London Passenger Mode Choice

Tim Hillel

October 26, 2019

Context

This case study investigates mode choice on an urban multi-modal transport network. The objective was to be able to predict how people will react to changes to the transport network and conditions, to allow for more efficient transport network management and investment planning.

Trip diaries from travel surveys are a commonly used source for historical trip data. However, like other revealed-preference (RP) data sources, they only contain details of the selected itinerary, and do not contain details of the choice-set. To address this, a data fusion framework was developed to add individual specific mode-alternative level-of-service (LOS) variables (e.g. in-vehicle travel time, public transport fares, fuel cost, etc.) to historical trip records. The approach uses an online journey planner to generate the routes for four transport modes (walking, cycling, public transport, and driving). A closely tailored cost-model is then used to calculate individual specific public transport fares and driving fuel and congestion charge costs for each trip.

The London Passenger Mode Choice (LPMC) dataset was created by using this framework with trip diary data from the London Travel Demand Survey (LTDS), a rolling household travel survey collected by Transport for London (TFL). This dataset is used to predict mode choice out of

- walking,
- cycling,
- public transport, and
- driving.

Full details of the framework, dataset, and the models it was used to develop are given in Hillel et al. (2018). Please cite this paper in any work using this dataset.

Data Collection

Three years of trip diary data were used to create the LPMC dataset, from April 2012 to March 2015. The data was processed to remove all trips outside of London, as well as routes which did not use TfL services. In total, the LPMC contains details for 81 086 trips made by 31 954 individuals across 17 616 households.

The sampling was performed by TfL when collecting the original travel survey. Expansion factors are provided for using the data for forecasting, though are not provided with the LPMC.

Variables and Descriptive Statistics

The variables of the dataset are described in Table 1, and the descriptive statistics are summarized in Table 2.
Note that the total public transport duration (if needed) can be calculated as the sum of dur_pt_access, dur_pt_rail, dur_pt_bus, and dur_pt_int. Similarly, the total driving cost can be calculated as the sum of cost_driving_fuel and cost_driving_ccharge.

Variable	**Min**	**Max**	**Mean**	**St. Dev.**
trip_id	0	81085	40542.5	23407.66
household_id	0	17615	8709.5	5070.38
person_n	0	9	0.8	1.06
trip_n	0	18	1.54	1.77
travel_mode	1	4	3.06	1.08
purpose	1	5	2.9	1.18
fueltype	1	6	2.83	2.22
faretype	1	5	2.27	1.74
bus_scale	0	1	0.65	0.47
survey_year	1	3	1.99	0.81
travel_year	2012	2015	2013.18	0.9
travel_month	1	12	6.69	3.33
travel_date	1	31	15.36	8.74
day_of_week	1	7	3.96	1.94
start_time	0	23.93	13.83	4.46
age	5	99	39.46	19.23
female	0	1	0.53	0.5
driving_license	0	1	0.62	0.49
car_ownership	0	2	0.98	0.75
distance	77	40941	4605.26	4782.35
dur_walking	0.03	9.28	1.13	1.12
dur_cycling	0.01	3.05	0.36	0.35
dur_pt_access	0	1.19	0.16	0.09
dur_pt_rail	0	1.47	0.09	0.18
dur_pt_bus	0	2.15	0.17	0.19
dur_pt_int	0	0.87	0.04	0.08
pt_interchanges	0	4	0.37	0.62
dur_driving	0	2.06	0.28	0.25
cost_transit	0	13.49	1.56	1.54
cost_driving_fuel	0	10.09	0.83	0.82
cost_driving_ccharge	0	10.5	1.07	3.18
driving_traffic_percent	0	1.25	0.34	0.2

Table 2: Descriptive statistics

Name	Description
trip_id	Unique numerical identifier for each trip
household_id	Unique numerical identifier for each household
person_n	Id of the person (within household)
trip_n	Id of the trip (within person)
travel_mode	Id of the chosen mode (1: walk, 2: cycle, 3: public transport, 4: drive)
purpose	Id of purpose of trip (1: home-based work, 2: home-based education, 3: home-based other, 4: employers' business, 5: non-home-based other)
fueltype	Id of fueltype of proposed vehicle for trip: (1: petrol car, 2: diesel car, 3: hybrid car, 4: petrol LGV, 5: diesel LGV, 6: average car)
faretype	Id of rail faretype for person making trip: (1: full-fare, 2: 16+, 3: child, 4: disabled, 5: free)
bus_scale	Bus fare scale of person making trip (0: free bus journeys, 0.5: half price, 1: full price)
survey_year	Survey year: (1: 2012/13, 2: 2013/14 or 3: 2014/15)
travel_year	Trip travel year, from 2012 to 2015
travel_month	Trip travel month, from 1 (January) to 12 (December)
travel_day	Trip travel day-of-month, from 1 to 31
day_of_week	Trip travel day-of-week, from 1 (Monday) to 7 (Sunday)
start_time	Linearised trip start time from 0-24, e.g. 16:45 becomes 16.75
age	Age of person making trip in years
female	1 if the person is female, 0 otherwise
driving_license	1 if the person has driving licence, 0 otherwise
car_ownership	Id of car ownership status of household (0: no cars in household, 1: less than one car per adult, 2: one or more cars per adult)
distance	Straight line distance between trip origin and destination in metres
dur_walking	Predicted duration of walking route in hours
dur_cycling	Predicted duration of cycling route in hours
dur_pt_access	Predicted total access and egress time for public transport route in hours
dur_pt_rail	Rail in-vehicle time for public transport route in hours
dur_pt_bus	Bus in-vehicle time for public transport route
dur_pt_int	Interchange time on public transport route in hours
pt_interchanges	Number of interchanges on public transport route
dur_driving	Predicted duration of driving route in hours
cost_transit	Estimated cost of public transport route in GBP
cost_driving_fuel	Estimated fuel cost of driving route in GBP
cost_driving_ccharge	Estimated congestion charge cost of driving route in GBP
driving_traffic_percent	Predicted traffic variability on driving route

Table 1: Description of the variables

References

Hillel, T., Elshafie, M. Z. E. B. and Jin, Y. (2018), 'Recreating passenger mode choice-sets for transport simulation: A case study of London, UK', **171**(1), 29–42.
URL: *https://www.icevirtuallibrary.com/doi/10.1680/jsmic.17.00018*