Consumer Data Research Centre

An ESRC Data Investment

movement strategies

Forecasting travel demand for the Elizabeth line using smartphone GPS data

Yi Cheung[1], Tao Cheng[1], Aidan McInerney[2] and Caroline Christie[2]
[1]University College London, [2]Movement Strategies

Project Background

Travel demand models of transport systems are able to provide many insights for urban planning and transport infrastructure management use. A travel demand model is usually represented by an origin-destination (OD) matrix to capture the flow between two locations. Using smartphone-based GPS data to extract OD patterns has been proven to be an efficient alternative to the traditional household-based survey approach to represent the travel demand of the city. With the popularisation of smartphones generating GPS data, more detailed information about individuals' travel behaviours can be provided at low cost. This study aims at developing a GIS-based methodology to estimate travel demand. It modelled the weekday morning peak ridership of the central section (10 stations – see Figure 1) of the Elizabeth line as an example.

Data and Methods

The data used in this research consisted of approximately 586 million GPS events generated by around 2.57 million users during December 2017 in London. The methodology is comprised of three parts. Firstly, we extracted OD flows of users who live and work in the 10 studied stations during the first 15 weekdays of the month. We referenced the activity-based approach to determine 3 parameters: the appearance time of home and work locations, the buffer distance of each station catchment, and the number of days each smartphone user was in London. This was done in order to capture sample commuters who live and work in 10 buffered zones by regression analysis. Then, all captured users were assigned home and work stations based on their most frequently occurring locations. After that, the home and work location of users were linked to become the OD flows of 10 stations, which were converted into an observed OD matrix for further calculation. Secondly, we applied 3 scaling factors for each of the 10 buffer zones, including actual population at work age (16-64), transport mode share statistics, and the transport route shifting probability on each OD flow. Finally, the observed OD matrix was multiplied by 3 scaling factors.

Key Findings

This study provided data identification and classification methods to: capture smartphone users' living and working's spatial distributions in London; to investigate smartphone users' commuting patterns near the 10 stations along the central section of the Elizabeth line; to estimate the potential commuters the section during weekday morning peak; and to assess the potential application and constraints of the methodology. Overall, the research observed nearly 8,000 people of over 421,000 filtered users who live and work in the 10 catchment areas. After scaling up the observed users by 3 scaling factors, the estimated entry and exit numbers of each station and the ridership flows eastbound and westbound were plotted. From the travel demand model, 11.1% (24,362) of the work population are employed in the same catchment as their home location, while 6.3% (13,749) of them commute to one of the other 9 stations of the Elizabeth line under study.

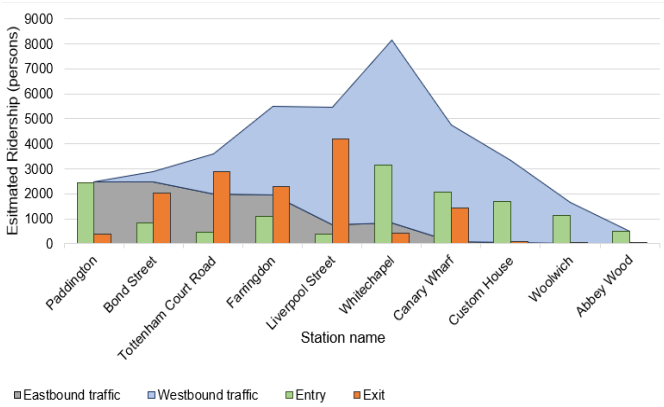

Figure 1: *The estimated ridership and traffic direction volumes during morning peak*

The relationship between dwell time and distance was also explored. The principle of 'gravitational' theory states that shoppers are more likely to shop in a more attractive town or shopping centre (i.e. in this case, number of visits) but that the attractiveness decreases with distance. This study found there is a relationship between dwell time and distance in a retail agglomeration however there is room for improvement.

Value of the Research

This initial attempt to developing a cost-effective approach in a travel survey sample succeeded in effectively increasing the sample size over 15 times in terms of sample fraction. It did so by using a large amount of smartphone GPS data and combining referenced travel behaviour datasets to forecast morning commuting patterns. Subject to the availability of reference datasets, the methodology of this research can be further customised and applied to extract travelling patterns for different time periods and events. By enlarging the scope of analysis to more locations, compiling more census data and more validation information, the accuracy and customized applications of this type of GPS-based travel demand model will be greatly improved.