TRACKING CELLULAR TELEPHONES AS AN INPUT FOR DEVELOPING TRANSPORT MODELS

ANTONY K COOPER*, PETER MU SCHMITZ* and STEPHAN C KRYGSMAN†

Email: acooper@csir.co.za Telephone: 012 841 4121
* Built Environment Unit, CSIR, PO Box 395, Pretoria, 0001, South Africa
† SSI, PO Box 5195, Tyger Valley 7536, South Africa

ABSTRACT

During 2005, the CSIR ran a small experiment in the Pretoria area to test the feasibility of tracking passively cellular telephones, and hence people. Based on the success of that pilot, we initiated a large project to assess various aspects of tracking cellular telephones and using the data to populate transport and other models. We report here on one of the pilots, known as DYNATRACK (Dynamic Daily Path Tracking), a larger experiment conducted in 2007 with a more heterogeneous group of commuters in the Cape Town area. We discuss the technologies used to track participants and construct their travel routes, problems with recruiting participants, the ethical issues, and the results of the project.

1. INTRODUCTION

The primary mechanisms for capturing data about commuters' travel choices have been travel diaries, log books filled in by commuters giving details of all their trips, or origin-destination surveys, mainly providing information on the main *home – work* trip, mode used and time of travel. Unfortunately, these techniques are not always accurate, are difficult for illiterate commuters to complete and do not capture all the variables that impact on travel behaviour (e.g. time of travel, route section, intermediate stops on the way to/from work, etc). There is also increasing resistance to using them because of the effort required and the risk of providing data that could make respondents vulnerable to crime [Behrens 2001]. Another technique, recall surveys, has the inherent problem that people do not recall easily the details needed for accurate transport models. All these forms of survey are expensive, as well.

During March 2005, the CSIR ran an experiment to test the possibility of populating transport models by tracking passively cellular telephones (cell phones), and hence people. Sixty volunteers from the CSIR in Pretoria were tracked over a 48-hour period by 'pinging' their cell phones every five minutes – a connection to the phone would be initiated, which would cause the phone to tell the network in which cell it was (ie: the Cell ID), but the interaction would then be terminated immediately. Using the Flowmap software [Geertman *et al* 2003], developed by University of Utrecht's Faculty of Geosciences in collaboration with the CSIR, these cell IDs were linked and snapped to the road network to give the most likely routes travelled by the participants [Krygsman *et al* 2005; Krygsman *et al* 2007].

Proceedings of the 29th Southern African Transport Conference (SATC 2010) 16 - 19 August 2010
Proceedings ISBN Number: 978-1-920017-47-7 Pretoria, South Africa
Produced by: Document Transformation Technologies cc Conference organised by: Conference Planners

This small experiment showed that it was feasible to track the movement of cell phones as an input for developing commuter transport models. It also raised a number of questions, and the GenDySI project (Generation and harnessing of DYnamic Spatial Intelligence) was initiated because of the need for better ways to capture data about commuters (eg: origin-destination models), to build better transport models and facilitate better transport planning [Cooper *et al* 2009; Hauptfleisch *et al* 2006; Schmitz *et al* 2006; Schmitz *et al* 2007; Schmitz *et al* 2010]. GenDySI included a pilot project, known as DYNATRACK (Dynamic Daily Path Tracking), which took the 2005 experiment further, applying it to a larger, and more heterogeneous, group of commuters.

We report here on the DYNATRACK project, including the technology used to track participants and construct their travel routes, problems with recruiting participants, the ethical issues, and the results of the project. The paper also discusses some general trends and recent developments on sourcing dynamic travel data from geospatial technologies such as cell phones and global positioning systems, and assesses the viability of using the technologies in travel surveys.

2. OVERVIEW OF THE CELLULAR TRACKING TECHNOLOGY

Connectivity in a cellular telephone network such as GSM (Global System for Mobile Communications) is provided in *cells*, each of which is covered by a base station (antenna), which is either an omni-directional antenna (an antenna located in approximately the middle of the cell covering the cell in all directions) or a directional antenna (an antenna with a narrow coverage that is on the edge of the cell and that transmits into the cell). Typically in South Africa, a mast carries three directional antennas, each with a 60° arc of coverage. In reality, the shape of each cell is influenced by the terrain [Lee 1986]. Figure 1 gives an example of cells in a small area of Cape Town.

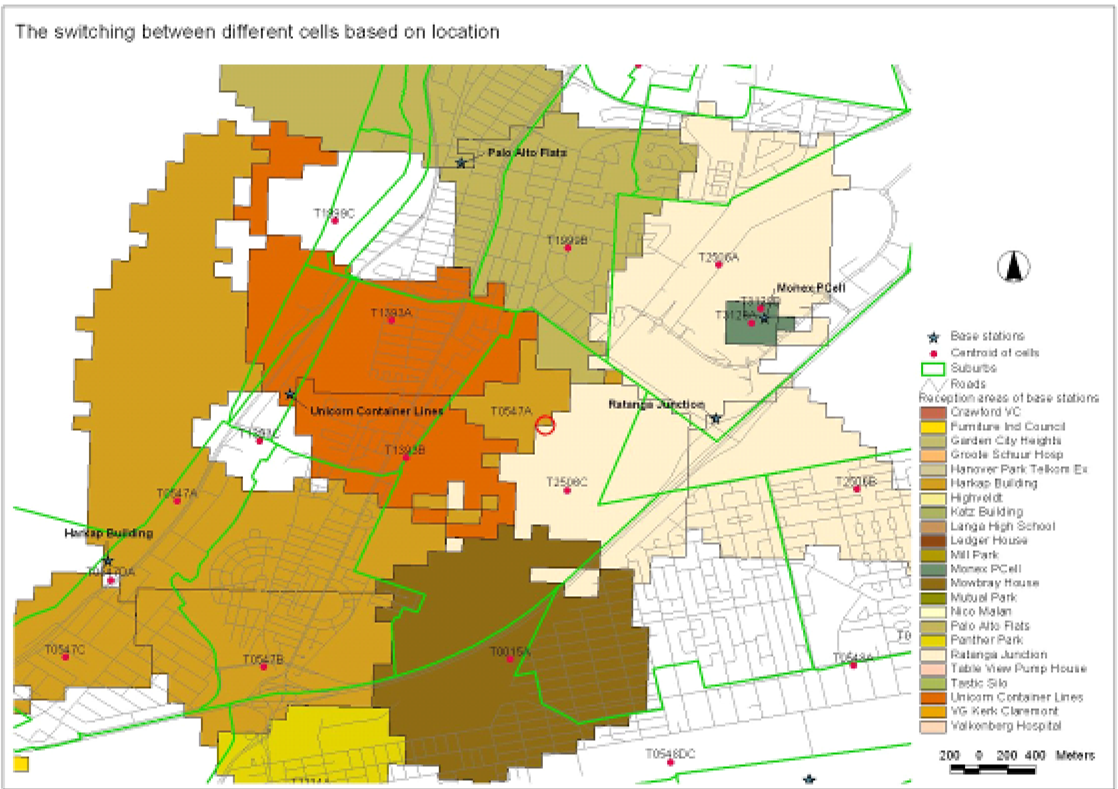

Figure 1: An example of actual cells in the Cape Town area.

When a cell phone is 'pinged', it responds to the network returning the cell ID of its preferred cell. This is usually the cell containing the cell phone, but in cases of sever congestion of the network in that cell (such as when a crowd has gathered for an event [Cooper *et al* 2009]), it might be a neighbouring cell. We use the centroid of the reporting cell to geocode the cell ID, that is, assign coordinates to it. This provides the approximate location of the cell phone at time *t* for geographical analysis, with its accuracy depending on the size of the cells (typically, about 1 km^2 in an urban area).

We then 'snap' these consecutive pairs of geographical coordinates and time stamps to the nearest road network using *Flowmap* [Geertman *et al* 2003]. It is dedicated to analysing and displaying flows (movement) between users (customers) travelling to facilities or services. Most thematic mapping and GIS packages have little functionality for treating this kind of information [Geertman *et al* 2003]. When snapping to the nearest road network, the assumption is made that the person using the cell phone is using that section of the road network. For technical details of the process, please see Krygsman *et al* [2005], Krygsman *et al* [2007], Schmitz *et al* [2006], Schmitz *et al* [2007] and Schmitz *et al* [2010]. Future research includes refining the model to differentiate between different modes of transport, such as by using fluctuations in the speed of travel.

3. SETTING UP THE DYNATRACK EXPERIMENT

We chose to run the DYNATRACK experiment in the Cape Town area, because of the good relationship we had with the relevant planning authorities there. The participants were recruited from amongst the employees of the City of Cape Town and the Western Cape Provincial Government, and their family and friends.

We obtained data for 163 participants from the city and 16 from the provincial government. The 163 participants from the city were divided up into six groups, to cover week day and weekend travel behaviour, starting on Monday, 12 February 2007. The 16 participants from the province were tracked starting on Thursday, 22 February 2007. Each group was tracked continuously for 48 hours, starting at 06:00 on their first day. Each cell phone was pinged at five-minute intervals.

In drafting the original proposal, we were conscious that the project could raise concerns over the invasion of people's privacy. The University of Pretoria allowed us to use the Research Ethics Committee of their Faculty of Humanities to review the project. The main focus of the review was ensuring there was proper informed consent from the volunteers who were tracked. Each participant then gave their written consent in advance to be tracked, and also had to give their final consent via SMS before their tracking actually started. Nevertheless other pilots in the GenDySI project showed how easily people forget they are being tracked [Cooper et al 2009]! For actual use in transport modelling, socio-economic data could be obtained and linked to each track, but this would have to be dealt with explicitly in the informed consent process. We did not do this as we were only assessing the feasibility of the technique.

4. RESULTS FROM THE DYNATRACK EXPERIMENT

Figure 2 shows the distribution of the cells pinged during the tracking exercise. Although most of the cells are located in Cape Town as expected, there were several cells pinged as far away as Langebaan in the north, Worcester and Ceres in the north-east and Hangklip in the south. The two centroids shown within the sea in False Bay are there because their cells are very large and extend well out to sea. Figure 3 shows all the possible routes used by the participants, based on snapping the cells used to the nearest roads, as explained in Section 2 above.

Figure 4 gives the results of the Monday group that was tracked for 48 hours. Most show the movement to and from the city of Cape Town since many of the participants worked in downtown Cape Town (where the main municipal offices are located). The map, however, also shows localised movement in Grassy Park, Kuilsrivier, Klein Constantia and the Strand. Owing to the removal of participants' private information, it was not possible to do follow-up studies to determine the reasons for the localised travel. These might well show the movements of family members who volunteered and don't travel much, such as the retired or stay-at-home spouses.

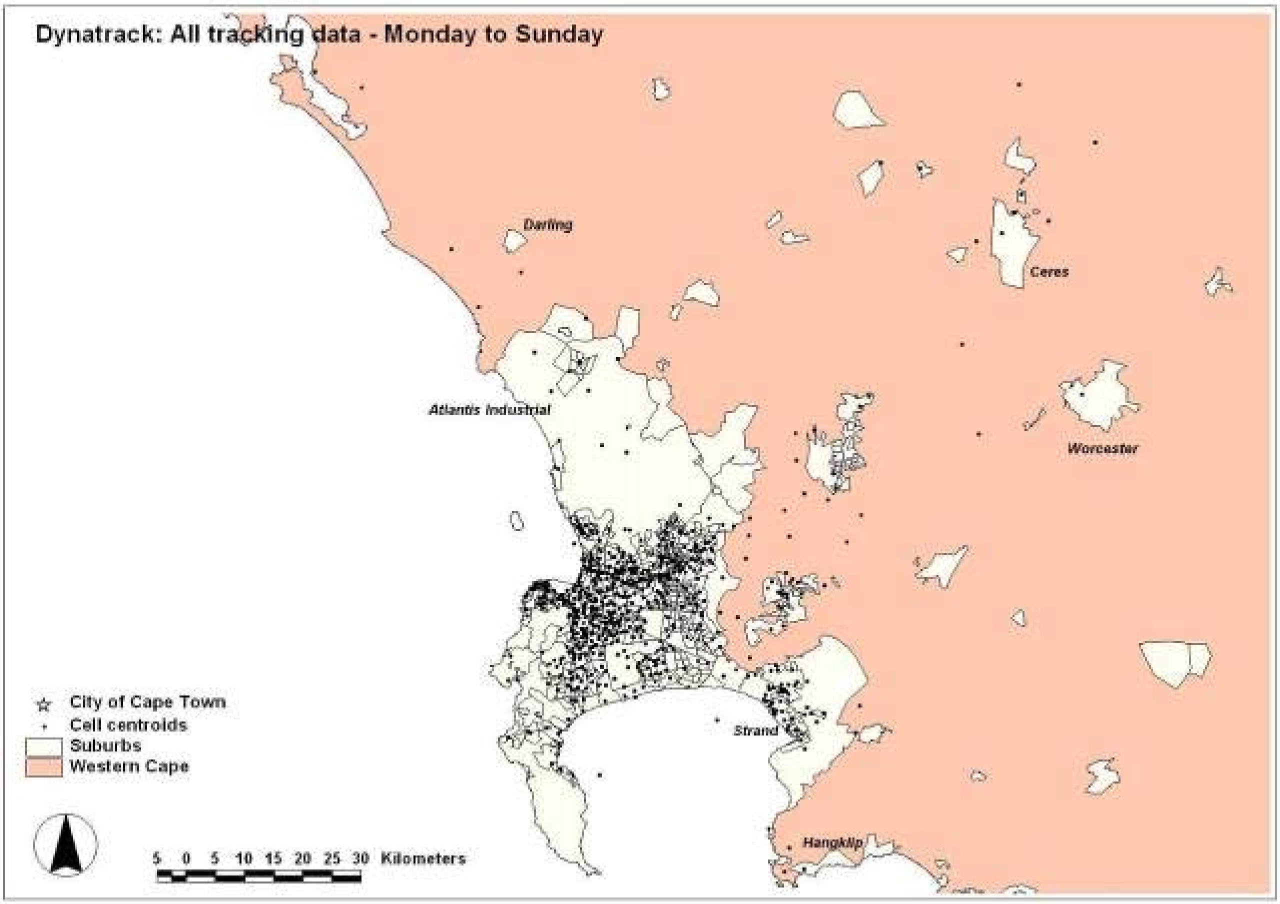

Figure 2: All the cells that were used during the tracking activity.

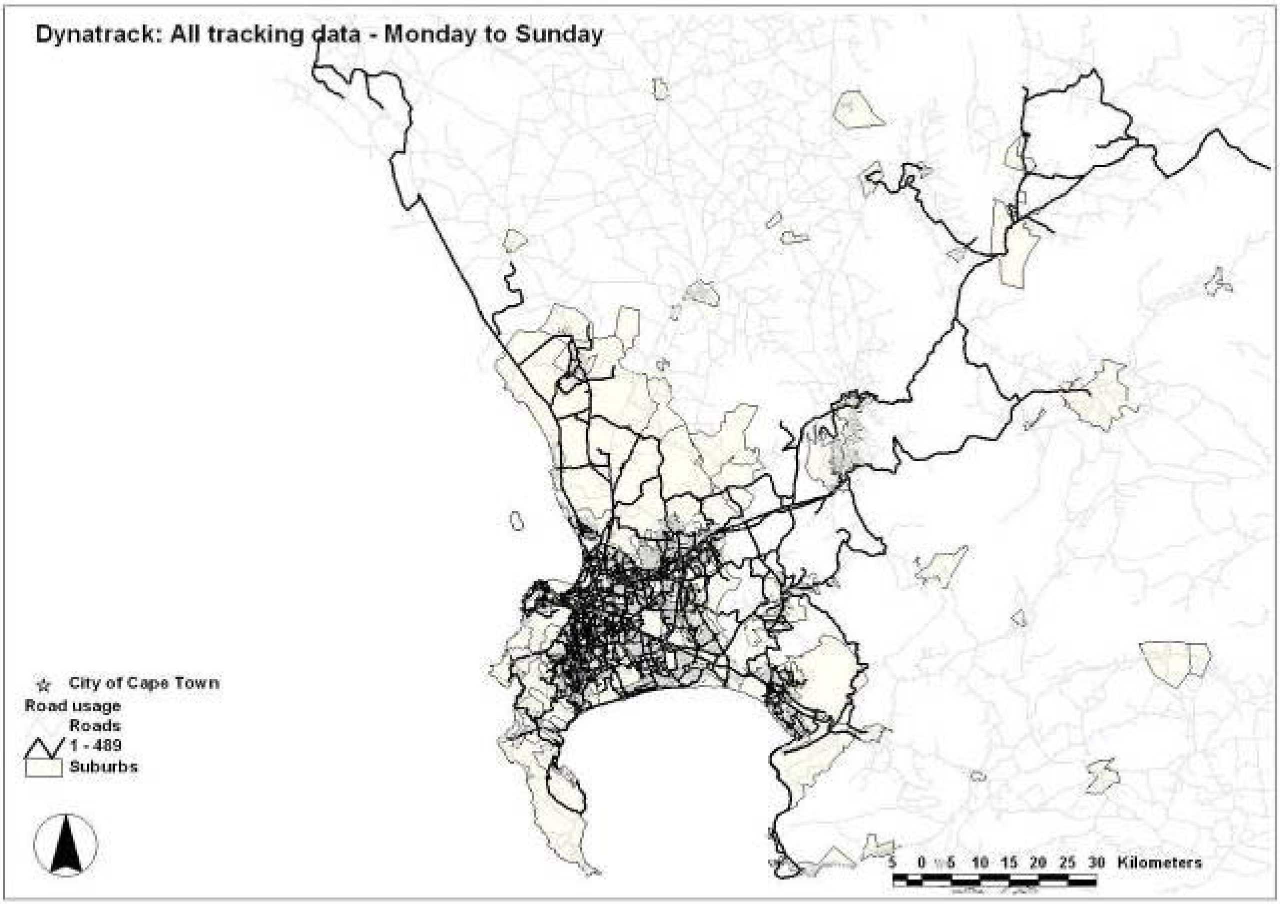

Figure 3: The modelled road usage based on the tracking data.

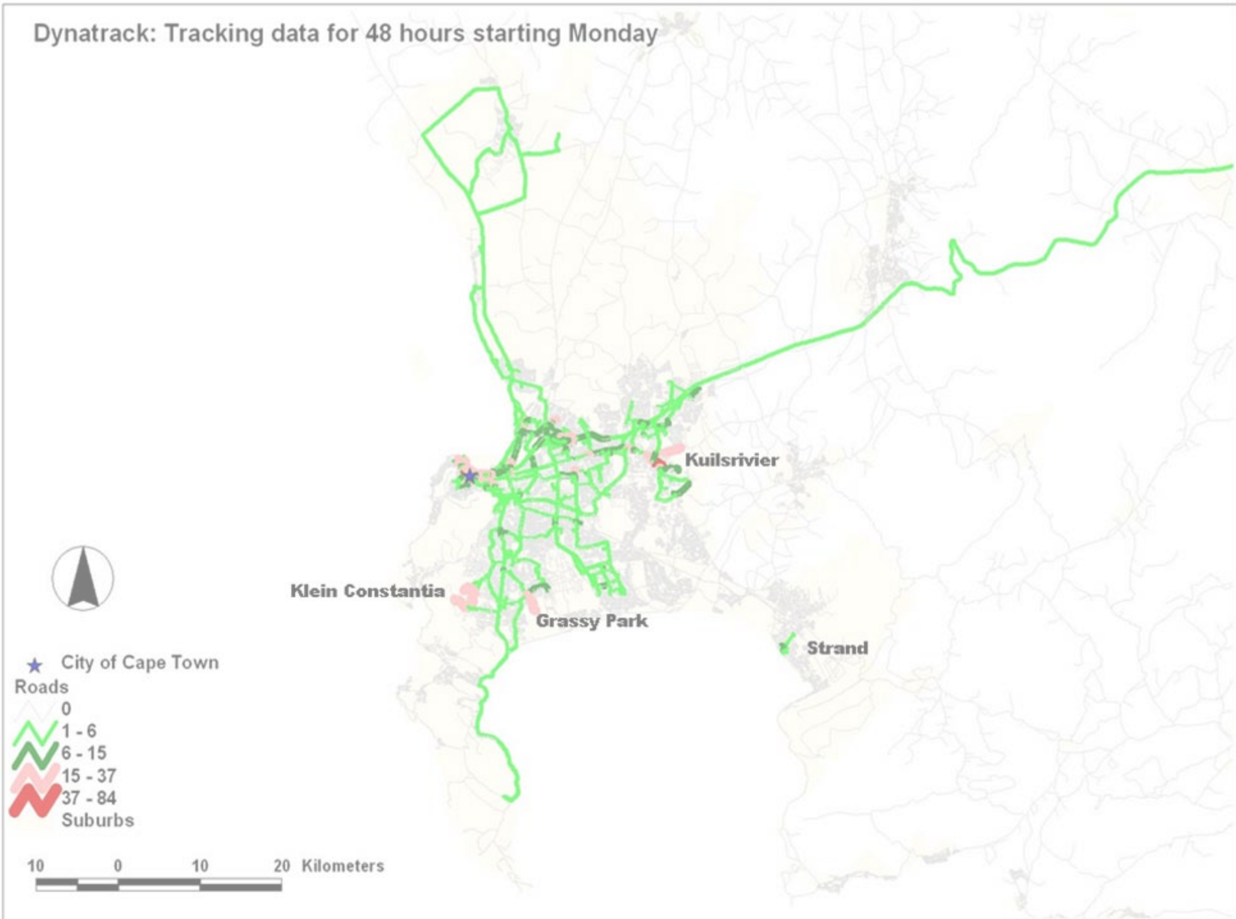

Figure 4: The modelled road usage for 2 days starting Monday

Figure 5 illustrates the morning peak traffic between six and eight o'clock in the morning, produced by combining all the travel data for Monday to Friday. It clearly indicates the movement of the participants to the downtown area of Cape Town, where the main municipal offices are located. Most of the movement occurred along the N1 and N2 national roads. Some of the participants travel to Cape Town from as far away as Atlantis and the Strand. When compared to Figure 4, one can see that those shown travelling from the Strand to Cape Town during morning rush hour did not do so on the Monday, and on that day there was only local travel in the Strand. The participant(s) travelling locally in the Wellington area could be treated as outliers and excluded from any Cape Town travel study.

The afternoon peak (Figure 6) shows a pattern similar to that of the morning peak. There are two exceptions, namely the participants using the N1 towards or from the interior and those travelling to or from Langebaan, indicating the participants either leaving or arriving in Cape Town. Again, the participants travelling locally in the Wellington area could be treated as outliers and excluded from any Cape Town travel study.

The overall working week road usage (Figure 7) is similar to that of the morning and afternoon peaks. This also shows the movement of participants outside of the two peaks to areas such as Simon's Town, Stellenbosch and Wellington. These movements may be work related or recreational and follow-up studies or interviews would be required to establish which they are.

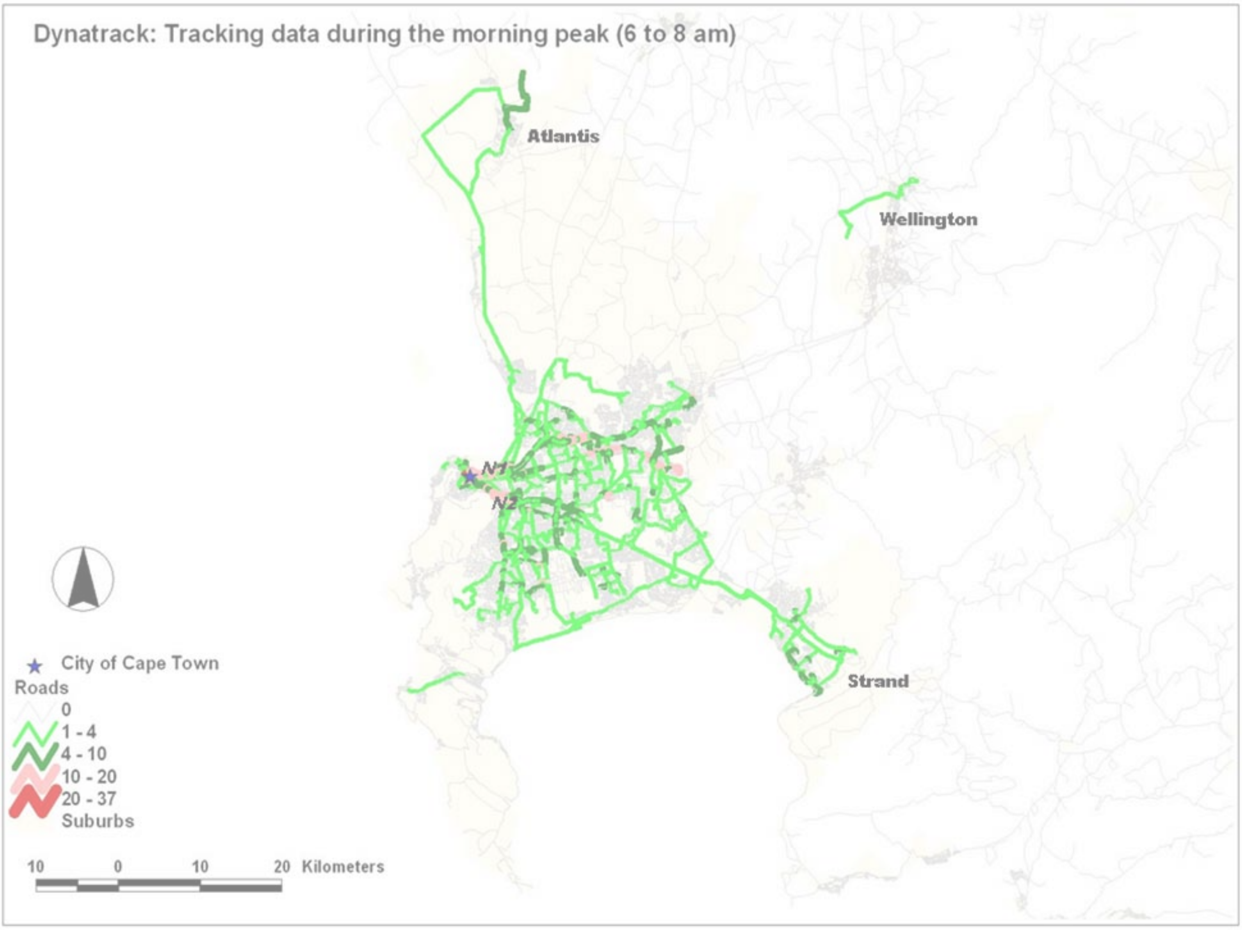

Figure 5: Morning peak hour road usage

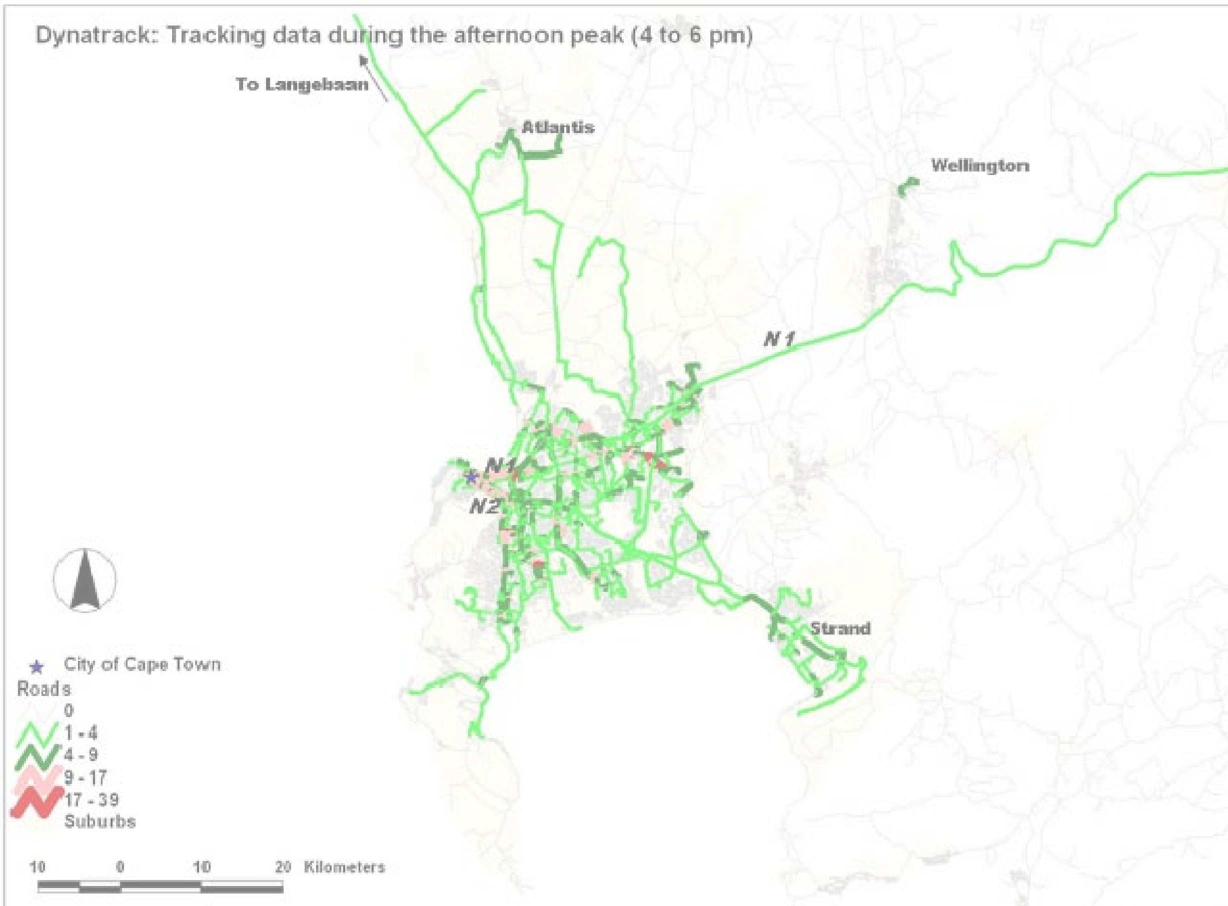

Figure 6: Afternoon peak hour road usage

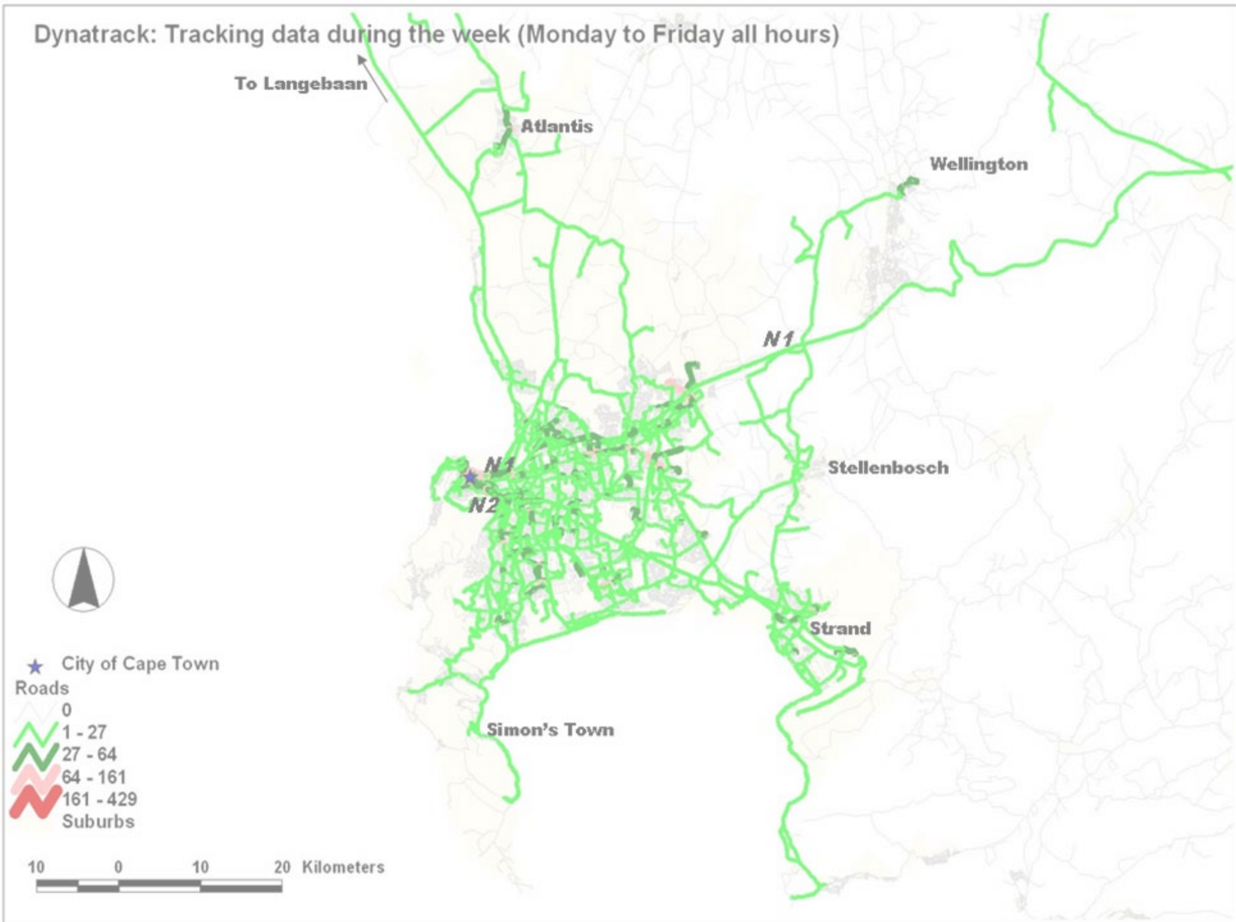

Figure 7: Working week road usage

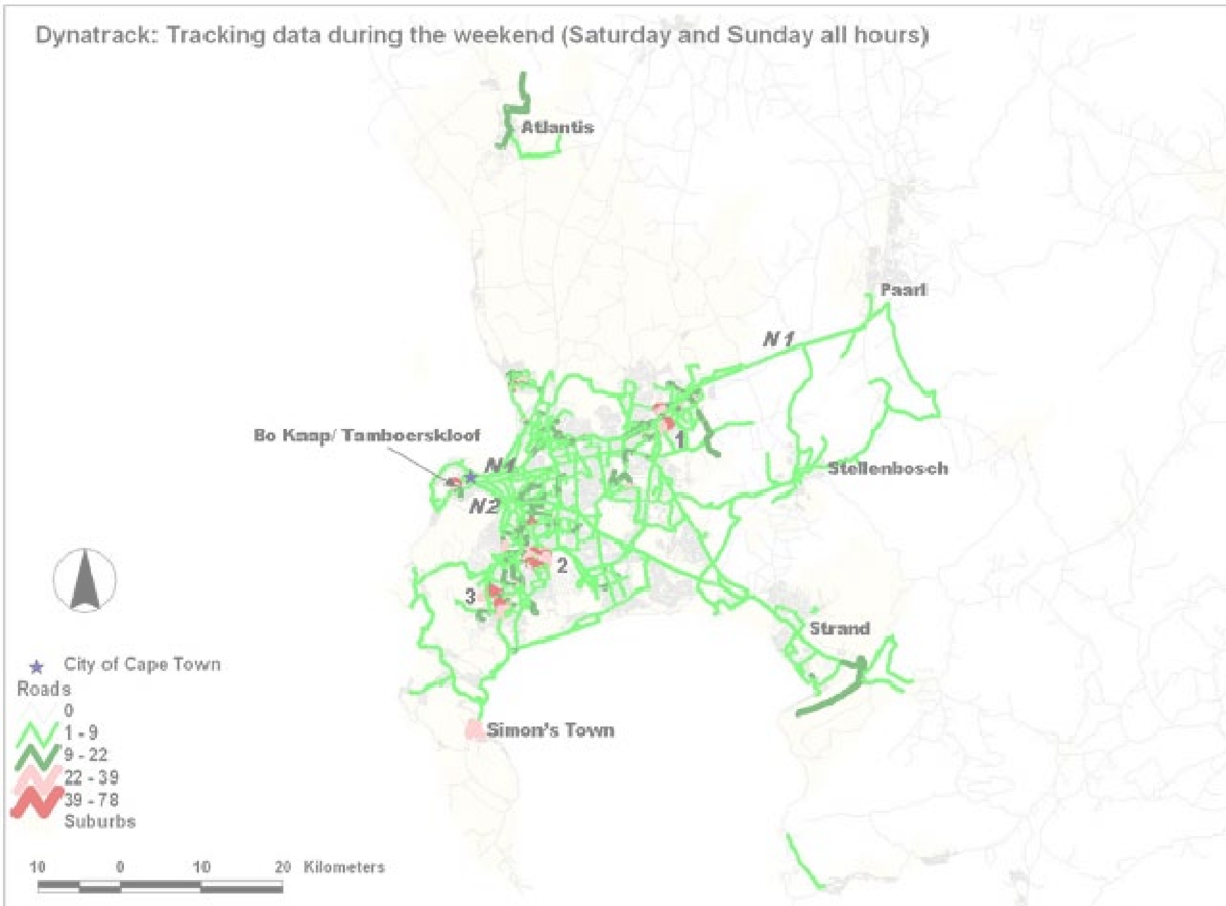

Figure 8: Weekend road usage

Figure 8 shows the travel patterns of the respondents tracked over the weekend. Three distinct nodes can be seen. The first node, indicated by 1 on the map, shows a movement to the Tyger Valley shopping mall which is north of the N1. The activity south of the N1 could be the result of the use of the recreation area at the Danie Uys Park and the Middestad Mall and Bellstar Junction shopping centres. The second node (2) is around the Kenilworth Race Course, Royal and the various shopping areas in Ottery and Wynberg. The third node (3) is in the Dieprivier, Bergvliet and Tokai suburbs. There are also trips into the wine lands in the Stellenbosch and Paarl area, as well as to Simon's Town, the Bo Kaap and Tamboerskloof, which all have recreational attractions.

5. ANALYSIS OF THE RESULTS

The experiment was not performed with a random or any other representative sample of the commuters of greater Cape Town, so it is not possible to draw any general conclusions about the commuting patterns in the area from the sample. Nevertheless, there were some interesting results that could be significant and worth investigating. These include:

- Noticeable differences between the aggregated patterns for the morning and evening rush hours (06:00-08:00 and 16:00-18:00).
- Noticeable differences between the weekday and weekend travel, with the weekday travelling appearing to be more widespread.
- Other results relevant for modelling include: time of the day of travelling, speed of travel, comparison between actual route taken and shortest route between origin and destination, number of daily trips and total time spent travelling.

6. ISSUES RAISED BY THE EXPERIMENT

We had great problems with recruiting sufficient volunteers to be tracked for the experiment, for various reasons. A key problem at the time was that we could only obtain suitable data through the service provider Vodacom, but MTN was the primary service provider for the City of Cape Town. Initially, we were going to run the experiment towards the end of November 2006, but we were able to recruit only 21 volunteers, and had to postpone the experiment to February 2007.

We were advised by several people who had conducted transport studies before that it was essential to offer incentives to volunteers. We ruled out corporate gifts because of the logistics involved in getting them to Cape Town, the uncertainty over the availability of suitable gifts, and the costs. We decided to offer draw prizes: a first prize of R 5 000.00, a second prize of R 3 000.00 and a third prize of R 2 000.00.

The final consent to participate was given by volunteers by responding to an SMS sent to them at 06:00, at the start of their 48-hour track. Unfortunately, 38 volunteers did not respond, and hence were not tracked. In retrospect, it was probably a mistake to send the SMS so early in the day. For reasons of privacy, we did not retain any data that can link those who were actually tracked to their consent forms or SMSs, so we cannot provide any details of the profile (age, gender, etc) of those actually tracked.

For the mapping, we used a commercial road network data base that allegedly was up to date. Unfortunately, there were glaring omissions in the data, such as the absence of the Huguenot Toll Tunnel (under the Du Toits Kloof Pass), which was opened on 18 March 1988. The GenDySI project also included research on generating automatically up-to-date road networks from remotely sensed imagery [Hauptfleisch *et al* 2006].

7. ETHICAL ISSUES

As a part of the overall GenDySI project, we conducted research into the ethical issues in tracking cell phones – see Cooper *et al* [2009] for details.

For all our pilots, we took great care to ensure that we had the informed consent of all the participants, particularly by taking the project through a research ethics committee (it is our understanding that the quality of the documentation, the responses to the questions of the committee and the approach to the review were significantly better than average for such research projects) and requiring participants to confirm their participation at the beginning of the tracking (though other pilots in the GenDySI project showed how easily people forget they are being tracked [Cooper *et al* 2009]). However, this has not been the case in other projects. For example, in Atlanta, GA, United States of America, two companies have been building traffic models by monitoring the locations of cell phones anonymously, but without informed consent, or the participants even being aware they are being tracked [Anon 2006]. Each company uses its own patented technology, but both involve introducing equipment into the network to bleed off information from the network [AirSage 2007, Intellione 2007]. It appears that both systems take two successive pings of a cell phone to create a vector without identifying information (ie: the vector is anonymous), and then aggregate these vectors to create traffic flow data. Hence, they would claim they do not need informed consent because the data are made anonymous before used. Of course, this assumes that one cannot relate any vector back to its source cell phone.

8. CONCLUSIONS

We report here on an experiment that we conducted to test the feasibility of tracking passively cellular telephones, as an input for populating transport models. While we were able to gather a lot of data, we experienced a number of problems with the project, especially with recruiting a sample large enough and representative enough of the commuting population of a city. It appears that with a representative or statistically valid sample, the technology and techniques used in this experiment can provide better data for populating a transport model than the traditional methods. The advantages of tracking cell phones include being passive (the users do not have to complete travel diaries or remember their trips); it is independent of the literacy levels and languages used by participants; cell phones are pervasive and hence can get data across socio-economic groups; and being mechanistic, it reduces bias in how participants complete the survey documents. Nevertheless, as with any other technique, the data obtained from tracking cell phones would need to be combined with other data, for both validating and supplementing the data.

9. ACKNOWLEDGEMENTS

The authors would like to thank the CSIR for funding this research through its Strategic Research Panel, project number PPTH/2005/036, “*GenDySI (Generation and Harnessing of Dynamic Spatial Intelligence)*”. We would also like to thank the many volunteers who participated in the DYNATRACK project; the City of Cape Town and the Western Cape Government for their support for the project; and the University of Pretoria, especially Professor Brenda Louw, for providing the project with access to a research ethics committee to assess the project.

10. REFERENCES

AirSage, 2007, “Redefining Traffic Information”, company's Web site, accessed on 23 April 2007: http://www.airsage.com/

Anon, 5 November 2006, “Tagging Phones to Track Traffic”, Wired.com [sourced from Associated Press], http://www.wired.com/techbiz/media/news/2006/11/72078, accessed 23 April 2007.

Behrens R, August 2001, “Looking beyond commuter travel in Cape Town: Methodological lessons from the application of an activity-based travel survey”. International Conference on Transport Survey Quality and Innovation, Cape Town, South Africa, 30pp.

Cooper AK, Ittmann HW, Stylianides T & Schmitz PMU, December 2009, “Ethical issues in tracking cellphones at an event”, special issue on ethics and operational research of OMEGA, Vol 37, No 6, pp 1063-1072.

Geertman S, De Jong T and Wessels C, 2003, "Flowmap: A Support Tool for Strategic Network Analysis", in S Geertman and J Stillwell, eds, Planning Support Systems in Practise. Berlin: Springer Verlag.

Hauptfleisch AC, Van Den Bergh F, Bachoo AK and Engelbrecht AP, 2006, "Comparison of canny and V1 neural network based edge detectors applied to road extraction", 17th Annual Symposium of the Pattern Recognition Association of South Africa, Parys, South Africa, 29 Nov - 1 Dec 2006, pp 6.

Hellinga B, Fu L & Takada H, June 2005, “Traffic network condition monitoring via mobile phone location referencing – an emerging opportunity”, 6th Transportation Specialty Conference, Toronto, Canada.

Intellione, 2007, “Life Made Simple”, company's Web site, accessed 23 April 2007, http://www.intellione.com/

Krygsman SC, de Jong T and Schmitz PMU, 2007, "Capturing daily urban rhythms: the use of location aware technologies", Computers in urban planning and urban management, 10th International Conference, Iguassu Falls, Brazil, 11-13 July 2007, 14pp.

Krygsman SC & Schmitz PMU, 2005, "The use of cell phone technology in activity and travel data collection", 24th Annual Southern African Transport Conference (SATC), Pretoria, South Africa, pp 11-13.

Lee, W.C.Y., 1986: Elements of Cellular Mobile Radio Systems. In “IEEE Transactions on Vehicular Technology”, Vol. VT-35, No 2, May 1986.

Schmitz PMU, Cooper AK & Krygsman S, December 2006, “Mobile cartography: The use of cellphone technology in activity and travel data collection”, ICA News number 47, pp15-16. Reprinted as “Mobile cartography: Using cellphones for data collection” in PositionIT, Jan/Feb 2007, pp 47-48.

Schmitz PMU & Cooper AK, “Using mobile phone data records to determine criminal activity space”, IQPC International GIS Crime Mapping Conference, Brussels, Belgium, 25-26 September 2007, pp 1-29.

Schmitz PMU, Cooper AK, De Jong T & Rossmo DK, 2010, “Mapping Criminal Activity Space using Cellular (Mobile) Telephone Data”, in preparation.