WORKING PAPER
ITS-WP-04-25

CAITR 2004

Mobile Phones as Traffic Probes

by

Geoff Rose

8-10 December 2004

ISSN 1440-3501

INSTITUTE OF TRANSPORT STUDIES

The Australian Key Centre
in Transport Management

The University of Sydney
and Monash University

Established under the Australian Research Council's Key Centre Program

NUMBER: Working Paper ITS-WP-04-25

TITLE: Mobile Phones as Traffic Probes

ABSTRACT: The provision of road-based travel time information often relies on speed data collected from inductive loops imbedded in the pavement. While this instrumentation is common on urban freeways, the loops installed on arterial roads are not configured to provide speed data. Road authorities desiring to disseminate dynamic, network-wide travel information to road users are therefore considering a range of data collection techniques. The use of mobile phones as traffic probes is appealing because the necessary infrastructure is already in place in most urban areas. Traffic speed information can be obtained by passively monitoring data transmission in the mobile phone network. International experience provides encouraging signs about the potential of mobile phones as traffic probes. Issues still to be resolved include potential public concerns about privacy, growing awareness of the road safety of implications of mobile phone use, and the need to better understand the quality of the data obtained from the mobile phone probes.

KEY WORDS: mobile phones, dynamic travel time information, traffic probes, real time traffic data

AUTHOR: Associate Professor Geoff Rose

CONTACT: Institute of Transport Studies (Monash)
The Australian Key Centre in Transport Management
Department of Civil Engineering
Building Box 60
Monash University Victoria 3800 Australia

Telephone: +61 3 9905 9627
Facsimile: +61 3 9905 4944
E-mail: itsinfo@eng.monash.edu.au
Internet: civil.eng.monash.edu.au/its/

DATE: 8-10 December 2004

1. INTRODUCTION

Travel time and traffic congestion information are valued by road users and road system managers. Road users rely on the information for journey planning and route choice decisions while road system managers are increasingly viewing travel time as an important network performance indicator. Public and private sector organisations are now involved in the dissemination of real-time traffic information to road users in real time via radio, roadside displays, internet sites and in-vehicle devices.

Travel time information systems rely fundamentally on the availability of real-time data on current traffic conditions. There are two alternatives for collecting that data, either by deploying fixed sensors (e.g. loops or roadside tag readers) or mobile sensors, often termed probe vehicles. This paper examines the emerging opportunities to collect network wide traffic information using mobile phones as traffic probes. The rationale for these systems is that with widespread access to mobile phones, many vehicles contain a mobile phone. By measuring the speed of mobile phones, the associated traffic speed can then be determined.

The aims of this paper are to:

- Examine qualitatively, the relative pros and cons of mobile phones as a source of dynamic traffic information compared to alternatives,
- Outline the approaches available for collecting traffic information from mobile phones and the type of information which can be obtained,
- Identify the state of practice in the use of mobile phones as traffic probes, and
- Identify unresolved issues which may have implications for the prospects of obtaining real time traffic information using mobile phones as probes.

The structure of the paper is as follows. The following section (Section 2) discusses the relative pros and cons of alternative sources of dynamic traffic information. Section 3 then describes how mobile phones can be passively monitored to provide traffic speed information. A number of issues are then identified which need to be addressed as part of the ongoing development of this technology (Section 5). The conclusions of the paper are presented in Section 5.

2. COLLECTION OF REAL-TIME TRAFFIC INFORMATION

The most common source of data for operational dynamic traffic information systems, such as Vic Roads DriveTime system (See Figure 1), is inductive loops. These point detectors are commonly installed on average about every 500 metres along freeways. They provide average traffic speed data every 20 to 60 seconds depending on the installation. As noted by Fontaine and Smith (2004), while inductive loops are a proven technology they have significant infrastructure and maintenance requirements and tend to be installed on only the most heavily trafficked roads in urban areas, specifically on freeways. Importantly these loops are usually installed to provide data for automatic incident detection systems but the data then also becomes available for other applications such as dynamic traffic information. The costs associated with installing the loops and the communication network needed to relay the data back to the control centre have also tended to confine this instrumentation to the inner parts of metropolitan areas. The recently upgraded Princess Freeway (a six lane facility) from Melbourne to Geelong, for example, contains no loop detectors over the approximately 26 kilometre stretch west of Werribee.

Figure 1: Roadside display in Vic Roads Drive Time System

The arterial network is rarely instrumented in a way that facilitates collection of data from which real-time travel time information can be disseminated. While traffic signal systems rely on inductive loops for input data, those loops are usually located at the stop line and provide degree of saturation rather than speed data. Current research (Luk, 2004) is, however, examining the potential to estimate travel times using the data collected by an adaptive traffic signal system based on earlier encouraging results (Karl and Whitmore, 2002).

Probe vehicles present the other alternative for gathering real time information. These systems provide the potential to increase system coverage without the infrastructure requirements of point detectors (Fontaine and Smith, 2004). These probe-based systems rely on monitoring a sample of the traffic stream and using the data from those vehicles as an estimate for the mean speed of traffic on the road. Probe systems take two general forms (Fontaine and Smith, 2004); first those where the probe vehicles are sampled at fixed locations and secondly systems where the vehicles are randomly sampled on the road network. Data provided by automatic vehicle identification systems, commonly installed on electronic tollways, are an example of the first type of probe system which can provide a basis for travel time estimation (Wang and Nakamura, 2004). An alternative, developed in Australia as an additional feature of the adaptive traffic signal control system SCATS, is to fit vehicles with an electronic tag and establish their location at points where tag readers are installed (Longfoot, 1991). Without needing a transponder fitted to the vehicle, TrafficMaster in the UK pioneered the collection of licence plate details using cameras installed every six kilometres on the motorway network in Britain (MacMorran and Billington, 1999). The alternative is to fit probe vehicles with a Global Positioning System (GPS) receiver and communicate the time stamped location information provided by that system to a traffic control centre. That data provides a basis for calculating travel times. ITIS Holdings have pioneered the use of this approach in the UK (ITIS Holdings, 2004). They targeted vehicles which are heavy uses of the road network to act as probes, e.g. AA vehicles, buses and trucks, since these provide a lot of data per vehicle.

The above probe vehicle options all have costs associated with either fitting or reading the electronic tags, installing the necessary field equipment to read the tags or number plates, or communication costs associated with relaying the GPS data to the control centre.

The use of mobile phones is potentially appealing particularly where it can rely on the hardware that is already installed to make the mobile phone system operational. This has intuitive appeal because on the surface it would suggest that the data could be made available in a more cost effective way without the need to install expensive and extensive hardware in the field. It also provides the potential for coverage over a wide area reflecting the ubiquity of the mobile phone network particularly in urban areas.

3. MOBILE PHONES AS TRAFFIC PROBES

There are two basic technology options for mobile phone traffic probes: handset-based solutions and network-based solutions. The handset-based solutions rely on additional technology in the handset, for example, a GPS receiver. The location data can then be communicated over the GSM (Global System for Mobiles) network using SMS (Short Messaging Service). The alternative, network based approach, is to add 'intelligence' to the network and make use of data transfer which is already taking place within the mobile phone system. Approaches which rely on additional capabilities in the handset have the potential for greater accuracy in the location fix on the phone but the trade-off is additional communications costs since the phone must then 'report' its location. The systems reviewed here are all network-based solutions. These have the advantage that they do not require any changes to handsets and are therefore able to deliver larger probe vehicle fleets because of the relatively slow uptake of GPS receivers in mobile phones. The network-based solutions considered here are also termed passive systems in that they do not initiative any additional data transmission within the mobile phone system. In contrast, active systems, which require mobile phones to be polled to establish locations, have the potential to use up bandwidth and also add to the operating costs of the system.

Importantly, all of the systems considered here require the mobile phone to be switched on to act as a probe. It is not necessary for the phone to be in use (for voice, data or video transmission) to act as a probe. Even when not in use, mobile phones periodically report to the network so their whereabouts is known in the event of an in-coming call.

The mobile phone system comprises handsets and base stations. The handsets communicate with the base stations. There is active communication between the phone and the nearest base station when it is in use and each phone is associated with its closest base station. Given a regular spacing of base stations, the boundaries of the areas covered by each base station are hexagonally shaped and are termed a cell (See Figure 1). The cellular nature of the network is the reason that the term 'cellular phone', or 'cell phone', for short is commonly used in the USA while in many other countries they are referred to as mobile phones. A process of cell handover occurs when phones move outside the cell boundary associated with one base station. For communication to continue if the phone is in use, or for the system to update the cellular location of the phone, the phone is 'handed over' to the next closest base station.

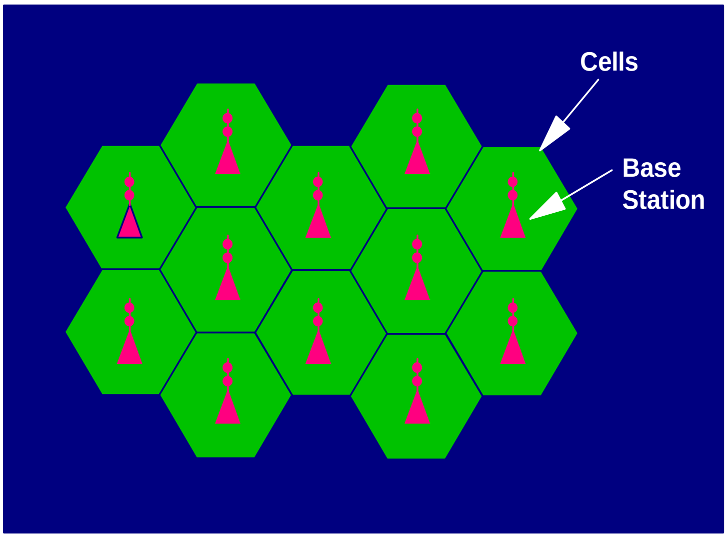


Figure 1: Nature of the mobile or cellular phone network

Due to commercial confidentiality reasons, alternative approaches for inferring traffic speed data from mobile phones, do not tend to be thoroughly described in the literature. Here we review approaches that rely on passive monitoring of the mobile phone to locate it and estimate its speed.

At least conceptually, the location of a particular handset could be determined by trilateration using signal strength or transmission delay from multiple base stations (Figure 2). Under this approach, the location of the phone could be established anywhere in a cell since the process of triangulation is not restricted to any part of the cell. Conceptually the easiest way to imagine speed data being obtained from a mobile phone is to locate it at two points in time and calculate the speed on the basis of the distance travelled divided by the time taken to travel that distance.

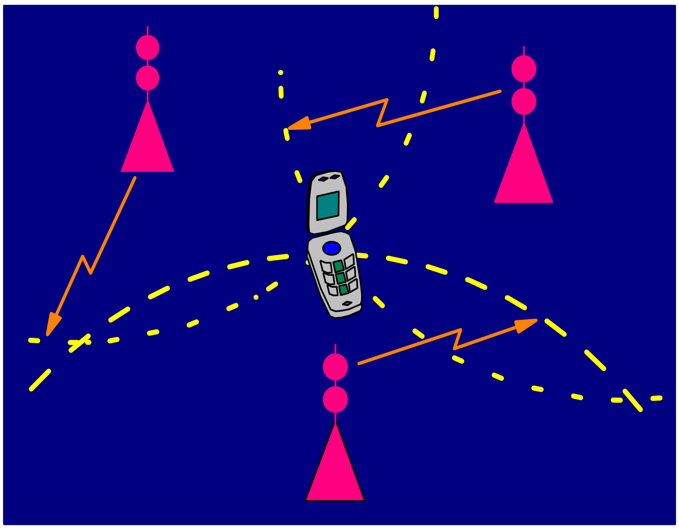


Figure 2: Establishing handset location by trilateration

Much of the interest in mobile phones as traffic probes stems from legislation enacted in the USA which requires mobile phone operators to be able to locate a mobile phone in the event of an emergency as is done for land line services. Mobile operators have the option of establishing the location of the phone using either network-based or handset-based tracking solutions (Yim, 2003). The handset-based approach could rely on having a GPS receiver in the phone so that precise location details can be transmitted automatically. While such an approach has the potential to produce precise location data it relies on technology which is not available in every handset. In contrast, the network-based approaches do not require any additional technology to be available in the handset but use the data from the mobile phone system to obtain a location fix on the phone. This approach has the advantage of potentially a much larger sample size from which the speed of phones could be determined. However, the regulations do not require operators to establish the location of the phone with pin point accuracy. The E911 standards set by the US Federal Communications Commission in 1996 (Yim, 2003) specify that network-based solutions should be able to identify 67 per cent of calls within 100 metres and 95 per cent of calls within 300 metres. Speed estimates, which rely on getting a fix on the phone at two points in time, are therefore impacted by those location errors.

An alternative approach is to monitor cell handover zones. This approach relies on the fact that the mobile phones which are travelling on a particular route change base stations at a fairly fixed location (Kummala, 2002). As the mobile phone moves between cells, the mobile phone network undertakes a 'handover' operation so that the phone is continuously associated with its closest base station. The cell handover zones provide scope for obtaining an approximate fix on a phone and then speed can be established using another fix at the next cell handover zone (See Figure 3).

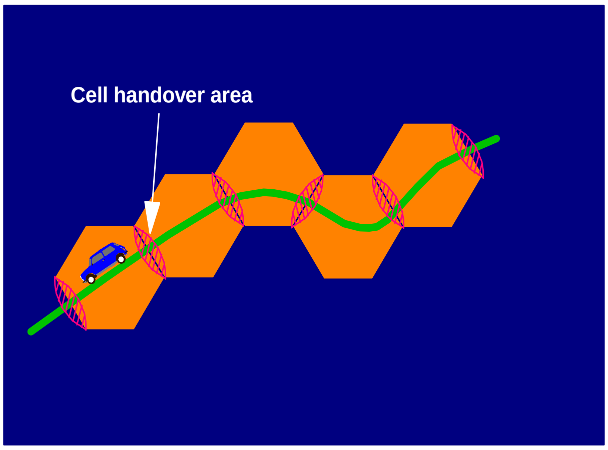
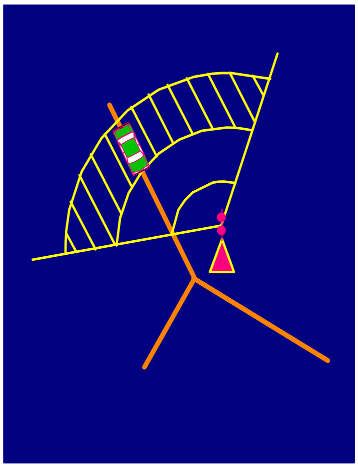


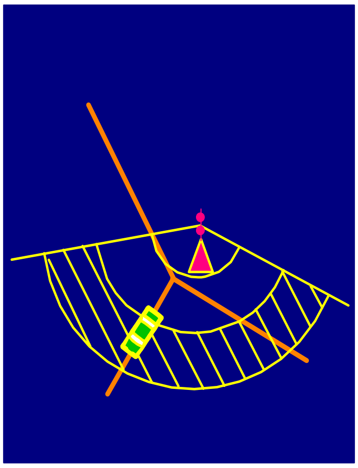
Figure 3: Vehicle traversing cell handover zones

When a phone is being used for communication, and is moving, it will be handed over between cells as needed. Therefore, a phone which is in use provides fairly continuous probe vehicle data. Even if the phone is not in use, but is at least switched on, it will periodically report its location to the network. It is possible to use those opportunities to obtain data which can provide a basis for speed estimation. In that case, the probe will not supply the same quantity of information as if it was in continuous communication.

Another approach relies on a feature of the GSM network known as Timing Advance (TA) through which the network compensates for the distance from handset to the base station (Applied Generics, 2004). The timing advance information can be used to narrow a handset's location to a 550 m wide concentric band radiating from the base station (Figure 4a). Over time, it is also possible to use data received from different quadrants monitored by a particular base station to trace the route of a particular vehicle through the road network (Figure 4b) to infer traffic speed. Under this approach the phone must be in use to act effectively as a probe since it is the analysis of the signal which provides the basis for speed and location determination. The brief burst of data communicated during cell handover is not adequate to establish the location or speed under this approach.



(a) Monitoring the Timing Advance, in combination with map matching, provides an approximate location fix



(b) Over time the Timing Advance provides a basis for vehicle tracking and speed estimation

Figure 4: Monitoring the Timing Advance provides a basis for traffic speed estimation

4. ISSUES STILL TO BE RESOLVED

While mobile phones appear to offer appealing prospects as traffic probes there are still a number of issues which need to be resolved.

Of particular concern is the accuracy and reliability of the data. Accuracy refers to the level of error on average while reliability is usually interpreted to mean the variability about the average (Richardson et al, 1995). Another issue is the 'timeliness' of the information. Inductive loops provide data every 20 secs and can therefore provide a foundation for the development of real time travel time information systems (Hearn, 1995; Paterson, Rose and Bean, 1999). It is not clear what time lag is associated with the data obtained from mobile phone systems and whether they can support real time applications such as incident detection, on the basis of changes in travel time, or near real time applications such as provision of travel time information to road users.

Clearly one of the shortcomings of the mobile phone systems which rely on locating the mobile phone at two points in time is that location errors in the position of the phone directly impact on the speed estimates that are obtained. Fontaine and Smith (2004) report an innovative simulation approach to examine the impact of those errors in the context of a system where the mobile phone system is used to obtain a location fix on each phone anywhere within the study region. Fontaine and Smith used a microscopic traffic simulation package to examine the impact of location errors in the mobile phones. The analysis began by using the traffic simulation model to record the 'true' location of each sampled vehicle. The accuracy of that location fix was then degraded by a known error level to produce the location as determined by the mobile phone system. Those locations were then map matched back on to the road network prior to estimating vehicle speeds. Using a rigorous experimental design, Fontaine and Smith examined the impact of a range of factors on the accuracy of the speed estimates. They found that different map matching algorithms, the frequency between position estimates, the level of error in the position estimates and road network geometry all had a significant impact on the accuracy of the speed estimates.

Importantly, Fontaine and Smith's (2004) finding about the impact of road network geometry on the accuracy of the mobile phone probe speeds is consistent with the results obtained in a major European field trial. Ygnace et al (2001) report results from a field trail which focussed on one section of urban motorway and another section of interurban motorway near Lyon in the south of France. They used speed data obtained from inductive loops as the 'ground truth' against which the speed estimates from the mobile phone system could be compared. The results demonstrated a correlation between the two speed estimates however the results were much poorer for the section of urban motorway. This was attributed to the nature of the network and the risk that phones which were off that part of motorway would be erroneously 'located' on the motorway. While data screening could eliminate some obvious outliers, for example, someone using their mobile phone while walking on a surface street near the motorway, there was still a high enough level of error in the location data to degrade the quality of the speed data. While the results show promise for the potential to obtain speed data from mobile phones it is also appropriate to highlight that they demonstrate that further research is also needed. In the final analysis average speeds from the mobile phone system and the inductive loops are reported as hourly averages for each hour across the day. The data aggregation inherent in reporting hourly averages removes a great deal of variation in the data. While the results indicate a correlation in the speeds obtained from the mobile phones and loops, there would clearly be a need for refinement of the mobile phone positioning system to obtain speed estimates which could be used in a real-time or near real-time applications.

Similar conclusions appear to have arisen in a later trial conducted in Finland where again the level of error in the position fixes on the phones undermines the quality of the resulting speed data. The approach taken in that field trial focussed on monitoring the cell handover areas. The results imply that further research would be needed to achieve a operational system for the delivery of reliable, real-time speed data (Kummala, 2002). The Finnish results were impacted by 'outlier' observations obtained from parallel roads, public transport, bicycles or pedestrians. Importantly, Kummala concludes that the location of the mobile phone base stations is not always optimal for traffic monitoring purposes and that the system was likely to be best suited for traffic monitoring on long observational links (stated as approximately 10 km) on which it is rare to find vehicles entering or leaving the link or stopping on it. This suggests this approach is more suited to an inter-urban motorway context rather than an urban setting. However, Kummala (2004) notes that an urban setting increases the number of potential observations due to the higher density of mobile phones but also increases the risk of falsely attributing observations to the road which is being monitored.

Yim (2003) reports the status of cellular probe research with a particular emphasis on developments in the USA. She notes that four of the six largest US cellular carriers have adopted handset-based solutions to meet the requirements of the E911 legislation. However, there has been little interest from the carriers part in collecting and processing travel time for commercial users. Interestingly that recent review from the USA reports no results from US field trials and relies on the results from the

Rhone Corridor study (Ygnace et al, 2001) to demonstrate the potential of network-based solutions. Yim argues that for further progress in the USA, critical issues are overcoming concerns about privacy, and developing business models to address the return on investment which carriers could obtain on research into cellular probes. On the basis of Yim's review, one would conclude that this area remains very much at the research frontier in the USA with no immediate prospects of operational systems.

Research undertaken by a small company in Israel led to the Estimotion technology which has since been purchased by ITIS Holdings in the UK (ITIS Holdings, 2004). This has been tested in Israel in the area round Haifa and is now being applied in the UK (Quayle, 2004). In parallel the Dutch information technology company, LogicaCMG has developed a system called 'Mobile Traffic Service' (MTS) (LogicaCMG, 2004a) which has been field trailed in part of a province in the south of the Netherlands. An area of approximately 800 km^2 was used for the test with that part of the province containing two towns with populations of about 200,000 inhabitants. A validation study (LogicaCMG, 2004b) compares the data obtained from MTS to data obtained from three reference systems:

- A floating car trave time survey
- A travel time survey conducted using a timed number plate survey, and
- Travel times estimated from speed data provided by inductive loop detectors installed on the motorways in the test network.

The results highlight the high correlation between the travel times from MTS and those obtained from the reference systems. This represents some of the most convincing evidence of the prospect of these systems to provide network-wide travel time information.

In addition to the level of error on average (the accuracy) the reliability (or variability about the average) of the data from mobile phone probes is of interest. Since the mobile phone probes are essentially providing a sample of speed observations on the network there are issues of the representativeness of that sample and the size of the sample. From the perspective of representativeness, there is an obvious difference between having four phones sampled from one vehicle as opposed to having four phones samples from four different vehicles. Ultimately, the sample size will depend on:

- Which mobile phone operator's network is being monitored – reflecting the market penetration of the mobile phone operator that sets an upper limit on the number of probes,
- Whether the phones are switched on – phones which are switched off are not able to act as probes, and
- Whether the phones are being used for communication – if the system relies on continuous monitoring of the signal from phones which are in use.

A recent intercept survey of drivers in Melbourne Vic Roads (2003) suggests that about 61 per cent of total travel on the road network in Melbourne (in terms of vehicle kilometres) is undertaken with a mobile phone switched on in the vehicle. This suggests a potentially large pool of mobile phone probes. Importantly there are differences in the availability of a mobile phone depending on gender, vehicle type and drivers age (Vic Roads, 2003) which could all have implications for the representativeness of the sample. Sample size, as noted above, will also depend on which mobile phone carrier supplies data for the system. Recent mobile carrier market share data (DCITA, 2004) highlights that the largest carrier has slightly over 40 per cent of the subscribers while smaller carriers have less than 20 per cent of the subscribers.

Other factors could also impact on the sample size of mobile phone probes. There is on-going public debate about the safety of driving and using a mobile phone with some research suggesting that the risk is no less if the mobile phone is installed for hands free operation. One potential risk is that either legislation could be introduced to restrict mobile phone use in cars (which is judged to be unlikely) or organisations could act to impose bans on mobile phone use to reduce their exposure under Occupational Health and Safety Legislation. One current example is the recently introduced policy of BP Australia which specifies that mobile phones are to be switched off then in a vehicle (implemented under a ‘Car On, Phone Off’ policy) (Anthony, 2004). Widespread adoption of that form of operating policy could have an impact on the number of vehicles available as traffic probes.

Another issue which could impact the sample size is concerns of individuals about privacy. At present a ‘firewall’ approach is adopted by some of the systems under which information which would enable the identification of the owner of the SIM card in the mobile phone is not required by the traffic probe application. However to respond to ‘perceived’ concerns about invasion of privacy, mobile phone operators may have to adopt either an ‘opt in’ or an ‘opt out’ policy whereby users would have to either explicitly agree to their phone serving as a traffic probe, or provide the option for people to ‘opt out’ of having their mobile phone monitored. While the magnitude of either is open for debate, they would both work to reduce the sample size of potential probes.

5. CONCLUSIONS

Mobile phones offer some appealing characteristics as traffic probes. Network-based solutions, which rely on passive monitoring of data already being communicated in the mobile phone system, have the potential to provide network wide travel time information. The existence of two operational systems suggests that this technology is maturing and moving beyond the pure research stage.

There are clearly a number of issues which still need to be resolved in relation to this technology. There is scope for greater understanding of the accuracy, reliability and timeliness of this data as well as further examination of the potential impacts on the size of the probe vehicle fleet of either real or imagined privacy and safety issues.

While mobile phones may be capable of delivering dynamic travel time information in their own right, the future may lie in fusing the data from these probes with travel time estimates developed on the basis of data from the traffic control system or other probe technologies. In that way a dynamic travel time information system may be developed which is greater than the sum of its parts.

REFERENCES

Anthony, E. (2004) Personal communication. BP Australia.

Applied Generics (2004) *RoDIN24 road traffic monitor*. Accessed from http://www.appliedgenerics.com/ (Accessed 26/7/04)

Department of Communications, Information Technology and the Arts (2004) “DATA Issue 7: Mobile Phone Subscribers”. Accessed from http://www.dcita.gov.au/ict/publications/data_magazine/issue_7/data_issue_7_-_mobile_phone_subscribers . Last updated: 19/11/04 (Accessed 2/12/04)

Fontaine and Smith, B.L. (2004) “Probe-based traffic monitoring system using wireless location technology: investigation of the relationship between system design and effectiveness”. Paper submitted for presentation at the 84th Annual Meeting of the Transportation Research Board.

Hearn, B. (1995), "A dynamic freeway information system for Melbourne", *Proceedings* of the International Conference on Application of New Technology in Transport System, Melbourne, Australia, 287-306.

ITIS Holdings (2004) “Estimotion“. Accessed from http://www.itisholdings.com/estimotion.asp (Accessed 26/7/04).

Karl, C.A. and Whitmore, M. (2002) Beat the Jam: Traffic Services for Shell Mobile in Hong Kong. *Proc*. Institution of Transportation Engineers 2002 International Conference, Melbourne, Australia. Institution of Transportation Engineers, Australia/New Zealand Section.

Kummala, J. (2002) *Travel time service utilising mobile phones*. Finish Road Administration, Helsinki. Finnra Report 55/2002. 67pp.

Longfoot, J., 1991, An Automated Network Travel Time System - ANTTS, *Proc*. of Vehicle Navigation and Information Systems Conference, Vol. 2, Warrendale, P.A., Society of Automotive Engineers, pp. 1053-1061.

LogicaCMG (2004a) “Mobile traffic services””. Accessed from http://www.logicacmg.com/transport/mts.asp (Accessed 26/7/04)

LogicaCMG (2004b) “MTS quality assessment results”. Accessed from www.mts-live.com (Accessed 26/7/04).

Luk, J. (2004) Personal Communication. ARRB Transport Research.

MacMorran, T. and Billington, , P. (1999) ‘Putting journey time on-line: Image capture for traffic monitoring’. *Traffic Technology International*. Annual Review 1999, 139-142.

Paterson, D., Rose, G. and Bean, S. (1999) “Improving dynamic travel time estimates for Melbourne’s Drive Time System”. CD Rom proceedings of the Intelligent Transport Systems Australia Conference, Adelaide, 10 pp.

Quayle, D. (2004) Personal Communication.

Richardson, A.J., Ampt, E.S. and Meyburg, A.H. (1995) *Survey Methods in Transport Planning*. Eucalyptus Press, Melbourne, Australia. 459 pp.

Vic Roads (2003) *Driving Around Melbourne.* Prepared under contract by Drummond Research Pty. Ltd. For the Vic Roads Road Safety Department, Kew, Victoria.

Wang, R. and Nakamura, H. (2004) Expressway travel time estimation by using the combination of detector and probe vehicle data. *Proc*. 10th World Congress on ITS. ITS America.

Ygnace,J-L, et al. (2001) Travel Time/Speed Estimates on the French Rhone Corridor Network Using Cellular Phones as Probes. Final report of the SERTI V program, INRETS, Lyon, France. 148pp.

Yim, Y. (2003) *The State of Cellular Probes*. California PATH Research Report. UCB-ITS-PRR-2003-25. 29pp.