Calculation of Origin-Destination data

Laurence Howe, National Traffic Control Centre (NTCC),

In the past, ANPR data has been used for the purpose of measuring travel times or (in the case of law enforcement organisations) to identify individual vehicles. There is now the possibility that it could be used in the strategic planning and analysis process by producing Origin-Destination data.

Using this method of gathering O-D data, it would be possible to carry out useful data analysis that has hitherto been impossible. In particular, the data could be updated dynamically so that it would never become out-of-date. The confidence in such data would be much greater than that associated with currently available data.

INTRODUCTION

Origin-Destination (O-D) data is an important component of transport planning. From the road transport point of view, it enables planners to estimate the current demand between any two locations. By tracking the change in demand over time, it is possible to predict future capacity requirements for the road network. O-D data can also be important in predicting possible revenue streams from road charging schemes. If road-charging schemes were to be introduced, it would be paramount to monitor the effect of the road charging on individual route selection. O-D data does not do this directly on an individual basis, but dynamically calculated O-D data could be used to track trends over time. This would allow any road-charging scheme to be refined, optimising the effect on road usage.

Conversely, if alternative travel links are to be introduced, it is important to estimate the size of the possible market. Current road travel patterns can be useful in estimating the number of passengers who might choose to switch to a new alternative, such as a high-speed rail link. Clearly, O-D data alone is not sufficient to estimate the potential market, but in conjunction with other analysis such as preferred means of travel and tolerance of variability, etc, it can play an important role in this process. It can also be used to estimate the amount of investment that could be switched from road infrastructure to the proposed non-passenger car travel link.

BACKGROUND

Traditionally, O-D data has been collected by interviewing travellers en-route. Typically, they have been asked a standard set of questions, such as 'Where have you come from?', 'Where are you travelling to?', etc. Locations may be on urban or rural feeder roads into the strategic network or at motorway service areas. Questionnaire information can be supplemented by observing the behaviour of vehicles at key intersections.

There are several drawbacks associated with the reliance on such research. Firstly the sample size for each O-D combination may be very small. Secondly, they are only a true reflection of the behaviour of those who mainly make a particular journey on repeated occasions. Thirdly, those who use the strategic network for non-strategic journeys are unlikely to call at motorway service areas and, thus, will not be included in the surveys. This means that it is difficult to estimate the proportions of strategic and non-strategic journeys. Even those who indulge in relatively long-distance commuting, such as from Worcestershire to London, are unlikely to call at a service area on their journey.

People who have more complex travel patterns are likely to fall through the net entirely. For example, someone who lives near Birmingham and regularly travels to Bristol, London, Sheffield and Manchester may make a return journey to each or may make a round trip. The journeys could take place over several days. The capture of such information would be of great benefit to any project established to consider alternative forms of transport. The simple questionnaire approach, even if it captures one of these journeys, is unlikely to unravel the complexity of the whole travel pattern.

ANPR DATA

An Automatic Number Plate Recognition (ANPR) camera reads the number plates of vehicles travelling along the carriageway that the camera monitors. The most common use of ANPR data is to calculate the average travel times of the traffic. Travel time information is used to help control room operators to manage the traffic and to inform drivers of expected journey times. In order to do this without infringing the anonymity of individual road users, a non-reversible cryptographic algorithm (**hash process**) is used, which scrambles the number plates to produce a result known as a **tag**. When the upstream camera and the downstream camera both report the same tag, this is known as a match. The hash process is such that a single tag may represent more than one registration number. However, the likelihood of producing a **false match** from two different number plates (one at the upstream location and one at the downstream location) which produce a journey time that could be interpreted as reasonably achievable by a single vehicle is extremely low.

ANPR data for journey time calculations were originally used over relatively small distances and were thought to be less usable over longer distances because of the increased probability of false matches. However, the structure of the tag has been modified to make the data more useful over much longer distances. Also, it is now

CONTACT DETAILS
Laurence Howe, National Traffic Control Centre (NTCC), Birmingham, can be contacted by email at Laurence.Howe@serco.com

possible to store all the tags generated and therefore to track (anonymously) individual vehicles around the strategic network. So it is within the bounds of possibility to identify travel patterns where people located near Birmingham travel to Bristol, London, Sheffield and Manchester, as hypothesised in the preceding section. It would be possible, for example, to identify vehicles leaving the Birmingham area, travelling south along the M5 and arriving in the Bristol area.

LANE COVERAGE

Because ANPR-based travel times need only to be a representative sample, not all lanes are monitored. Generally, on 2-lane carriageways, only Lane 1 is monitored and on 3 lane carriageways, only Lane 2 is monitored, although there are exceptions. Therefore the raw tag data from the ANPR cameras must be processed before any estimate of total demand can be made. Fortunately, for each carriageway monitored by ANPR cameras, the total traffic flow is measured by a traffic counting **loop site** (inductive loops buried beneath the surface of the carriageway). This information can be used to supplement the ANPR data and create a useful estimate of the total number of point-to-point journeys recorded by the ANPR cameras, as follows.

At each ANPR camera location, the total number of vehicle tags recorded in a given period, $\mathbf{N_T}$, can be counted. Over the same period of time, the associated loop site counts the total number of vehicles, $\mathbf{N_V}$, passing the ANPR camera(s). The inverse tagging ratio, R, for the camera location can be calculated by dividing the total number of vehicles recorded by the loop site by the number of tags (the number of vehicles whose number plates were captured by the ANPR cameras).

$$\mathbf{R} = \frac{\mathbf{N_V}}{\mathbf{N_T}} \qquad (1)$$

For each pair of camera locations, a separate ratio must be calculated for the upstream ($\mathbf{R_{up}}$) and downstream ($\mathbf{R_{down}}$) locations. The subsequent analysis assumes that there is a uniform probability of any vehicle to be travelling in any of the available lanes at each location. This is, of course, not strictly true. For example, for a 3-lane carriageway, HGVs are more likely to be found in Lane 1 than Lane 2, and are not expected to be found in Lane 3. At the other extreme, there will be some cars that spend the majority of the time in Lane 3. Between these two extremes, there will be cars that spend the majority of the time in Lane 2 and others which progress from one lane to another according to the conditions. However, as will be demonstrated, this is a reasonable assumption for the purpose of the calculations. If the uniform probability for all vehicles is assumed, the total number of vehicles making the journey from the upstream location to the downstream location, $\mathbf{N_j}$, can be calculated from the total number of matches, $\mathbf{N_m}$:

$$\mathbf{N_j} \ \mathbf{N_m} = \mathbf{R_{up}} \ \mathbf{R_{down}} \qquad (2)$$

To verify the goodness of this assumption, consider a 3-lane carriageway with ANPR monitoring of Lane 2 only. It is assumed that 20% of the vehicles spend 75% of the time in Lane 1 and 25% in Lane 2 (HGVs, busses and slower cars), 60% of the vehicles spend 50% of the time in Lane 2 and 50% in another lane (average car drivers) and 20% of the vehicles spend 75% of the time in Lane 3 and 25% in Lane 2 (faster then average drivers). By applying Equation 2, the result is an over-estimate of 9%. For a 2-Lane carriageway, using a similar model, the over-estimate is 10%. If the upstream location is a 2-Lane carriageway and the downstream a 3-Lane carriageway, or vice versa, there is no estimation error. Changing the proportion of drivers in the three categories gives a worst error of 12% and a best error of 0%. It is therefore hypothesised that Equation 2 gives a good approximation to the true total number of vehicles travelling from the upstream location to the downstream location.

PRACTICAL APPLICATIONS

The total number of vehicles travelling between any two ANPR locations on the strategic road network could be calculated via Equation 2. Using this equation, vehicles travelling to a succession of destinations could be tracked and totals calculated for different combinations of destinations. Same-day return journeys could be identified and counted. Vehicles making the same journey at the same time of day could be counted for different locations, including long-distance commuters. In addition, the journey patterns of vehicles with foreign number plates could be established and classified by frequency. It would be very easy to establish changing patterns, such as the increase or decrease in particular journey types or the increase or decrease in the number of vehicles from other countries. In the latter case, it would be possible to target specific facilities or services used by non-UK based road users.

CALCULATING AVERAGE JOURNEY LENGTHS

Once it is possible to establish the number of vehicles travelling between two ANPR locations it becomes possible to calculate the average length of the journeys made on different parts of the strategic road network at different times of day. This can be achieved by fitting the data to a model of constant loss. The analysis is based on the following argument.

Consider a number of successive Junctions on a stretch or road. Of those vehicles that started at Junction 1, the fraction continuing on beyond Junction 2 will be $\mathbf{f}$, those continuing on beyond Junction 3, $\mathbf{f}^2$, etc. So, for a stretch of carriageway comprising $\mathbf{n}$ Junctions, the fraction, $\mathbf{F_n}$, of those that started at Junction 1 and remain on the network for all $\mathbf{n}$ Junctions will be given by:

$$\mathbf{F_n} = \mathbf{f}^{n-1} \qquad (3)$$

It can be shown mathematically that, in such a case, the average number of Junctions traversed by all vehicles, $\mathbf{N_{av}}$, is given by:

$$\mathbf{N_{av}} = \frac{\mathbf{1 - f^n}}{\mathbf{1 - f}} \qquad (4)$$

When n is relatively large, Nav equates to the sum of an infinite geometric series

$$\mathbf{N_{av}} = \frac{\mathbf{1}}{\mathbf{1 - f}} \qquad (5)$$

This underpinning assumption of the above argument is that, along a defined stretch of carriageway, a constant proportion of traffic leaves the network at each junction. Clearly, there will be a variation in the proportion at each individual junction, depending on the type of junction. However, over the entire stretch of carriageway, there will be an average loss, which will vary, depending on the local topography, the day of the week and the time of day. Using this analysis, areas could be identified where predominantly strategic (long) or local (short) journeys are made on the strategic road network for various time periods.

CONCLUSIONS

The use of Journey time data derived from ANPR cameras has great potential for transport planning. When used in conjunction with appropriate vehicle count data from loop sites, it offers the possibility of providing up-to-date O-D data on a continuing basis and identifying complex journey patterns. It is also possible to use it for identifying commuting and other regular patterns of travel. The travel patterns of foreign vehicles can also be mapped, enabling better provision to be made for them. Finally, the average length of journeys can be determined for different parts of the country and for different population densities.