DATA & MODELLING

THE LATEST TECHNIQUES AND APPLICATIONS FOR THOSE AT THE CUTTING EDGE OF THE TRANSPORT MODELLING WORLD

- Nanosimulation comes of age
- Modelling 'shared space'
- Forecasting traffic in real time
- Using new data sets for transport modelling
- Using technology to model... technology
- The rise of electronic data capture

Sponsored by

Mott MacDonald
intelligentdata
LOCAL TRANSPORT TODAY
Using BIG data in transport modelling

Both innovation and ingenuity are required to turn an increasing amount of data that is not obviously ‘transport oriented’ into useful information that can assist with model building in circumstances where relevant data is otherwise not available or in poor supply, explains Omnitrans’ Erik de Romph.

Companies such as TomTom possess enormous databases with vehicle-movements based on GPS data to predict traffic jams in real time and, with the introduction of public transport smart cards, large databases with public transport movements are a potential source of useful matrix building data.

The most interesting new data source, however, is GSM (Global System for Mobile Communications) location data, which is location-based information retrieved from mobile phones. Each mobile phone is at all moments connected to a certain GSM cell site antenna. Such a cell, provided by a mobile provider, knows what mobile phones are present in that cell. The basic version only knows that a particular mobile phone is present, the more advanced cells know from what direction (angle) the phone is connecting to the cell antenna. The most advanced cells also return an indication of the distance from the mobile phone to the cell antenna. This results in fairly accurate positioning of a phone. When the phone moves, for example during a car ride, the mobile phone switches over from one cell to another cell. Monitoring the movement of an anonymised mobile phone through the provider’s network gives the desired data on mobility.

With such data, telecom providers can locate every cell-phone in their network and many new research projects are underway with the objective of utilising this ‘big data’ in some relevant way. But, unfortunately, few of these projects are focused on utilising the resource for transport modelling, even though there is the potential to extract some very useful information relating to the movements that people are making - an obvious possibility is to see if Origin-Destination (O-D) matrices can somehow be extracted. If yes, the cost of data collection would decrease and the accuracy of models and their validation would improve. It would....
View.dat screen showing the number of travellers travelling into Amsterdam on 31 March, 2013 and their origin

also be possible to consider ‘base year’ matrices being updated on a more frequent basis than currently might be the case.

In 2012 Orange Telecom initiated a project called ‘Data for Development Challenge’. It provided a GSM dataset based on network positioning data originating from calls and SMS exchanges between five million of Orange’s customers in the Ivory Coast. This data set, which had been made anonymous, comprised 2.5 billion data records and, in an open challenge, research teams around the world were encouraged to use this data to help address issues relating to ‘the development of society’ in novel ways.

Rising to this challenge, a consortium comprising Goudappel Coffeng and Omnitrans International (from The Netherlands) and KDD-Lab (Italy) was formed to build ‘the best possible’ transport model of the Ivory Coast using only publicly available data (e.g. OpenStreetMap), including the voluminous telephone data supplied by Orange. An objective of the project was to provide ‘proof of concept’ that such data is of value to the transport modelling community.

Special GSM analysis tools, developed at KDD-Lab, were used to process the locations of callers and recipients and tie them to a region. Each region was defined by the GSM cell site antenna’s reception area.

Furthermore, looking at the origins and destinations combined with the time of departure and arrival including the frequency of these trips, showed it was possible to approximate home and work locations. By making this link, it was possible to build average O-D matrices for the morning and evening peaks, and all-day (24 hour) periods; these matrices were assigned to the network using a traditional static assignment.

Although this ‘model’ has not been validated with local traffic counts, or related to the socio-economic characteristics of population or land use, the outputs do provide an insight into travel patterns across the country, and this gives us the potential to provide a modelling tool to test the effects of infrastructural changes in the relatively short term. Nonetheless, the objective of the ‘proof of concept’ was achieved, showing that, in developing countries where the ‘conventional’ data required for transport planning is virtually non-existent, new sources of data can be exploited to provide insight into travel behaviour that it would otherwise have been impossible to achieve.

Out of over a hundred contenders to the Orange challenge, our contribution was chosen among the best. The Ivory Coast project provided us with valuable experience for a new service that has been set up in The Netherlands called View.dat. Using similar GSM data, the anonymous records from five million phones with on average a hundred contact moments per phone are processed each day. The data is aggregated into hourly origin-destination matrices and linked to socio-economic data. In this way numbers of phones are changed into numbers of people and an estimate of the total population is calculated.

The information is interactively available through a website (www.viewdat.nl) where it can be analysed in various ways per city, such as number of visitors and where these visitors came from.

In the above picture a specific hour is chosen for the centre of Amsterdam and the picture shows how many of those people came from the surrounding cities or areas.

The prime objective of View.dat as a service is showing the possibilities of mobile network data to cities. An example of usage is city marketing. For the first time ever a city marketer knows with an enormous sample size of millions of people when people are visiting his city and where they are coming from. Also, the same data can be used to deduce where inhabitants are going.

For transport modelling new data sources will become very important. Not just for validating O-D matrices estimated in a traditional way, but for a new way of transport modelling all together.