THE USE OF CELLPHONE TECHNOLOGY IN ACTIVITY AND TRAVEL DATA COLLECTION

S C KRYGSMAN and P SCHMITZ

CSIR Transportek, PO. Box 320, Stellenbosch, 7599.
Tel: 021 888 2547. Fax: 021 888 2694. E-mail: skrygsman@csir.co.za

ABSTRACT

The potential use of cellphone technology to address the critical data needs of developing countries particularly in terms of activity and travel data collection is explored in this paper. Transport planners and decision-makers increasingly rely on detailed information of travel at an individual level to inform transport planning and policy formulation. Origin-destination (OD) surveys rely on (selected) households being visited to collect information on place of residence, place of employment, transport mode used, time of departure/arrival, etc. (mainly related to the main commute trip). In a developing country, activity-travel surveys can be particularly difficult to administer and complete due to lower levels of literacy, language barriers and limited funding.

Global Positional Systems (GPS), Global System for Mobile Communication (GSM), Location-Based Services (LBS), computer assisted data collection technologies, satellite imagery and computing technology offers interesting and potentially very useful tools to support and complement traditional activity and travel data collection. The proliferation and adoption of cellphones (and other consumer devices such as GSM technology and A-GSM) offers the opportunity to leverage existing hardware and software already in the hands of study participants for collecting detailed travel behaviour information.

This paper reports on a *proof-of-concept* study to track individuals over a two-day period making use of their cellphones. A small sample (n = 83) was recruited from the CSIR site (Pretoria) for the study and their cellphones were tracked over a two day period. For each individual, a temporal and spatial location was determined every five minutes by sending a "signal" to the cellphone. The signal recorded the time (z), and the individual's geographic position namely latitude (x) and longitude (y). Thus, the temporal and spatial location of each individual was recorded *570 times* during the 48 hours.

The tracking yielded a comprehensive record of each individual's spatial and temporal movements. Various rules and algorithms were applied to transform and label the data to make it more suitable for travel analysis. The 'interpreted' data were subsequently analysed for locational (and temporal) accuracy, measurement errors, and in terms of how well it provided for an accurate record of an individual's daily activity path. Reasons for any errors have been explored and possible solutions to these have been put forward.

This proof-of-concept study revealed that it is indeed possible to obtain detailed individual travel behaviour data from cellphones. While the results are potentially very useful for transport planning, retail analysis, etc., the research did reveal the need for significant data manipulation to eliminate measurement errors and derive useful transport planning data. The paper concludes with recommendations for future research.

1. INTRODUCTION

Transport decision-makers and transport planners increasingly rely on detailed information of individual-level travel and activity patterns in disaggregate form for transport planning decisions and policy formulation. Historically, most of this information was collected with origin-destination (OD) surveys which required that (selected) households be visited and information be collected on place of residence, place of employment, transport mode used, time of departure/arrival, etc. (the information related mainly to the main commute trip). Such surveys are often expensive, contain relatively little information about the complete daily travel pattern and can be of dubious quality if the exercise is not planned and executed well. Furthermore, as the shortcomings of the conventional aggregate transport planning approach received more attention, theoretical, conceptual and methodological advancements in activity/travel behaviour research have necessitated the need for more detailed, disaggregated activity and travel data (Bhat and Koppelman, 2000; Kitamura, 1997; Pas, 1990; Pas, 1995; Pipkin, 1986; Timmermans, 2000).

Location-aware technologies (LAT) offer interesting and potentially very useful techniques and tools to augment, complement and in some instances even replace traditional OD-surveys and activity-travel diaries. These technologies include Global Positional Systems (GPS), Global System for Mobile Communication (GSM), radio-frequency identification (RFID) and are supported by computer assisted data collection technologies and tools, satellite imagery and mobile computing technology. This technology can overcome the problems associated with traditional data collection diaries and questionnaires such as respondent burden and incomprehension, erroneous recoding (e.g. wrong time and location) while providing supplementary information not previously available, such as the exact space-time activity/travel trajectory. In a developing country such as South Africa, cellphone technology can be particularly useful to overcome the know problems associated with trip under-reporting, missing and or incomplete data (in specific locational data) and finally incomplete, missing, or inconsistent trip details.

GPS in particular have increasingly been used by vehicle tracking companies, logistics operators and wildlife researchers to track vehicles, monitor freight, and track animal movement in wildlife reserves over prolonged periods. While the process and technology of such 'tracking' is relatively well established for vehicles, freight pallets and individual animals, the technology and methodology for a person-based service is less well understood and established (Asakura and Hato, 2005). Furthermore, the cost of GPS excludes the technology from being used in large scale surveys of human activities and travel behaviour of commuters. The proliferation of cellphones and their inherent locational positioning and temporal features provides for an interesting technology to address some of the data and technology concerns.

This paper reports on a 'proof-of-concept' study which tracked individuals over a two day period making use of their personal cellphones. The objectives of the study were threefold, i.e. to assess the technical feasibility of obtaining spatial *and* temporal data from the cellphone tracking, to assess the consistency of the data (namely the provision of a full record of all activities and travel over an extended period of time) and, finally, to assess the quality of the data and the potential of integrating the data with auxiliary data sources (such as land use, transport networks and other facilities, etc.).

In this paper, a very brief literature review is provided summarising the main developments and applications of new technologies, and cellphones in particular, in travel surveys. This is followed with the specification of an approach to operationalise the use of cellphone technology in developing countries to obtain travel data. Then, the case study is discussed

in some detail, as well as the tracking process and data obtained from the network operator. The analysis of the results is presented, with a discussion of the results and directions for future research concluding the paper.

2. BACKGROUND AND BRIEF LITERATURE REVIEW

Location-aware Information Communication Technologies (ICT) and more specifically cellphones provide a useful alternative (or rather a supplementary technological solution) to address some of the problems associated with comprehensive data collection requirements . These technologies may augment (or even replace) complex and cumbersome travel/activity diaries which are difficult to administer, complete, and which require an accurate recording of space-time trajectories. However, the uptake of these technologies by the transport planning fraternity has been slow, with only a few studies in Europe, Japan and the USA and none in the developing world being reported. The technology is of course a relatively new phenomenon and it is really only in the last ten years or so that the technology has become established (Stopher, 2004; Wermuth et al., 2004; Wolf, 2004).

In comparing the GPS and GSM in collecting spatial data, Kracht (2004) summarizes the accuracy and availability of GPS / GSM, technology as follow (Kracht, 2004):

Table 1. Comparison of GPS and GSM information.

Technology	Information accuracy			Information availability		
	Origin	Route	Destination	Origin	Route	Destination
GPS	+++	+++	+++	++	++	++
GSM	++	++	++	+++	+++	+++
Questionnaire	++	-	++	++	-	++
	+++ = good, ++ = sufficient, + = poor, - = no data					

While GPS does provide more accurate location information, information availability suffers from technological restrictions (i.e. cost, complexity, heavy systems, etc.). As illustrated above, however, these technologies seem to be very complementary. Kracht discusses the use of both technologies in a case study in tracking individuals using all modes of travel in Berlin, Germany. As the city is well connected via various public transport modes, many trips are undertaken by public transport. Where GPS services do not always accommodate person-based in-vehicle tracking GSM could provide an alternative.

The literature, while still sparse on the subject, does reveal that there are potentially many benefits to using cellphones to collect travel behaviour, notably in a reduction in missing trips reporting, improved accuracy in the reporting of trips and activities (time and space), better routing and speed data, and the ability to capture longer periods of travel (Sharp and Murakami, 2004). The technology also provides many benefits for respondents such as low input requirements and little (if any) respondent burden (such as trying to recall activities/trips, etc.). The potential benefits of mobile phones as compared to GPS was that (i) GPS is not always an unobstructed signal (i.e. in tunnels or inner cities), especially when people are using public transport, (ii) the market penetration of mobile phones is very high (with a very high potential), (iii) mobile phones are generally cheaper than GPS and do not require additional cumbersome hardware. Other disadvantages of GPS involve the cost of equipment and deployment, and the fact that data can only be downloaded

after the survey.

Cellphones, however, also have some inherent problems in data collection, notably with regard to pinpointing a location with accuracy. Complex positional algorithms supported by much additional data requirements may be required to overcome this. Accuracy is very dependent on the density of base stations and the quality of the signals in the study area. In urban areas with a higher density of base stations, resolution can be up to 50 – 100 meters. In outlying areas, the accuracy may be considerably less and in developed countries with fewer base stations, the positional data may be rather poor.

The research supported the potential usefulness of cellphones in obtaining detailed activity and travel data in developed countries. While the technology potentially holds much promise for use in developing countries, many technology and data quality issues still remain. (The proof-of-concept study was undertaken to address some of these issues.)

3. APPROACH FOLLOWED IN THE STUDY

Data collection for activity-analysis and transport planning is no small task, especially in a developing country. Figure 1 illustrates the role of GSM technology in overcoming the main problems associated with conventional OD surveys (limited data for more comprehensive needs based analysis for travel requirements) and complex activity/travel dairies (difficult to compete and administer) while retaining on the advantages of these tools. The passive tracking capability of mobiles phones allows, theoretically, for exact spatial and temporal record of daily activities and travel. By not requiring user input, it overcomes some of the problems associated with surveys in developing countries while also providing information not previously available, such as route choice and the sequence and duration of activities.
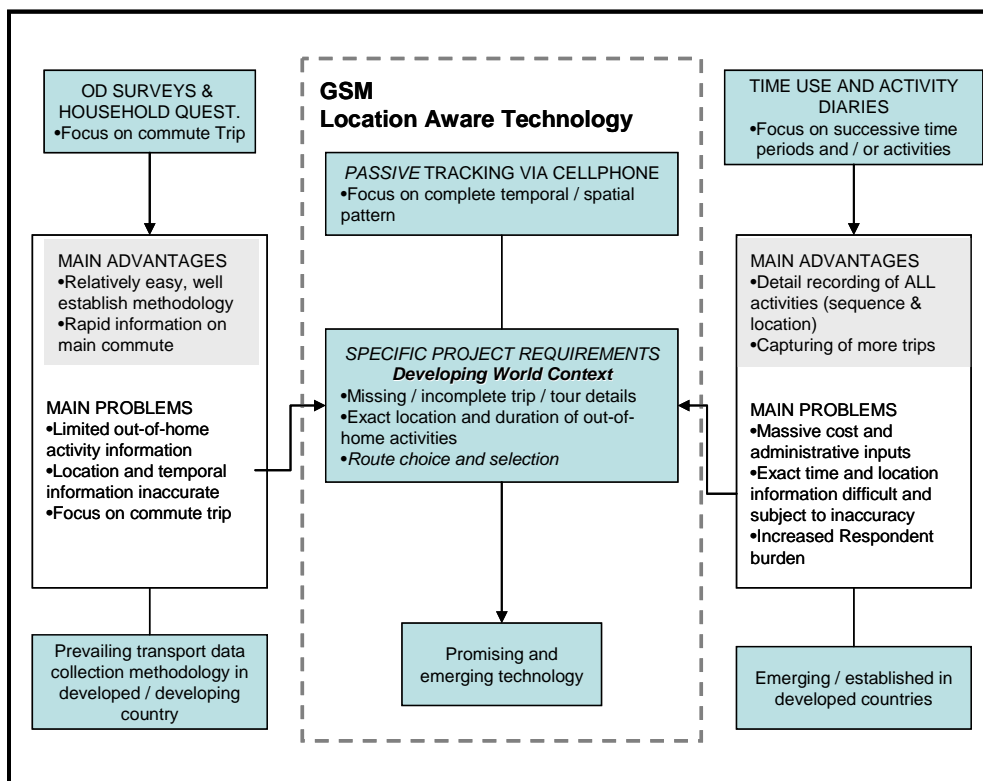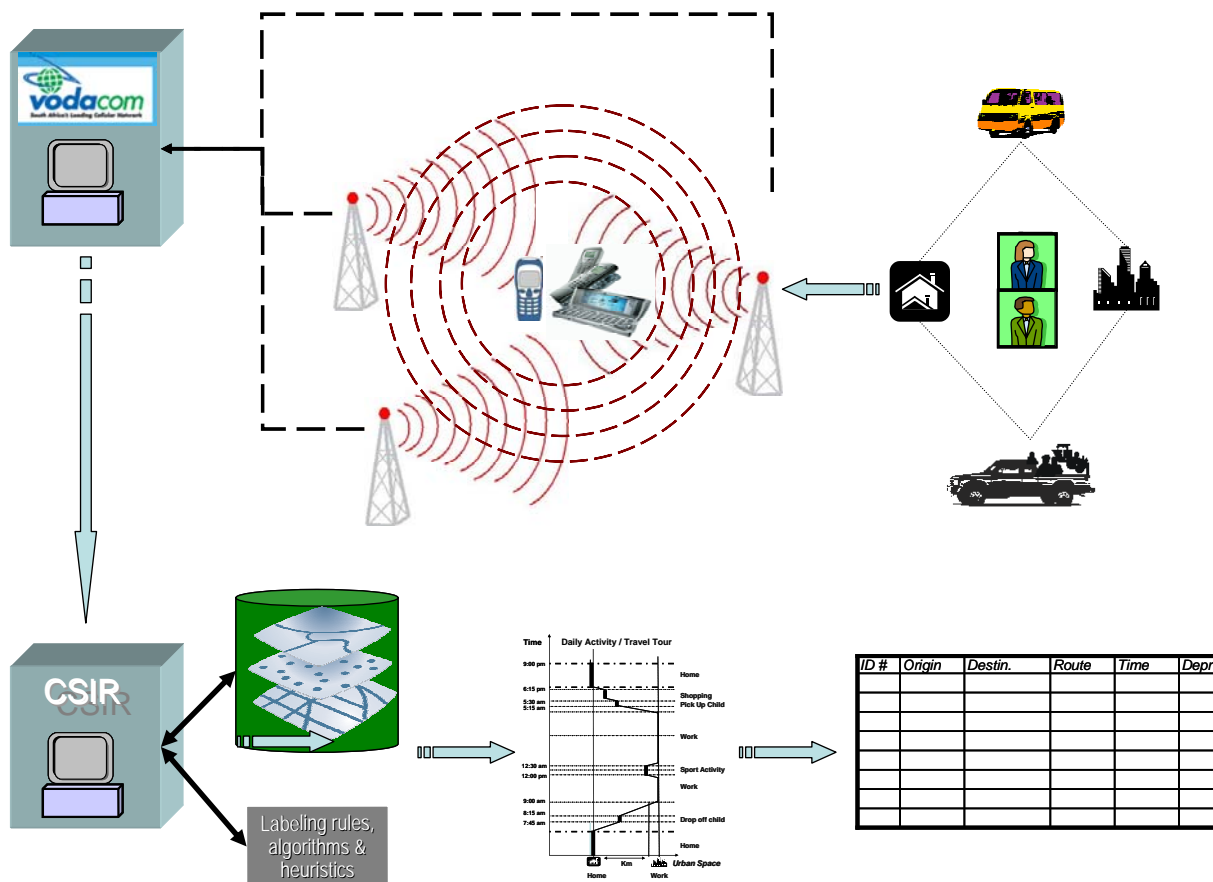

Figure 1. GSM technology: supplementing OD surveys and activity dairies.

In order to test the use of cellphone technology in a developing country, the stakeholder and data supply chain had to be mapped. Figure 2 shows the supply chain necessary to obtain and transfer individual positional information activity and transport analysis purposes.

Figure 2. Stakeholder and data supply chain.

Permission must obviously first be obtained from the study participants. Once obtained the cellphone number is forwarded to the service provider and tracking commences at an agreed time and date. The cellphone (whether being used to make a call or not) is in constant contact with a GSM network as long as it is switched on and the location information is accessed by sending the cellphone a signal. The service provider (or partner application provider) receives the information and adds geographic coordinates and time stamps. Positional accuracy (at a first level) is provided by the 'catchment' radius of the cellphone towers (which may be divided into segments). The information is subsequently transferred to a processing unit which adds contextual information to the data and 'interrogates" the raw data with labelling algorithms and cleaning rules. A daily activity-travel path is derived and data useful for applications, such as transport planning, is then extracted.

4. CASE STUDY AREA, SURVEY SAMPLE, LAND USE AND TRANSPORT SPATIAL DATA

As the objective of the project was primarily to consider the technological feasibleness of collecting data with the GSM network, the choice of study area was not guided by whether the area and sample population was (statistically) representative of the region or general population. Rather, the study area was chosen to minimise issues such as privacy concerns, legitimacy of the exercise, financial considerations and ease of administration.

The importance of these issues, however, should not be overlooked as previous researchers have pointed out (Wolf, 2004).

The sample was made up of 83 employees of the CSIR (all based at the Pretoria campus). The home location of these respondents were obtained from a questionnaire and mapped in GIS. It showed that the residential locations are spread over a rather large area with some employees having to travel considerable distances to the CSIR. During the time of the survey, several respondents undertook work related long-distance trips (i.e. Hartbeeshoek in North West province (distance of 120 km), Stellenbosch in the Western Cape (distance of 1600km) and East London in the Eastern Cape (1000km).

In addition to the above spatial and temporal information, verification information was also collected (such as home locations) to compare the tracked information with respondent noted information. This included GPS coordinates and activity diaries for a few selected respondents.

5. ACCURACY AND QUALITY OF THE DATA

Generally, the tracked path compared well with the GPS tracked path and the activity diaries collected. In some instance cellphone records proved more accurate data as some respondents omitted short, incidental trips (such as stopping at the local video store). However, in some instance recoded activities were wrongly coded as activities when in fact it was "noise" in the data. Nearly all the respondents 'recognised' their daily activity pattern form the maps and could verify the various locations visited (i.e. work, home and other).
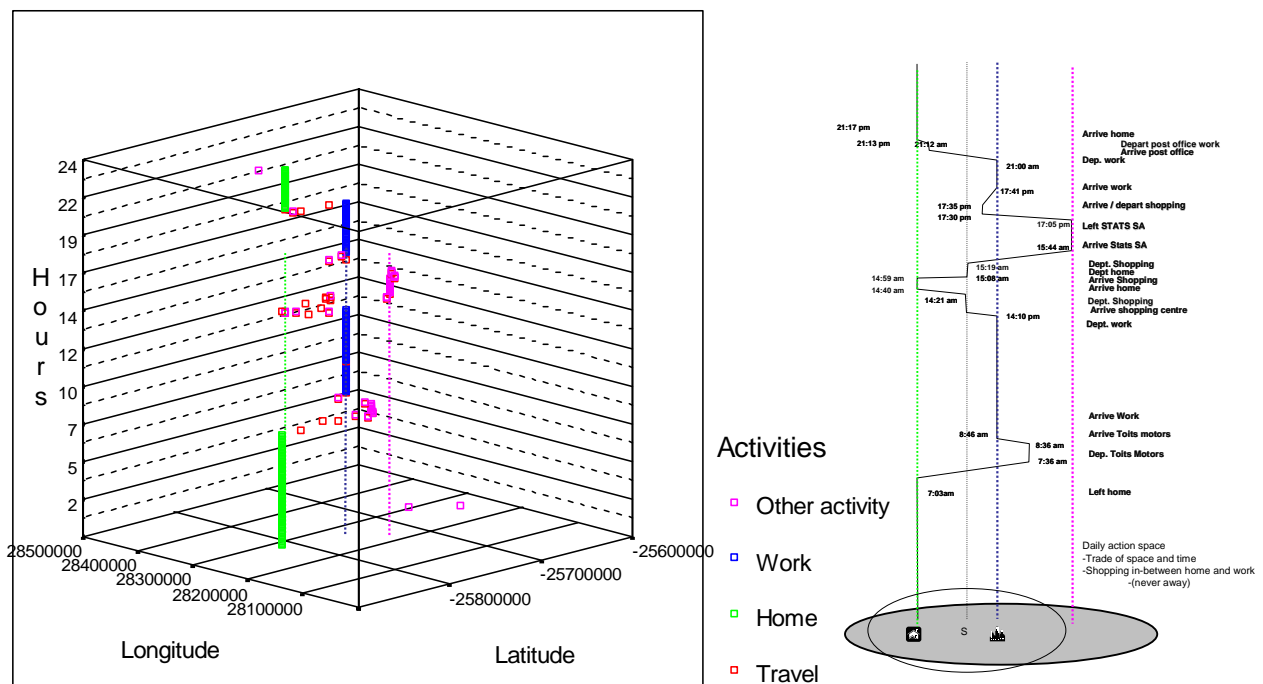

Figure 3. Comparison of cellphone tracked record with daily activity pattern.

Figure 3 compares the cellphone tracked records with a activity/travel dairy of a respondent. In this instance the match between the two is very high with even a short incidental stop on the way to home (i.e. a short 5-minute stop at the post office) capture and recorded as an 'Other Activity'. Figure 3 also shows two 'noise' locations in the data. These can be relatively easily cleaned.

6. INTERPRETED ACTIVITY AND TRIP PATTERNS

The interpreted spatial and temporal paths for 4 individuals are plotted in Figure 4. The interpreted pattern shows the home origin, work destination and other activity locations. Both figures show that the interpreted signal points are intuitively appealing. That is, stationary points at night and during early morning are classified as *home* while points during working hours are classified as work. Points between home and work are classified as 'other activity' while points with significant space between them are classified as moving.
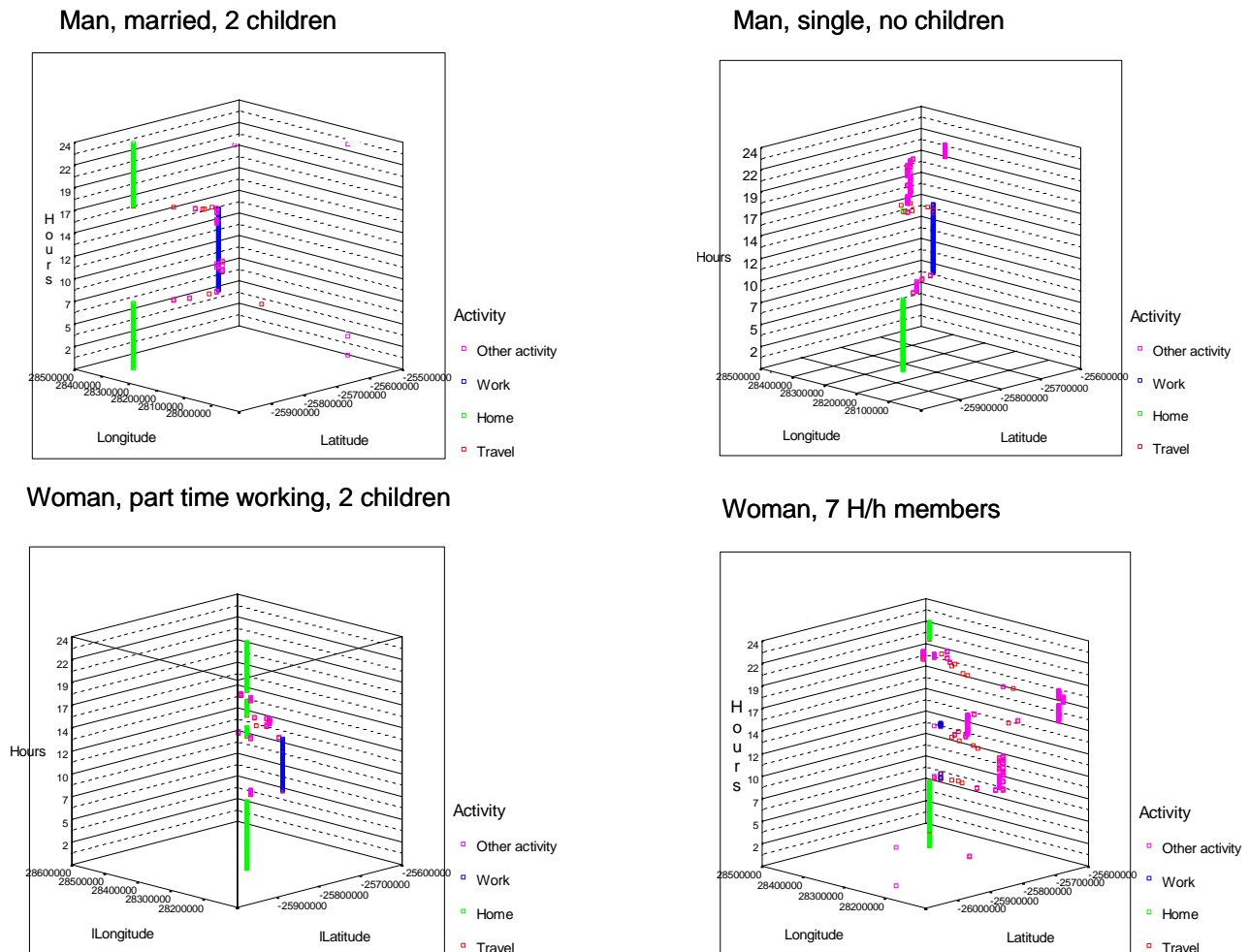

Figure 4. Interpreted temporal / spatial path for four respondents.

From the classified points, it was also possible to derive some descriptive statistics such as the number of out-of-home activities and trips per day, work duration, home duration, total daily travel time, travel time for work as well as departure and arrival time distributions, etc. Longer term tracking exercises allows researcher to explore longer terms daily rhythms in activity and travel behaviour.

Interpreted activity / travel pattern data can subsequently be mapped in a GIS and overlaid with various other feature data layers, including the (satellite) land use classification data, transport zones information, a land use map and the transport zonal data. Important attribute data can be appended to each record, relating for example to land use, neighbourhood, closest street, transport zone the individual belongs to, etc. Figure 5 shows the interpreted patterns of two individuals with complex patterns overlaid with a GIS base map of the road network. The figure shows th visit of one of the respondents to the airport
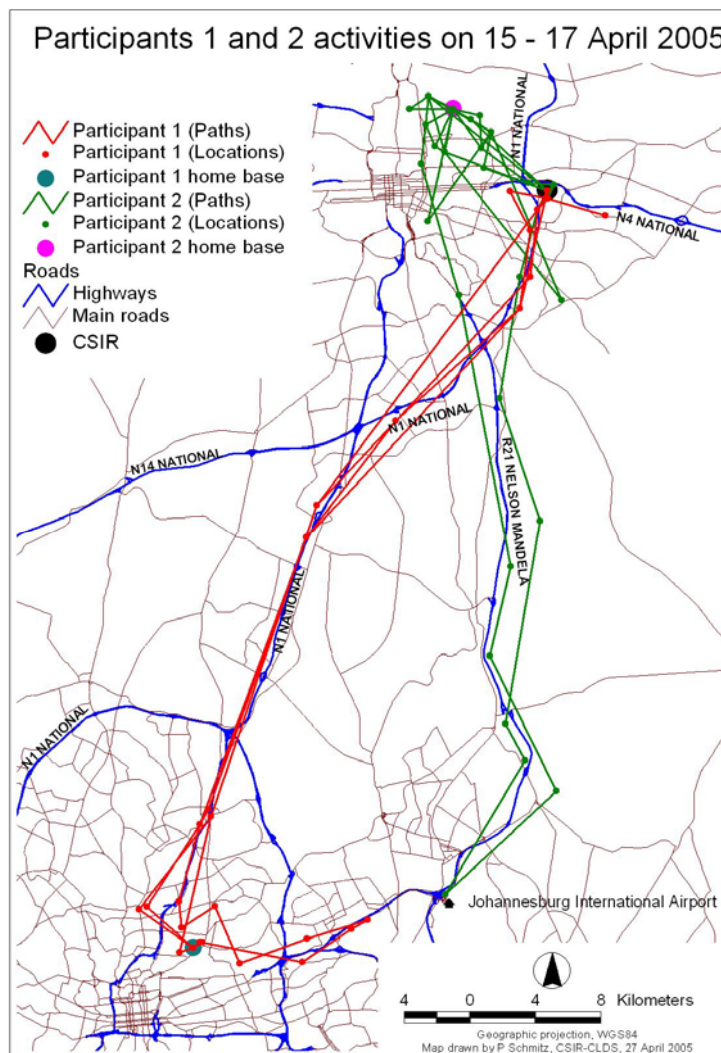
Figure 5. Cellphone records overlaid with road network.

7. CONCLUDING DISCUSSIONS AND DIRECTIONS FOR FUTURE RESEARCH

This paper presented the findings of a *proof-of-concept* study to obtain detailed information on individuals' daily travel and activity patterns using cellphone tracking. The proliferation and adoption of consumer devices such as GSM technology (and Assisted-GSM in the future) offers the opportunity to leverage existing hardware and software already in the hands of study participants for collecting detailed travel behaviour information and at the same time overcoming the well-known problems with complex and burdensome activity-travel diaries, specifically in a developing country (Wolf, 2004).

The proof-of-concept study showed that it is possible to obtain individual-level travel behaviour data from cellphones. However, while information such as the origin and destination of trips (and the linking thereof to transport zones) can be 'extracted' from cellphone data without significant labour and technological intervention, obtaining a more richer description of individuals' entire daily activity and travel patterns proved considerably more complex requiring (at least initially) much more methodological innovation.

To extract detailed daily activity and travel patterns entailed a complex (iterative) process relying on labelling rules and algorithms to clean, verify and ultimately transform the raw data to travel pattern data. Applying the rules and algorithms to the data resulted in a very real activity and travel path for each individual from which valuable relevant information

(departure time, arrival time, activity duration, activity sequencing, etc.) could be extracted for use in conventional transport modelling studies and more advanced activity-based research.

While the exercise was generally a success, the researchers did experience some problems in cleaning and labelling the raw data and several areas for further research are considered. These include improving the cleaning and labelling rules and algorithms, improving location accuracy (though triangulation and/or signal strength determination), allocating movement points to routes and integrating the data with various other contextual layers.

Furthermore, a need was also identified to replicate the exercise with larger and preferably a randomly drawn sample with a variety of work destinations. In addition, it is recommended that cellphone tracking is undertaken together with more conventional household surveys (to verify the tracked information and inform the rules and algorithms). Information on home and work locations and other daily activity engagements can be used to verify the tracking data and provide information on individual socio-demographic characteristics not available via cellphone tracking.

Cellphone data collected by way of passive tracking only provides locational and temporal data. When used in transport and land use studies, the data should still be supplemented with GIS layers of land use and other spatial features. Depending on the application (e.g. retail services, transport, logistics) cellphone data will have to be supplemented with additional data, such as home interviews, telephone surveys and respondent stated and revealed preference surveys. Cellphone tracking, however, does take care of one of the more complex data requirements, i.e. accurate temporal and spatial activity and (possibly) route location information. It is also possible to use cellphone tracking as add-on surveys with other more routine surveys such as OD and national travel surveys. The cellphone information can then be used to identify under-reporting of trips/activities (and hence be used to derive correction factors for the larger, conventional surveys), obtain accurate location information and correct for incomplete, missing, or inconsistent trip details.

The area holds much promise as consumer driven services are driving location based services and leading to improvements in the accuracy and quality of service. The integration of GPS in cellphones is an area of continuing technology improvements which promises to provide a wealth of information for users of cell phones, personal digital assistants and other mobile devices. Similarly, these services will provide transport planners and other infrastructure researchers with the detailed information required to make more informed decisions.

The research illustrated that it is possible to use cellphone technology to address the critical data needs of developing countries. However, as illustrated by several technology projects in developing countries, no technology is a single panacea for all data and research requirements (or activity-based data requirements). The technology requires careful implementation and appropriate business models. Furthermore, the integration of the technology in transport information collection and transport planning requires an understanding of where the technology ends and transport planning begins.

8. ACKNOWLEDGEMENTS

in providing the geo-located information. The author(s) remain solely responsible for any inaccuracy or oversights reported in the paper.

9. REFERENCES

[1] Asakura, Y. and Hato, E., 2005. 'Tracking survey for individual travel behaviour using mobile communication instruments'. Transportation Research Part C 12, Elsevier, pp. 273-291.

[2] Bhat, C. R. and Koppelman, F. S., 2000, "Activity-based travel demand analysis: History, results and future directions", TRB, Washington, D.C.

[3] Kitamura, R., 1997, "Applications of Models of Activity Behaviour for Activity-Based Demand Forecasting", Texas Transportation Institute, ed., Travel Model Improvement Program, Washington, DC. pp. 119-150.

[4] Kracht, M., "Tracking and interviewing individuals with GPS and GSM technology on mobile electronic devices", Seventh International Conference on Travel Survey Methods, Costa Rica.

[5] Pas, E. I., 1990, "Is Travel Demand Analysis and Modelling in the Doldrums?," In Developments in dynamic and activity-based approaches to travel analysis, Avebury Grower Publishing Company, pp. 3-33.

[6] Pas, E. I., 1995, "The Urban Transport Planning Process," In The Geography of Urban Transportation, 2 edn, S. Hanson, ed., The Guildford Press, New York, pp. 53-77.

[7] Pipkin, J. S., 1986, "Disaggregate Travel Models," In The Geography of Urban Transportation, S. Hanson, ed., The Guilford Press, New York.

[8] Sharp, J. and Murakami, E., 4 A.D., "Travel Survey Methods and Technologies Resource Paper", NATIONAL HOUSEHOLD TRAVEL SURVEY CONFERENCE: Understanding Our Nation's Travel, Transportation Research Board, Washington, DC.

[9] Stopher, P., 2004, "GPS, Location, and Household Travel," In Handbook of Geography and Spatial System, 1 edn, D. A. Hensher et al., eds., Elsevier, Oxford, pp. 433-449.

[10] Timmermans, H. J. P., 2000, "Theories and models of activity patterns," In Albatros: A Learning-Based transportation Oriented Simulation System, European Institute of Retailing and Services Studies (EIRASS), Eindhoven, pp. 6-70.

[11] Wermuth, M., Sommer, C. and Kreitz, M., 2004, "Impact of new technologies in travel surveys," In Transport Survey Quality and Innovation, P. Stopher & P. Jones, eds., Pergamon, Oxford, pp. 455-482.

[12] Wolf, J., "Applications of new technologies in travel surveys", August 2004, International Conference on Transport Survey Quality and Innovation, Costa Rica.