

Real time simulation of daily travel patterns

Charlotte Sturley
University of Leeds

Project Background

The proliferation of smartphone ownership and app usage has enabled new forms of data collection relating to transport behaviour. Currently, data are collected via transport surveys, which have small sample sizes, are time consuming to carry out and often contain inaccurate information. Smartphone technology has the potential to generate individual level in real time.

Catch!, a project funded by Innovate UK, aims to understand travel behaviour in greater detail using data crowd-sourced via a smartphone application and to provide insight to Local Authorities to improve the transport network.

This particular part of the project involves developing a solution to extrapolate sample collected data to the whole population.

Data and Methods

Data collected via the WeCycle smartphone application was provided by TravelAi. This contained a database of GPS tracks of journeys recorded by the app. Home and work locations of app users were inferred from the GPS data. A database of app users was also supplied which contained a set of attributes including gender, year of birth and weight.

Census data were obtained via the UK Data Service. Demographic data relating to age, gender, economic activity and place of work were extracted at small area level. The 2011 Output Area Classification, which classifies areas into socio-economic groups, was sourced from the Office for National Statistics.

A microsimulation approach was taken to extrapolate sample app data to the whole population at city level. The city of Newcastle-upon-Tyne was chosen as a case study for the microsimulation model. The population of Newcastle was represented at an individual level within the model.

Population synthesis was carried out to generate a synthetic population of individuals from the aggregate Census tables. A Monte Carlo simulation approach was applied. Travel characteristics were attributed to individuals in the synthetic population by linking them to individuals in the app user database based on demographic characteristics.



Figure 1: Example GPS tracks

Key Findings

A methodology has been developed to generate a synthetic population of individuals from Census data for a Local Authority area. A process to link a synthetic population of individuals to crowd-sourced smartphone application data has also been defined. The process for building a microsimulation model has been automated in R so that the outputs are reproducible.

There are some limitations to the model. Currently, there are only a small number of app users. Therefore broad age-gender categories were used. With more data these could be refined.

Value of the Research

This project has demonstrated a methodology for linking new forms of data to traditional datasets to provide greater insight into individuals' travel behaviour.

The synthetic population generated in this research could be used as the basis of an agent-based model to simulate the movement of individuals across a city over time. Such a model could be used to identify pinch points in the transport network and inform scheduling of transport services.

The research also highlights a number of considerations when working with crowd-sourced smartphone application data, notably around data quality. For example, the accuracy of the GPS data and missing demographic data relating to app users.