th World Congress on Intelligent Transportation Systems*
[2] UK Office for National Statistics (ONS) (2011), *2011 Census*
[3] OpenStreetMap (2017), www.openstreetmap.org.
[4] Transport for London (2017), Official website
[5] Hadjidimitriou, S.N., Mamei, M., Dell'Amico, M., Kaparias, I. (in press) "Classification of Live Bus Arrivals user behaviour", *Journal of Intelligent Transport Systems*
[6] Kaparias, I., Bell, M.G.H., Belzner, H. (2008), "A new measure of travel time reliability for in-vehicle navigation systems", *Journal of Intelligent Transport Systems*, 12, 202-211