WestminsterResearch

http://www.wmin.ac.uk/westminsterresearch

The uses of data from smartcard systems for public transport management.

Peter White
Mousumi Bagchi

School of Architecture and the Built Environment

THE USES OF DATA FROM SMARTCARD SYSTEMS FOR PUBLIC
TRANSPORT MANAGEMENT

Peter White and Mousumi Bagchi
Transport Studies Group
University of Westminster

The growing provision of smartcard systems has been documented in recent articles in
this journal - for example the development of ITSO-compatible systems, and of
contactless cards, in issue 2 of 2007. Examples such as 'Oyster' in London and
'Octopus' in Hong Kong are well-known and now account for high share of public
transport use on those cities. While speed of operation and the ability to check validity
of cards are an obvious reason for this adoption, it is also important to bear in mind
the extensive potential for management information from such cards.

What data are collected?

For each journey, a record is usually obtained of the serial number of the card used,
the fare stage or station at which a boarding took place, the route number, and time of
boarding. Where a card is unique to a named individual (such as a pensioner's
concessionary pass), the trips recorded correspond to those of that person; where a
card is transferable (as for some stored value cards) then journeys by more than one
individual may be recorded. To date, in Britain, only the boarding stage has been
recorded for bus travel, whereas entry and exit reading on the London Underground
permits both ends of the journey to be recorded.

In addition to journey details, characteristics, details of the individual holder may also
be recorded. While name and address are clearly confidential, certain classificatory
data may be useful in identifying trip frequencies and patterns of different user
groups. For example, broad age categories can be defined (child, working age adult,
pensioner concession), type of area of residence (such as inner or outer zones of a
city) and possibly gender. Such variables are akin to those traditionally collected
through manual travel diary surveys (such as the National Travel Survey, or NTS, in
Britain), enabling cross-tabulation of data while maintaining individual
confidentiality. An anonymised serial number may be assigned for research purposes,
unique to a particular card, but not corresponding to its actual number.

Some applications

At the simplest level, smartcard data may be used to estimate total boardings on a
system within a defined period, giving greater accuracy than is possible where
operating staff manually record each occasion a pass is used (such as a bus driver
pressing the appropriate button on an ETM). The validation process also enables
classification of the type described above to be applied.

A current example in Britain is the need to improve accuracy of estimates of
concessionary travel, given the recent introduction of free bus travel for all aged over
60, and hence the need to accurately estimate compensation payable to operators. Not

only can total trips be recorded, but also the proportion of cards that have been issued which are 'active' (i.e. used at least once within a given period). Hence, more reliable estimates may be made of trip rates per holder, and effects of policy changes (such as the planned extension from free travel within the district of residence to the whole of England) may be monitored consistently.

While operators have sometimes found it difficult to justify investment in smartcard technology purely on commercial grounds, systems may be introduced largely at public expense in cases such as concessionary travel, the technology then becoming available for a wider range of ticket types.

In contrast to travel diaries, for which a period of one week is about the feasible maximum (as found in the NTS), much longer periods may be analysed, of up to seven weeks or longer, depending on the length of time for which data is retained. This enables variations in individual travel patterns to be examined (such as the extent to which users consistently travel at the same times during the peak, or vary from day to day).

There a number of operators with a joint ticketing systems, such as travelcards. Smartcard data enable more accurate apportionment of trips and revenue, improving the confidence that partner operators may have in this process, and also satisfying the requirements of regulatory authorities (such as the Office of Fair Trading in Britain) that a fair method of revenue apportionment is being used.

Inferring linked trips and other features

An advantage of recording trips made on the same card in the same day is that it becomes possible to examine the successive linking of trips during that period, as well as estimating overall trip rates. Strictly speaking, what many operators record is 'boardings', rather than trips, e.g. on bus systems in Britain a separate fare is paid (or a card validated) each time a passenger boards a bus. However, as well as making journeys between one activity and another (such as home and work) in which only one bus ride is involved, users may also travel on more than one bus, interchanging en route. This is particularly evident in large cities where the complexity of the network means that through journeys are not always possible. In addition, users may derive time savings by boarding the first bus to arrive and transferring to another service rather than waiting for an infrequent through route. By analysing successive boardings on the same card, reasonable inferences may be drawn as to the proportion of these 'linked trips' (for example, successive boardings of different routes within a 30 or 40 minute period), and hence estimates made of total trips as well as boardings.

This concept may be taken further by examining all trips on the same card in a day. This enables return legs of the same trip chain to be identified (for example, a journey in the morning peak, and return leg in the evening peak).

This can be taken somewhat further by inferring origin and destination (O&D). Where exit reading is practiced, then this may be derived directly from the smartcard data. Where only entry reading is followed, then inferences may be made – for example, that the alighting point for a journey made in the morning is the same as the boarding point in the evening). This may also be applied to earlier technology than smartcards,

where similar data is collected. A number of researchers have examined this approach[1] . However, it is not readily applicable to more complex trip chains.

The fine degree of detail possible in recording times at which trips commence is of particular value if assessing scope for fare structures which vary by time of day, from simple peak/off-peak differentiation to more complex structures, and monitoring the effects of introducing or modifying them.

Some limitations, and scope for complementary survey methods

One must, however, bear in mind some limitations of smartcard data. It does not record ultimate origin or destination (e.g. at the level of street address), and in some cases a substantial walk link may be involved, especially for urban rail use. For this purpose, manual survey data collection will continue to be necessary when operators are planning networks at the detailed level, as distinct from monitoring levels of use. No record of trip purpose is obtained, nor any attitudinal data collected.

A limitation in some cases is the continued use of cash fare payment for some journeys (albeit now very low in cases such as buses in London), and use of ticket types other than smartcards (for example, multirider tickets cancelled manually). In some cases certain modes may have comprehensive coverage of smartcard readers (such as buses and Underground in London), whereas other modes or operators are not covered (notably the privatised train operating companies in London, although agreement has now been reached to extend the Oyster system to them).

Hence records obtained through smartcards give a detailed picture of part of the public transport market, but not a fully-comprehensive one.

These findings point to a complementary role for smartcard and annual data collection. Smartcards should enable a good deal of routine data collection to be automated, such as estimating trip rates by concessionary pass holders. At the same time, certain data types (such as exact O&D, or trip purpose) require manual data collection methods to be used. One can also see complementary roles, in which manual data is used to check or calibrate that from smartcard systems. For example, a sample of smartcard holders could be interviewed to establish their trip purposes by time of day. This data could then be used in conjunction with automatically-collected smartcard data to provide a profile of trips by purpose. Likewise, a direct user survey could be carried out to establish patterns of interchange, and typical intervals between commencing one leg of a trip and the following leg, as a means of determining a time interval to be used in inferring linked trips.

This paper is based on work conducted at the University of Westminster, primarily in the form of a PhD thesis by Mousumi Bagchi for which Peter White was the principal supervisor. A fuller description of work examining two pioneering systems in the north of England and the data they produced has also been published [2],

Footnotes

[1] See, for example, Trepanier, M. and Chapleau,R. (Ecole Polytechnique de Montreal) 'Destination Estimation from Public Transport Smartcard Data; Paper at 'INCOM06' conference, St Etienne, France, 2006 ; and Hofmann,M and O'Mahony, M. (Trinity College Dublin) 'Public Transport Passenger Path and Behaviours : Extraction of Origin/Destination information' Geo Informatics [date/issue to insert]

[2] Bagchi.M and White,P.R. 'The potential of public transport smartcard data' Transport Policy, Vo1 12 (2005) pp 464-474

PRW/PB 18.05.07   [1590 words]

'ETsmcdPaperMay07.doc'/d:/p4 - end