Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities

John Steenbruggen · Maria Teresa Borzacchiello · Peter Nijkamp · Henk Scholten

Abstract The use of wireless location technology and mobile phone data appears to offer a broad range of new opportunities for sophisticated applications in traffic management and monitoring, particularly in the field of incident management. Indeed, due to the high market penetration of mobile phones, it allows the use of very detailed spatial data at lower costs than traditional data collection techniques. Albeit recent, the literature in the field is wide-ranging, although not adequately structured. The aim of this paper is to provide a systematic overview of the main studies and projects addressing the use of data derived from mobile phone networks to obtain location and traffic estimations of individuals, as a starting point for further research on incident and traffic management. The advantages and limitations of the process of retrieving location information and transportation parameters from cellular phones are also highlighted. The issues are presented by providing a description of the current background and data types retrievable from the GSM network. In addition to a literature review, the main findings on the so-called Current City project are presented. This is a test system in Amsterdam (The Netherlands) for the extraction of mobile phone data and for the analysis of the spatial network activity patterns. The main purpose of this project is to provide a full picture of the mobility and area consequences of an incident in near real time to create situation awareness. The first results from this project on how telecom data can be utilized for understanding individual presence and mobility in regular situations and during non-recurrent events where regular flows of people are disrupted by an incident are presented. Furthermore, various interesting studies and projects carried out so far in the field are analyzed, leading to the identification of important research issues related to the use of mobile phone data in transportation applications. Relevant issues concern, on the one hand, factors that influence accuracy, reliability, data quality and techniques used for validation, and on the other hand, the specific role of private mobile companies and transportation agencies.

J. Steenbruggen (✉) · P. Nijkamp · H. Scholten
Department of Spatial Economics, VU University,
De Boelelaan 1105, 1081 HV Amsterdam,
The Netherlands
e-mail: john.steenbruggen@rws.nl

P. Nijkamp
e-mail: pnijkamp@feweb.vu.nl

H. Scholten
e-mail: hscholten@feweb.vu.nl

M. T. Borzacchiello
European Commission, Joint Research Centre, Institute for Environment and Sustainability, Spatial Data Infrastructure Unit, Via Enrico Fermi, 2749, T.P. 262, 21027 Ispra (VA), Italy
e-mail: maria-teresa.borzacchiello@jrc.ec.europa.eu

Keywords Traffic management · Transportation applications · Mobile phones · GSM network

Springer

Introduction

Location-based digital information—often originating from mobile phone data—has gained much popularity in recent years as a real-time operational vehicle for urban, environmental and transport management. Interesting applications are inter alia the use of private or public spaces by individuals (see, e.g. Calabrese et al. 2010), the concentration of people in a city (see, e.g. Reades et al. 2009), the activity spaces of commuters (see Ahas et al. 2006), non-recurrent mass events such as a pop-festivals (see, e.g. Reades et al. 2007), the entry of tourists in a certain area of attraction (see e.g., Ahas et al. 2007, 2008), or the estimation of spatial friendship network structures (see Eagle et al. 2009). Especially in the transportation sector, the potential applications are vast, and consequently, the use of cell phone data has shown a rapid increase in urban transport applications. These data offer a rich source of information on continuous space–time geography in urban areas. They can be used for daily traffic management, but also for incident management, for instance, in case of big fatalities, terrorist attacks, or mass social events such as festivals or demonstrations.

In the present paper we will analyse in particular the use of cell phone data for incident and traffic management in urban areas. The main question to be addressed is how to anticipate and control unexpected events in a transportation system, either on road segments or entire networks. Effective and timely control measures call for real-time detailed data on traffic movements. The possibility offered by micro-electronic devices to identify the geographic positions and flows of people opens unprecedented ways of addressing several policy issues such as urban security, incident control, organization of services for citizens, traffic management, risk management and so on.

In particular, the opportunity to gather real-time data about location and movements by means of mobile (or cell) phone activities may have an enormous impact on traffic management, given also the interests that private telecommunication companies might have in this market. Moreover, it immediately calls for real-time applications to city management, especially concerning the optimization and the regulation of the transportation system.

Intelligent Transportation Systems are based on the concept of a dynamic equilibrium between traffic demand and transportation supply. This might be achieved by means of a system able to orient its performance to the request that people have to move, in order to maximize the capacity of the system and to minimize the waste of energy and resources (Cascetta 2009).

Consequently, a system able to forecast the demand and to anticipate its evolution is needed. Presently, many efforts have been made to obtain models capable of forecasting traffic demand (econometric demand forecasting models, neural and Bayesian networks, stochastic processes, etc.) and to understand the way it moves on transportation networks (traffic flow models, etc.). The problem is that all these efforts have been only marginally tested on real and complex sites, since the cost needed to gather the huge amount of data required is, in most cases, unaffordable. As an example, the US Government has recently funded the very big NGSIM project (US Department of Transportation 2008)) aimed at providing, to the world's research community, data to test and to develop all possible traffic-related models. Albeit invaluable for very specific transportation applications, these data are collected by cameras only on short stretches (few hundreds of meters) of a set of roads in North America.

There are different techniques to collect traffic data: vehicles' trajectories are mostly collected by means of remote-sensing and object-tracking from video or photo cameras; positions of vehicles are obtained by applying Global Positioning System (GPS) technology, whose advantages are the high accuracy, the precise timing of the system and the high sampling frequency of the measures (Punzo et al. 2009), while the shortcomings are due to the fact that only a limited number of vehicles, equipped with GPS device, can be tracked. Loop detectors are the most widely used technique for traffic volume detection. The system is constituted by one or more magnetic loop detectors put in the road infrastructure, connected to a device able to pick data, located at one side of the road. To have detailed information about how loop detectors use magnetic properties to count traffic volume, we refer to Papageorgiou (1991).

In recent years, a new typology of data deriving from mobile phones, and in particular from the GSM network, has attracted the attention of researchers, due to the huge amount data that may be collected at the individual level, and to the possibility to obtain

high levels of accuracy in time and space. These features make mobile phone data ideal candidates for a large range of applications, in particular in the transportation field.

The history of GSM network is rather recent: in 1982, the European Commission on Postal and Telecommunication Administrations created the GSM (Groupe Spécial Mobile) to develop Second Generation Standards for digital wireless telephone technology (GSM Association 2009). In 1987 a memorandum of understanding was signed among 13 countries to develop the cellular system. The GSM (Global System for Mobile Communications) network was launched for the first time in 1991 and already in 1993 there were over a million of subscribers in 48 countries operated by 70 carriers (Emory University 2009). At present, 80% of the mobile market makes use of GSM technology in more than 212 countries, reaching over 3 billion people, (PR NewsWire 2009). Recent market surveys show that in various countries cellular phone penetration attains and, in some cases, exceeds 100% (Caceres et al. 2008).

Since mobile phones move with people and vehicles, the big market penetration is one of the advantages of the use of mobile technology for estimating traffic related parameters, once known the location of the device.

The first occasion leading to seriously consider the location potentialities of the mobile network stems from European and American regulations regarding electronic communications networks and services, according to which public telephone network operators receiving calls for the emergency calls number should make a caller's location information available to authorities in charge of handling emergencies (European Commission 2002a). These regulations motivated telecommunication companies to investigate the network capabilities of determining the location of fixed and mobile users.

Therefore, from the middle of the 1990s, several studies and projects have been carried out, and, in particular, over the past decade a number of research studies and operational tests have attempted to develop wireless location services in sectors like tourism, energy distribution, public transportation, urban planning, disaster management, traffic management, etc. Indeed, many fields nowadays require the use of location technology, and in several cases this need is inducted by the increasing speed of the

technology growth. The motivation for this paper is the need to systematize the literature regarding the use of mobile phone data in the field of the estimation of traffic parameters.

More specifically, against this background the aim of this contribution is to provide a review of past studies, projects and applications on wireless location technology, by highlighting the advantages and limitations of the process of retrieving location information and transportation parameters from cellular phones, and by trying to clarify: (a) which data types can be retrieved from the GSM network and how they are currently used; (b) whether it is possible to individuate a *fil rouge* among the number of studies in the field; (c) which are the main research issues connected with the use of telecom data in transportation applications.

The remainder of the paper is organized as follows: in the next section a short description of the most used mobile phone location methods is provided, while the literature review is presented in a subsequent section. Next, an illustrative application to the city of Amsterdam is offered. Various unsolved research issues and conclusions are discussed in the last two sections.

Mobile phone location methods

In order to understand the mechanisms allowing the derivation of the location of a mobile phone from the signals it sends to the network, it is worth clarifying how the GSM network works (www.gsmfordummies. com, accessed 29 July, 2009). It is relevant to note that in the present study novel kinds of network such as UMTS (Universal Mobile Telecommunications System) will not be considered, but they could be input for further research along the same lines.

As shown in Fig. 1, physically the Base Transceiver Station (BTS) is the Mobile Station's (the mobile phone, aka handset) access point to the network. A cell is the area covered by one BTS (not visible in the figure). The network coverage area is divided into a set of cells, named Location Areas (LAC). The BSC (Base Station Controllers) is a device that controls multiple BTSs. It handles the allocation of radio channels, frequency administration, power and signal measurements from the Mobile Station. The heart of the GSM network is the Mobile Switching Centre (MSC).

It handles call routing, call setup, and basic switching functions. An MSC handles multiple BSCs and also interfaces with other MSCs and registers. Location Management from a GSM network is possible by means of a system of databases, the HLR (Home Location Register) and the VLR (Visitor Location Register). The HLR is a large database that permanently stores data about subscribers, including the current location of the mobile phones. The VLR is a database that contains a subset of the information located on the HLR. It contains similar information as the HLR, but only for subscribers currently in its Location Area. The position of a mobile phone is derived from an automatic process that maintains the network informed about the phone location, depending on the phone status.

By means of a system involving the exchange of signaling messages between the phone and the network, the so-called Location User process is able to determine the position of the cellular at the Cell-ID level. The operator knows the coordinates of each cell site and can therefore provide the approximate position of the connected mobile. To overcome this approximation, two methods are mentioned in the literature (Promnoi et al. 2008); (a) Received Signal

Strength (RSS) methods, a technique that estimates the position of a mobile phone by matching the signal strength with the neighboring reference points; (b) triangulation, based on the difference of the arrival instant of the signal from the same handset to a set of different receiving base stations. The mobile station measures the arrival time of signals from three or more cell sites in a network. The network measures the transmission time of these signals from the relevant cell sites. By combining these two pieces of information it is possible to estimate the position of the mobile phone.

These methods need the network to be synchronized and require additional network elements which are not strictly necessary for the GSM communication, namely the SMS (Short Message Service) or IP traffic. For this reason, cell-based location data, without any form of improvement through RSS or triangulation, are currently the most used techniques.

Apart from location information, GSM network provides mobile phone activity parameters, which offer the information about the rate of use of the network. The most used activity parameters to estimate transportation parameters are handovers, cell dwell time and communication counts. A good
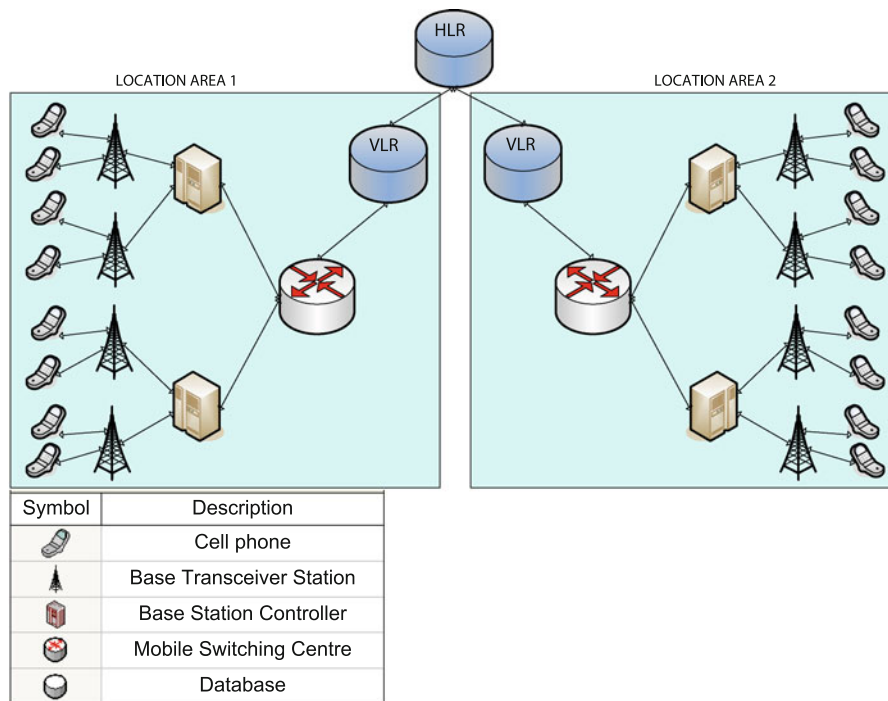

Symbol	Description
	Cell phone
	Base Transceiver Station
	Base Station Controller
	Mobile Switching Centre
	Database

Fig. 1 GSM network scheme

comprehensive review is given in Caceres et al. (2008) and Ratti et al. (2006).

Handovers (also called hand-off) refer to the switching mechanism of an on-going call to a different channel or cell. It is the mechanism of managing a permanent connection when the phone moves through two cells of the network. Hereby the phone call changes from one base station to the other without quality loss. This information is stored in the above mentioned HLR and VRL network databases. Together with the Mobile Switching Centers (MSC) they provide the call routing and roaming capabilities of the GSM network.

The Cell Dwell Time (CDT) represents the duration that a cellular phone remains associated to a base station between two handovers. This parameter is used in the literature referring to each individual cell, and thanks to the comparison among multiple adjacent cells it allows estimating traffic congestion.

The actual use of the network also provides useful indicators. The standard unit of measurement of telephone traffic used by most network operators is an Erlang, where one Erlang equals one person-hour of phone use. Such data is aggregated and made anonymous in terms of usage time and depends on the number of communications and their duration.

Telecom operators also measure a range of additional traffic features, for instance, for billing, network planning and network quality control. These include the number of new calls, the number of terminating calls, the average call length or the number of SMS messages. It is worth to clarify that data can be collected from the mobile phones not only when a call is made by the user, but also when the device is simply switched on.

In the next section it will be explained how such cell-phone parameters have been used in the literature so far to retrieve traffic parameters.

Review of projects using mobile phone data for traffic parameters estimation

In this section a short description of the most important field test deployments and simulation studies aimed at the estimation of traffic-related parameters is provided. Details are offered in Table 1, in which the projects are listed, where possible in chronological order.

Since there has been quite a number of review studies in the field, for a thorough description of the main projects and simulations studies, the interested reader is referred to Fontaine et al. (2007) and Caceres et al. (2008), which only exclude the most recent projects. Hereafter only the main features of the studies will be highlighted.

First attempts from US

The first recorded big project investigating mobile phones as vehicle probes is the CAPITAL project (Cellular APplied to ITS Tracking And Location), which started in 1994 (University of Maryland Transportation Studies Center 1997). It has been the first big project using an extensive set of data from a mobile company, and obtaining position through triangulation methods promoted by the academic, the public and the private sector as well (Table 1). Unfortunately, the location accuracy of about one hundred meters was not sufficient to obtain reliable traffic information (see Table 2).

Several other studies followed in North America. It is worth to mention here the US Wireless Cooperation Tests with deployments in San Francisco and Washington DC, using the RadioCamera technology (Yim and Cayford 2001; Smith et al. 2001), which, however, suffered from having a small sample size and from being able to track only the phones being in an on-call status. Together with CAPITAL, these early generation systems based on wireless signal analyses and triangulation had significant problems in determining true location of the cellular phone and were largely unsuccessful (Fontaine et al. (2007)).

European efforts

After the early American attempts there was a shift from wireless signaling analyses to handoff-based techniques. A first European effort to use the mobile cellular network for road traffic estimations based on handovers was initiated in Italy, in a simulation study by Bolla and Davoli (2000). Claiming to be the first attempt in this field, this study analyzes the use of location information to estimate on-line traffic conditions of important roads and highways by exploiting the presence of mobile phones on board of vehicles. The presence of a cellular terminal could be

Table 1 Summary of studies and field test deployments (in bold the citations of the review studies from which specific information has been derived)

Years	Project title (where available) (reference)	Promoters	Location	Data source	Used cell-phone parameters	Target traffic estimations	Results
1994–1997	*CAPITAL* (University of Maryland Transportation Studies Center 1997)	Federal Highway Administration; Virginia Department of Transportation; Maryland State Highway Administration; University of Maryland; Bell Atlantic NYNEX; Raytheon, Farradyne	Washington DC, USA	Data from Bell Atlantic NYNEX mobile's cellular network	Position through triangulation	Traffic speed	Only 20% of probes generated speeds; 100 m position accuracy; No consistent traffic monitoring
1999 2002	(White and Wells 2002; **Caceres et al.** 2008)	Highway Agency UK; Transport Research Laboratory; BTCell net (O2)	Kent, UK	Billing data from BTCellnet	Initial and ending position of the mobile phone	OD matrix	Small sample size; Phones on call; Groups able to transmit their position to server
2000	(Yim and Cayford 2001; **Fontaine et al.** 2007)	US Wireless Corporation; University of California-Berkeley	San Francisco Oakland, USA	44 h of wireless data from US wireless	Position of the mobile phone—call duration	Traffic speed	60 m mean location accuracy; 60% of locations could not be matched to roads; No usable data generated
2000 2001	(Smith et al. 2001; **Fontaine et al.** 2007)	US Wireless Corporation; Virginia Department of Transportation; Maryland State Highway Administration; University of Maryland; University of Virginia	Washington DC, USA	160 phone calls tracked every 2 s, generating 4,800 data points every minute	Cellular phone position	Traffic speed	5% of 10-min intervals had no data; 6–8 mph mean speed estimation error; Some intervals had errors >20 mph; Over 20% had significant differences from reality

Table 1 continued

Years	Project title (where available) (reference)	Promoters	Location	Data source	Used cell-phone parameters	Target traffic estimations	Results
2001	*STRIP* (Ygnace 2001; Ygnace and Drane 2001; **Caceres et al. 2008; University of Virginia Center for Transportation Studies 2006**)	INRETS SERTI French Government SFR carrier (Vodafone France)	Lyon, France	Mobile phone data from SFR	Position from in-vehicle mobile phones Number of phone calls	Journey times Traffic speed Directions of movement	Inter-city speeds overestimated by 24–32% Little speed variations on inter-city motorway Strong relation between call volume and number of accidents
2002	(Kummala 2002; Virtanen 2002; **Caceres et al. 2008**)	Finnish Road Administration Radiolinja	Finland	Mobile phone data from Radiolinja	Time required by each phone to cross a road section from the moment it enters the service area of a base station (cell) until the next one	Travel time Journey time	Validation with License Plate Recognition (LPR) More accurate results produced when the traffic was monitored over longer stretches around 10 km Location of the base stations not always optimal
2003	(Thiessenhusen et al. 2003; **Caceres et al. 2008**)	Institute of Transport Research of German Aerospace Center Vodafone	Munich, Germany	Mobile phone data from Vodafone	Handover	Traffic flow	Errors between 20 and 30 km/h Phone flows (calls) are closely related to vehicular flows Results based on small sample size

Table 1 continued

Years	Project title (where available) (reference)	Promoters	Location	Data source	Used cell-phone parameters	Target traffic estimations	Results
2004	(Rutten et al. 2004)	Mobile Traffic Service (MTS) LogicaCMG Vodafone	Noord-Brabant, Netherlands	MTS (Mobile Traffic Service) data from LogicaCMG and Vodafone	Handover Location update	Traffic flows Traffic congestion	High correlation between travel times generated by MTS and by the reference systems. Errors are generally low between 3 and 4%, with errors of 10–20% in journey times of 20–25 min Conclusions based on limited results that are publicly available
2003 2005	(University of Virginia Center for Transportation Studies 2006; Smith 2006; Fontaine et al. 2007)	Federal Highway Administration Virginia Department of Transportation AirSage inc. University of Virginia	Hampton Road, Virginia	Anonymous mobile phone data from Sprint US carrier	Handovers	Travel speed Travel time	68% of speed estimated had errors >20 mph No reliable measures Airsage claimed results caused by lack of access to full data
2005	MIT SENSEable City Laboratory (Ratti et al. 2006)	MIT A1 Mobilkom	Milan, Italy	Anonymous data from a European telecoms carrier	Erlang Network counters	Call density OD of calls	Real-time visualization dynamics metropolitan area
2006	(Maerivoet and Logghe 2007)	ITIS Holdings Proximus	Flanders Antwerp Belgium	Anonymous Cellular Floating Vehicle Data (CFVD).	Handover	Travel speed	Technology able to accurately detect traffic trends over time and per road segment Predictions, however most accurate in the case of free traffic flow rather than congested conditions

Table 1 continued

Years	Project title (where available) (reference)	Promoters	Location	Data source	Used cell-phone parameters	Target traffic estimations	Results
2005 2006	(Bar-Gera 2007)	Estimotion Ltd. ITIS Inc. Ben Gurion University	Tel-Aviv Israel	Cell phone data provided from Estimotion Ltd.	Handover	Travel time	Limited data during off-peak hours WLT estimates different between floating car and loop data by 10–30% during congested conditions
2006	**Real-time Rome** (Calabrese and Ratti 2006)	SENSEable City Laboratory MIT Telecom Italia	Rome Italy	Cell phone data provided from Telecom Italia Mobile	Erlang	Touristic, pedestrian and vehicle density Travel speed	Broad range of research directions demonstrating a large spectrum of useful applications
2007	(Liu et al. 2008)	Minnesota Department of Transportation University of Minnesota Sprint PCS	Minnesota, USA	Cell phone data from Sprint PCS Mobile network	Handover	Travel times	Segments with high speeds results within 10 MPH of ground truth conditions In segments with low to moderate speeds the results become occasionally scattered Under- and over-estimating travel times between AM and PM peak hours were found for some roads
2007	(Pattara-Attikom and Peachavanish 2007; Pattara-Attikom et al. 2007; Hansapalangkul et al. 2007; Hongsakham et al. 2008)	NECTEC NSTDA Thamasat University	Bangkok Thailand	Data from probe mobile phone in vehicles	Cell Dwell Time	Traffic congestion	Accuracy level congestion estimates between 73 and 85% Sample size includes phones in active and idle modes

Table 1 continued

Years	Project title (where available) (reference)	Promoters	Location	Data source	Used cell-phone parameters	Target traffic estimations	Results
2008 2010	**Mobile Millennium** (Liu et al. 2008)	Nokia, Navteq, and UC Berkeley, California Departments of Transportation	Hayward and Fremont area, California	GPS equipped cell phones	Virtual trip lines, position and speed of mobile phones	Speed and travel time	Provides real-time traffic conditions to end users. Free public traffic-information system. Deployment, currently tested
2007 2009	**Current City Amsterdam** (Vaccari et al. 2009)	Senseable Future Foundation University Salzburg Vrije Universiteit Amsterdam KPN Dutch Department of Traffic Management, (Rijkswaterstaat)	Amsterdam, The Netherlands	Data from Dutch KPN carrier	Network parameters	Spatial network signatures	Use of telecom network data for analysis of spatial network activity patterns

Table 2 Main field test project characteristics

	CAPITAL	US-wireless	STRIP	Real-time	Mobile-century
Promoters					
Public Agencies	X		X	X	X
Research Institutions	X	X	X	X	X
Mobile phone carrier/company	X	X	X	X	X
Location technology	Triangulation	RadioCamera	Signaling messages	Ad-hoc algorithm based on ToA, TA, CDT, TDoA	GPS
Location accuracy	107 m	60 m	100–150 m	Not available	Same as GPS, 10 m
	(**Fontaine et al.** 2007)	(**Yim and Cayford** 2001)	(**Caceres et al.** 2008)		(**Spinak et al.** 2009)
Useful traffic info	No	No	No	No	Yes
Use of new technology	No	RadioCamera	Abis/A probing	Yes	No
Validation	No	Loop detectors	Loop detectors	No	Loop detectors

detected at the vehicle's entrance in monitored roads. This quantity was then used to estimate average vehicle density, flow and speed in every cell.

Another early European example can be found in the UK (White and Wells 2002), where the Transport Research Laboratory (TRL) developed a system to generate journey times and traffic speeds from OD matrices based on billing data from the telecom network. This study only uses a subset of all monitored phones, resulting in a very small sample size.

In more recent years, in a number of European countries (e.g. France, Belgium, Germany, Spain, Austria, Finland, Italy, UK and The Netherlands) different field tests, simulation studies and evaluations took place. Most of these projects focused on

Fig. 2 Day-night pattern and weekend pattern for the traffic at WTC and Rembrandtplein (http://www.currentcity.org/)

how to obtain reliable travel times and travel speeds from the telecom network.

An extensive study of cellular probes has been carried out within the framework of the STRIP project (System for Traffic Information and Positioning) in Lyon, France (Ygnace 2001). This project evaluated the feasibility of 'Abis/A probing' location technology for travel time estimates. Abis/A Probing system is a network-based solution that gathers data from the cellular service providers. The system uses Abis and A interfaces, which include algorithms and databases of information to identify the location of a cellular phone. Results were compared with data from loop detectors, both on an inter-city motorway and an intra-city freeway, with major errors in the second case. A significant relationship between the number of outgoing calls and the level of incidents was found (Caceres et al. 2008).

Another example of travel time estimates was carried out in Finland by the FINNRA (Finnish Road Administration) in 2002 (Kummala 2002; Virtanen 2002), with the aim of estimating traffic data from mobile phone data exploiting the signaling messages exchanged between the phones and the network, eventually using License Plate Recognition (LPR) to validate the results (Caceres et al. 2008). There were more accurate results produced when the traffic was monitored over longer stretches of about 10 km. The data were affected by some problems such as parallel
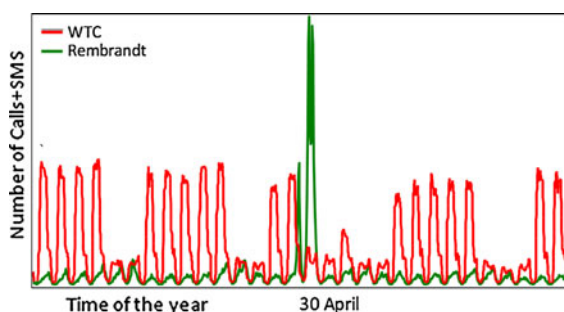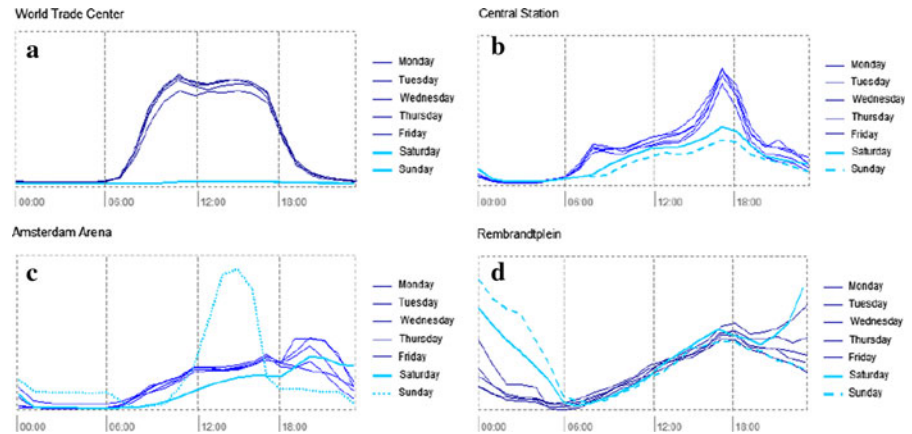
Fig. 3 Call intensity
(measured in Erlang during
a 24-h period) in different
Amsterdam city areas:
a business district, World
Trade Center; **b** transport
hub, Central Station;
c football stadium, Arena;
d entertainment nightlife,
Rembrandt square. *Source*
(http://www.currentcity.org/)

roads and pedestrians. Also the location of the base stations was not always optimal to support traffic management information systems.

Telecom companies projects

In 2003, the telecom carrier Vodafone, in collaboration with the Institute of Transport Research of the German Aerospace Center, used double handovers (a combination of data from two successive handovers from the same mobile phone, which is possible if the call duration is long enough) and signaling data from the network to generate information traffic flows and traffic speeds around Munich (Thiessenhusen et al. 2003).

LogicaCMG developed in 2004 the Mobile Traffic System (MTS) to monitor traffic speed and to provide road authorities with the possibility to manage traffic flows and traffic congestion (www.logica.com, accessed July 28, 2009). The system was tested in the province of Noord-Brabant in Netherlands and validated with field data from floating cars, number plate surveys and induction loop detectors.

The British company ITIS Holding developed in 2006 a pilot project based on the 'Estimotion' technique in the province of Vlaanderen in Belgium. They monitored traffic on highways to verify traffic speed between two arterials. Also here the objective was to assess whether data collected from mobile phones (e.g., travel times) provided accurate traffic information. The validation study compared traffic data from cellular floating vehicles with other traffic sources such as single inductive loop detectors and GPS-equipped probe vehicles. The general conclusion was that the technology was fairly able to

accurately detect the traffic trends over time and per road segment. The prediction was however most accurate in the case of free traffic flows rather than in congested conditions (Maerivoet and Logghe 2007).

In the TrafficOnLine project in 2006 in Germany, the already mentioned idea of double handovers was used (Birle and Wermuth 2006). In order to validate the results, double handovers, loop detectors and floating car data (FCD) from taxis equipped with GPS were compared. As a result, it was shown that mobile phones can provide a reliable detection of traffic congestion, depending on the covered area. Better results were obtained for motorways compared to urban roads. To improve the results in urban environments, information of existing buildings, which were responsible for handovers in overlapping coverage and signal strength of adjacent cells, were used. Problems were related to a small sample size, because only phones that made sufficient long calls within an entire cell were included. It was concluded that reliable data only could be generated in case a single roadway link exists into the border zone between two cells, so that it can be uniquely identified.

Recent projects outside Europe

Enlarging the view outside Europe, in North America a number of field studies have been carried out on the use of handovers to estimate traffic features. In 2003, Airsage deployed a monitoring system in the Hampton road region in Virginia based on cellular handoffs and transitions between sectors of cells to produce traffic speed and travel time. The University of Virginia performed the evaluation in 2005 and found

significant errors. It was concluded that, as of December 2005, the Hampton Airsage system could not provide the quality of data desired by the Virginia Department of Transportation (University of Virginia Center for Transportation Studies 2006; Smith 2006).

In 2005, in collaboration between ITIS holding and Delcan Corporation, another project was initiated based on the 'Estimotion' technology in Maryland (Delcan Corporation 2009). They used handovers to detect traffic events like congestion and accidents. The data were tested during 2006 by the University of Maryland, which found that average errors were approximately 10 mph on freeways and 20 mph on arterials. The quality degraded significantly during a.m. and p.m. peak periods.

In 2007, the Minnesota Department of Transportation carried out a field test around Minneapolis in collaboration with the telecom operator Sprint PCS network (Liu et al. 2008). The travel times and travel speeds were compared against ground truth conditions.

In 2008, around the San Francisco Bay Area, the Mobile Millenium project was started (Amin et al. 2008), whose aim is "to design, test and implement a state-of-the-art system to collect traffic data from GPS-equipped mobile phones and estimate traffic conditions in real-time" (http://traffic.berkeley.edu/theproject.html, accessed July 29, 2009). The project has organized a big field test deployment consisting in tracking the location and changes in position of informed users, carrying Nokia mobile phones equipped on purpose inside their vehicles. In exchange, participants received, free of charge, traffic information on the screen of their mobile. This project is still going on.

The study of Bar-Gera (2007) in Tel-Aviv compared the performance of the WLT data, detection loop data and floating car data to validate travel times. Intervals without congestion showed little variation of mean travel times.

In 2007, in a field test in an area around Bangkok in Thailand, some researchers have developed a methodology for detection and estimation of road congestion using CDT (Pattara-Attikom and Peachavanish 2007; Pattara-Attikom et al. 2007). CDT from multiple adjacent cells was used to estimate traffic congestion. The sample size includes mobile terminals in active mode (on call) and idle modes (turned on). They classified measurements in three levels of

traffic congestion based on duration. The results showed that the duration of CDT estimated the degree of congestion with an accuracy level between 73 and 85%. However they concluded that many issues need to be solved before actual implementation can take place.

Research on O-D matrix estimation

Regarding applications not concerned with travel times or travel speeds, which seem to be the main traffic parameters researchers are looking at, two recent (2007–2008) simulation projects on the OD matrix can be found in Spain (Caceres et al. 2007) and Korea (Sohn and Kim 2008) both focusing on a generation of traffic flows. The project in Spain concluded that turned-on phones (active and idle modes) of only one operator should be sufficient, and proposes an adjustment factor to transform phone data in vehicle data.

The project in Korea used a simulated environment for validation. They found that the accuracy of the estimation was less depending on the standard deviation of probe phones changing location than other factors like market penetration and cell dimension.

Research on urban behaviour

The Real Time Rome project (Calabrese and Ratti 2006) is one of the first examples of urban-wide real-time monitoring system that collects and processes data provided by telecommunications network and transportation systems, in order to understand patterns of daily life in the city of Rome. They address a broad range of research directions like: how do people move through certain areas of the city during special events (gatherings), which landmarks in Rome attract most people (icons), where are the concentrations of foreigners in Rome (visitors), and is public transportation effective where people are (connectivity).

In Reades et al. (2007) the authors analyze how cell phone data in Rome can provide a new way of looking at cities as a holistic dynamic system. This approach can provide detailed information about urban behaviour. Erlang data normalized over space and time are used to derive spatial signatures, which are specific time patterns of use of the mobile

network distinctive of a certain area. They found a mix of clusters suggesting a complex set of relationships between signatures. The visualizations generated an overall structure of the city with a correspondence between the levels of telecommunication and types of human activities. Finally, in Girardin et al. (2008) the use of cell phone network data and geo-referenced photos for the presence and movement of tourists with user-originated digital footprints are explored.

In the following table (see Table 1) an overview of the main information such as data source, promoters, and typology of results of the mentioned field projects is given, while in Table 2 a focus on the main characteristics of major past and recent projects is offered.

Illustrative application for Amsterdam

In 2007 the Current City consortium (SENSEable City Laboratory MIT; Salzburg University), in cooperation with the Dutch Ministry of Transportation, has realized a test system in Amsterdam (The Netherlands) for the extraction of mobile phone data and for the analysis of the spatial network activity patterns. This project is strongly connected to the earlier projects Mobile Landscapes in Graz (Ratti et al. 2007) and Real-time Rome (Calabrese and Ratti 2006). Later on, this will be explained in more detail, as it is the project from which the authors will start their further research. The project does not focus directly on traffic patterns, but explores space–time relationships of telecom data and assesses its suitability to derive census proxies and dynamic patterns of the urban area, which in turn can be utilized to derive mobility indicators, showing the possibility of extracting near real-time data from cell phone use and to reconstruct the spatial–temporal patterns of the telecom network usage (www.currentcity.org, accessed July 29, 2009).

The main objective of this project is to address the problem of Incident Management (IM). The Dutch ministry of Transport, Public Works and Water Management is responsible for maintaining over 3,200 km of main roads, ensuring that the infrastructure is safe and in a good state, and that the flow of vehicles is as smooth as possible. Approximately 12% of the traffic jams on Dutch roads are the result

of incidents such as crashes and vehicles shedding their loads (Ministry of Transportation and Water Management 2008). On a yearly basis there are about 100.000 incidents (Berenschot 2008), varying from small accidents to major multi-vehicle incidents causing casualties and vast damages to the road and its supporting structures. Incident Management (IM) refers to the entirety of measures that are intended to clear the road for traffic as quickly as possible after an incident has happened and to ensure safety for emergency services and road users (Ministry of Transportation and Water Management Netherlands 1999).

Several measures are currently considered to improve IM practices, under the guidance of the so-called "smart objectives" for the application of IM measures to the Dutch road network. Situation awareness for IM is the ability to understand the status and consequences of an incident in support of decision making. Situation awareness is essential to reach almost any other objective of IM improvement. The main purpose of this project is to provide a full picture of the mobility consequences and area consequences of an incident in near real time to create situation awareness for IM actors. Situation awareness has multiple facets. The project focuses on situation awareness for: (1) mobility and how it is affected by an incident, (2) the area surrounding the incident and (3) the site accessibility. The lack of a real-time assessment of the mobility consequences of an incident as well as of its consequences on the surrounding area hampers the decision making ability to respond to an incident and to manage its consequences. The project intends to exploits anonymous data from mobile telecom operators to create a real-time situation awareness of incident consequences, specifically:

- To detect how far the consequences of an incident reverberate on the road network and on the other mobility modes;
- To anticipate on which other roads or transportation modes there will be congestion caused by an incident;
- To assess the accessibility to the incident site;
- To measure the risks for surrounding areas in case of incidents involving e.g. chemical releases.

More in detail, the project uses anonymised data of the KPN Mobile network. The data, represented by

Erlang measurements and SMS counts, are used by the carrier to manage network quality. In the study area over 1,200 cells were identified, grouped in 8 LACs. It involves the city of Amsterdam and its surroundings, for an area of about 1,000 km^2.

The first research goal in this project was how telecom data can be utilized for understanding presence and mobility in regular situations and during events where entire regulated flows of people are disrupted by an incident or an exceptional occasion like a football match, a music concert, a large celebration, serious traffic jams or a demonstration. This outcome could then be used to understand how a city or a mobility system can be measured, simulated and actuated to improve the quality of services provided to inhabitants (Vaccari et al. 2009). A first step to these goals is to create so-called normality maps (weekday–weekend and day–night patterns) over a longer period of time to be able to automatically detect anomalies.

The data have firstly been processed to generate different visualizations of the urban dynamics of Amsterdam. The primary features of the data are the weekday–weekend and the day–night pattern which affect all data. The weekday–weekend pattern is more or less pronounced depending on the area itself and appears to follow a rather predictable activity pattern within a certain range of variations. These patterns are in part the result of presence of people in a certain area and of people's mobility, but also of callers' behaviour, ranging from the obvious lower network traffic during the night to subtler behavioural caller changes that depend on the callers' context.

Figure 2 shows the effect of events on the network traffic. Around Queen's day (30 April), a major city gathering, the network activity peaks in certain areas such as the Rembrandtplein (the blue line), where street parties and celebrations take place, while it subsides in areas such as the World Trade Center (WTC, the red line) which shows typical weekend behaviors.

A more detailed data analysis in the project Current City has been carried out for a selected number of areas that are characterized by different land-use patterns and known differences in terms of how people use the area. The definition of the areas was based on the indications of the best serving coverage map overlapped to land use. The weekly patterns can be seen from the graph in Fig. 3. The diagram shows for each day the average traffic (Erlang) over a period of 5 months (1 January–30 May 2008). The data are normalized on the averages for comparison. Most areas, with the exception of Rembrandtplein and Arena, show a week-weekend pattern. Rembrandtplein does not respect the same pattern, and has a stable-increasing traffic during the weekends. The Arena, the area around the Ajax stadium, has a peak of activity on Sundays during soccer games.

The project Current City has presented the use of telecom data for the analysis of spatial network activity patterns based on a 1-h interval. Next steps in the project are a reduction to a 15 min time interval, a more detailed analysis of data validation and an improvement of visualizations. At the same time, some applications for crowd management, evacuation support for disaster management, incident management based on network activity patterns and traffic management for the inner city of Amsterdam where there are no detection loops will be developed.

Main research issues

Lessons

Road traffic analysis and prediction are two of the most attractive areas of use for mobile network data. Steady growing traffic volumes have led to enormous congestion and mobility problems, especially during the rush hours, both in urban areas and the highway networks.

While traditional measuring methods, such as road loop detectors, camera detection or floating probe vehicles, are effective and precise, there are practical and financial limitations to their use. Detection loops installed under the road pavement are regularly installed on highways but their application in urban environments appears as unfeasible given the number of roads that need to be monitored and the complexity of installation. Similar concerns can be raised for detection cameras, which are a feasible option for a limited number of measurement points. There is, however, an increasing need for less expensive monitoring systems and effective and reliable information systems.

It is not surprising therefore, that there is a growing interest in data derived from cellular

networks to support the traffic parameters estimation without requiring expensive and complex installations of ad-hoc measurement systems.

Looking at the results of the previous section, the first evidence that can be pointed out is that all projects so far are independently carried out, lacking any kind of cohesion among each other. Most studies are from telecommunications or electronics researchers, not from transportation researchers, and sometimes there are ambiguities in the definition of the traffic parameters to be obtained. More or less each of them proposes a different method to obtain a traffic parameter, given a mobile phone parameter. This means that the *fil rouge* mentioned in the introduction unfortunately has not been individuated.

However, all authors of the main reviews and applications in the field agree in considering that the following main issues affect any kind of study that would imply the estimation of traffic parameters from mobile phone data: issues regarding sample size and reliability, privacy, the role of private companies, and the role of transportation agencies (Caceres et al. 2008; Rose 2006). Usually these aspects are considered separately in the literature, but actually they are strictly tied one with another.

Sample size reliability and accuracy

The possibility to exploit huge amounts of data from each person who carries a mobile phone in his/her pocket seems to solve the problem of small sample sizes, or at least it appears that having a sufficient sample size has a very competitive cost effect compared to expensive loop detectors field tests or camera surveys.

However, it is not unusual that having lots of data could result in an indiscriminate use of them, regardless of their quality or of their peculiar meaning. According to the reviewed literature, there are different aspects to be clarified in order to identify the factors on which the right sample size depends, and they all relate to the moment of the data collection, or at least, to the modality of obtaining this data.

First of all, the survey method or technology may influence the composition of the sample, which may be constituted by on-call phones only or by idle phones as well. Of course, having one or the other case drastically changes the size of the sample. The use of a sample of only on-call mobiles would guarantee higher accuracy, due to the stronger signal that the network receives from an active phone.

The survey area also has an impact on the sample size: if the data is collected on a motorway stretch, it is more likely that all the mobile phones surveyed are those inside the vehicle, which is not true for surveys carried out on streets in densely urbanized areas. Reliability of the sample is also related with the possibility to exclude from the survey the mobile phones carried by people that are not inside the vehicles, but simply walking, or travelling by bike, or by public transport, or inside a building.

Another issue that affects sample size and its reliability regards the difference between data coming from the real GSM network, without informing the subscribers using their ordinary mobile phones (e.g., data used in RealTime Project by MIT), or data coming from ad-hoc surveys, in which mobile users are informed and perfectly aware that they are being observed and agree in being tracked (e.g., the Mobile Century Project). In the first case, data are collected at a GSM network level, and therefore they cover big portions of the transportation network as well. However, data should be anonymised, and hence it is not possible to have any kind of control on them: this is the reason why so far this kind of data have been used only to obtain information about the behavior of aggregated groups of people and to study urban density and activity patterns, not to retrieve detailed traffic information. In the second case, the sample may be more reliable and useful to obtain traffic information, but of course it is smaller.

Finally, it is also possible that the cell phone in a car is used by a passenger. The presence of two, three passengers, each of them making a call with their own mobile, leads to uncertainty of counting the same car several times as the number of mobile phones that are inside it.

As can be seen in Table 2, the accuracy issue, related to the precision with which the location information and traffic parameters are provided, is not at all negligible. In the first place, accuracy is affected by the methodology used for the collection; of course, the coupling of mobile positioning systems and GPS methods would improve very much the accuracy of the location information; such is the case of the Mobile Millennium project, which benefits from the precision of GPS systems, which may,

however, not always be used for cost reasons. Therefore, it is important to create a balance between the accuracy needed for the application concerned and the costs to be afforded in order to achieve that level of precision. It is worth noting that, depending on the application carried out, each project and study, that involve the collection of data, cannot avoid to mention and justify the level of accuracy reached, as instead is the case in several projects.

Clarifications are undoubtedly needed about the techniques to use in order to post-process the acquired raw data and isolate only the usable ones.

Privacy

Besides technological and market developments, the adoption of wireless location technologies is influenced by security and privacy issues. In terms of privacy it is especially the tracking of people or goods transported which raises many privacy issues (Beinat et al. 2008). The use of mobile phone data from GSM network involves the cooperation of the carrier that provides them. This falls within the legal framework governed by regulations to protect the privacy of phone subscribers (Caceres et al. 2008). As defined by (Westin 1970) "Privacy is the claim of individuals, groups or institutions to determine when, how, and to what extent information about them is communicated to others, and the right to control information about oneself even after divulging it". In this definition a person's privacy corresponds to the control of that person's information. This issue has been widely discussed in the literature, and it is one of the main problems that could hinder the opportunity to fully exploit the potential of WLT.

Legislators have addressed personal information in various laws, which have implications to location and sensor services. To protect personal data from an economic perspective, extensive attention is paid in the European law, in general and more specifically for use in electronic communications. Article 7 of the Charter of Fundamental Rights of the European Union (2000/C364/01) focuses on some general issues on the respect for private and family life: 'Everyone has the right to respect for his or her private and family life, home and communications.' Directive 95/46/EC provides the legal framework for the protection of individuals with regard to the processing of personal data (European Commission

1995), while Directive 2002/58/EC addresses location privacy specifically, stating that location data can be processed only after being anonymised or after having gained the consent of the user, why should be perfectly informed of the use that will be made of their personal data (European Commission 2002b).

A way to solve this problem could be for telecom carriers to adopt an 'opt in' policy, for which users have to explicitly agree if their mobile phone may serve as a probe or must be excluded from the monitoring. The perception of users is that mobile phone data are extremely related to their private life, and a diffuse mistrust is spread among users who of course would avoid giving the permission to handle such private data (Ahas et al. 2008), thus not allowing mobile carriers to release detailed data.

For all these reasons, coming both from common sense and from legal acts, the phone location data should be received and handled in an aggregate and anonymous manner in accordance with current regulation like any other kind of information taken from the cellular network. In this way the use off cell phone data does not break the law on private data protection, as anonymous data does not associate information with specific users. Technological workaround to use anyway mobile phone data from individual users are still in an experimental stage (see e.g. Herrera et al. 2010), but this is an issue still under research.

The role of private mobile companies

In order to exploit the advantages with respect to traditional survey methods, Wireless Location Technology has to be carried out in agreement with private mobile carriers, so as not to have to organize ad hoc surveys with a limited number of informed users, but using all the universe of subscribers. In this case, another question arises. How many mobile carriers are active in an area? Most of the projects found in the literature, which could make use of this kind of data, have agreements with only one mobile carrier. What about the rest of the population, which makes use of different mobile companies for their communications?

This issue is mostly taken into account only with "coefficients" that consider the market penetration of that particular mobile company. Therefore, sample size depends also on the willingness of the mobile carriers to make such data available.

Another aspect to be discussed is that most of the research carried out in this field is, indeed, a confidential matter of private companies, or restricted by agreements and patents. Maybe in the future, it will be possible to make use of this information, and new horizons will open for researchers, such as it happened when the military "Selective Availability" of GPS signal was ended by US in 2000, and a big amount of highly-accurate and reliable location data became available to civil institutions.

The role of transportation agencies

Governments and public authorities play an important role in stimulating both the development and implementation of wireless location technology to support traffic management, or to support their demand. There are a number of issues that need to be addressed like regulation on privacy, road safety, data ownership, performance requirements, interoperability, market structure and general economic services. In Fontaine et al. (2007) it is argued that transportation agencies have historically not defined suitable performance requirements for wireless location systems. Many deployments have lacked a well developed independent evaluation that quantitatively assessed the system performance. As a result, most projects were developed as a 'technology push' rather than technology which support the demand side. The symbiosis of business needs and IT capabilities creates the potential for a surveillance infrastructure, namely dataveillance. "Dataveillance is the systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more persons" (Clarke 1988). Dataveillance is a key concern in the adoption of location and sensor services where the government acts as the data collection hub. While regulations already provide a strict framework that, on paper, provides a high level of protection for individuals, this does not eliminate the concern that data collected for a legitimate traffic management use may eventually find other applications, either in the future or under different public order and safety circumstances.

On the other hand, transport agencies need to balance between a broad range of issues for creating the good conditions to stimulate the market to develop a new technology. These include the individuation of suitable performance requirements for

wireless location systems so that validation studies can base their efforts on these target values. For transport agencies it is useful to collaborate in the early stage of promising research and development projects to understand the possibilities and limitation of the technology.

Conclusions

In this paper a broad overview of the present state of the art of the research in the field of the use of data from GSM networks for the estimation of traffic parameters has been provided. Although not going into the analytical details of how data are extracted from the cellular network, and how traffic parameters are estimated from cell-phone parameters, an articulated discussion of the main issues involved in this field of research has been given, raising many research questions, partly derived from the literature, but not yet or only marginally addressed, and partly coming from personal considerations by the authors.

Since the GSM network was commercially launched in 1991, there have been indeed many studies and field tests carried out during the last 15 years with the original start of the CAPITAL project in 1994. The literature can be subdivided into two types of references: individual research groups that have prepared ad-hoc surveys for testing their own data processing and estimations, and big projects with the use of extensive datasets of cell phone data ad-hoc surveyed or coming from agreements with telecom operators.

The following general conclusions can be drawn:

- Travel speed and travel time are the most studied estimation issues for traffic management purposes;
- Projects are often initiated by technology providers, telecom operators and transport agencies. Validation studies are mostly carried out by research institutions;
- The adoption of GSM data is still limited and it is a field still largely dominated by research and development. Technology is promising but not yet developed to the degree necessary for large scale utilization;
- Most of the studies focus on stretches of roads, or loops, and not on a road network level;

- Recent studies show more promising results; however transportation agencies have historically not defined suitable performance requirements for wireless location systems, which may cause ambiguities in validation studies to draw clear conclusions.
- Active systems, like GPS-equipped phones used in the Mobile Millenium project, where thousands of users agree to place these phones in their vehicles in order to transmit positioning data and receive free live traffic information, look very promising.
- Extraction of telecom network data for the analysis of the spatial network activity patterns used in the projects Real-time Rome and Current City Amsterdam opens new possibilities in using such aggregated data for traffic management.

Data from cellular phones undoubtedly open new and important developments in transportation engineering but this requires a careful analysis. Hence, there are a number of steps needed to achieve a significant confidence in the use that can be made of these data. First of all, indeed, data should be validated. For traffic management related activities, this validation can be made by comparing data obtained from cell phones with data collected with other "on-site" systems like for example video-cameras. Obviously this should be done in different road conditions (freeways, arterials, urban roads) in order to understand the possible range of applicability. Another way to validate data collected could be to compare the estimated density of people with census data during periods with a higher probability to have people at home (for example, the early evening or the Sunday afternoon, depending on the social context).

This paper is the first step of a study whose aim is to further investigate data deriving from the project Current City Amsterdam; the next phases of the research will include the development of a validation methodology of this data using loop detector data as ground truth, and the study of new applications in the field of traffic management. Especially for contingency management (e.g. traffic accidents, network disturbances caused by terror attacks or nature catastrophes) the use of cellular phone data may be of strategic importance in the future.

References

Ahas, R., Mark, U., Järv, O., & Nuga, M. (2006). Mobile positioning in sustainability studies: Social positioning method in studying commuter's activity spaces in Tallinn. In U. Mander, C. A. Brebbia, & E. Tiezzi (Eds.), *The sustainable city IV. Urban regeneration and sustainability* (pp. 127–135). Southampton, Boston: WIT Press.

Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2007). Mobile positioning data in tourism studies and monitoring: Case study in Tartu, Estonia. In: M. Sigala, L. Mich, & J. Murphy (Eds.), *Springer computer science: information and communication technologies in tourism* (pp. 119–128).

Ahas, R., Saluveer, E., Tiru, M., & Silm, S. (2008). Mobile positioning based tourism monitoring system: Positium barometer. In P. O'Connor, W. Höpken, & U. Gretzel (Eds.), *Springer computer science: Information and communication technologies in tourism* (pp. 475–485). Berlin: Springer.

Amin, S., Andrews, S., Apte, S., Arnold, J., Ban, J., Benko, M., Bayen, A.M., Chiou, B., Claudel, C., et al. (2008). Mobile century—using GPS mobile phones as traffic sensors: A field experiment. In *World congress on ITS* (pp. 16–20). http://www.its.berkeley.edu/newsbits/winter2008/Mobile_Century.pdf. Accessed July 29, 2009.

Bar-Gera, H. (2007). Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: A case study from Israel. *Transportation Research Part C: Emerging Technologies, 15*(6), 380–391.

Beinat, E., Biderman, A., Calabrese, F., Dal Fiore, F., Hawelka, B., Hermans, B., et al. (2008). *Analysis report real-time city Amsterdam calibration*. CurrentCity project.

Berenschot. (2008). *Eindrapport Kwalitatieve jaarmeting IM*, J.-H. Leopold en H. Doornbos, 18 December 2009.

Birle, C., & Wermuth, M. (2006). The traffic online project. In *Proceedings from 13th ITS world congress*, *special session: Cellular-based traffic data collection*. London, UK.

Bolla, R., & Davoli, F. (2000). Road traffic estimation from location tracking data in the mobile cellular network. In *Proceedings from wireless communications and networking conference* (Vol. 3, pp. 1107–1112). Chicago, USA.

Caceres, N., Wideberg, J. P., & Benitez, F. G. (2007). Deriving origin-destination data from a mobile phone network. *IET Intelligent Transport Systems, 1*(1), 15–26.

Caceres, N., Wideberg, J. P., & Benitez, F. G. (2008). Review of traffic data estimations extracted from cellular networks. *IET Intelligent Transport Systems, 2*(3), 179–192.

Calabrese, F., & Ratti, C. (2006). Real time Rome. *Networks and Communication Studies*, 20(3–4), 247–258.

Calabrese, F., Reades, J., & Ratti, C. (2010). Eigenplaces: Segmenting space through digital signatures. *Pervasive Computing, IEEE 9*(1), 78–84.

Cascetta, E. (2009). *Transportation system analysis: Models and applications* (2nd ed.). Berlin: Springer.

Charter of Fundemental Rights of the European Union (2000/C364/01).

Clarke, R. (1988). Information technology and dataveillance. *Communications of the ACM, 31*(5), 498–512.

Delcan Corporation. (2009). *Baltimore multimodal traveler information system/regional traffic probe deployment*. http://www.delcan.com/pdfs/PDS-Trans-Systems/PDS-Baltimore_TrafficProbe.pdf. Accessed July 29, 2009.

Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *PNAS, 9*(36), 15274–15278.

Emory University. (2009). *History and timeline of GSM*. http://www.emory.edu/BUSINESS/et/P98/gsm/history.html. Accessed July 29, 2009.

European Commission (1995). *EU Directive 1995/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data*. Official Journal of the European Communities L 281, 23/11/1995 P. 0031–0050.

European Commission. (2002a). *EU Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services (universal service directive)*. Official Journal of the European Communities of 7 March 2002, L. 108, 51.

European Commission. (2002b). *EU Directive 2002/58/EC on the processing of personal data and the protection of privacy in the electronic communications sector (directive on privacy and electronic communications)*. Official Journal of the European Communities L 201, 31/07/2002 P. 0037–0047.

Fontaine, M. D., Yakkala, A. P., & Smith, B. L. (2007). *Probe sampling strategies for traffic monitoring systems based on wireless location technology*. Final contract report vtrc 07-cr12, 2007. http://www.virginiadot.org/vtrc/main/online_reports/pdf/07-cr12.pdf. Accessed July 29, 2009.

Girardin, F., Blat, J., Calabrese, F., Dal Fiore, F., & Ratti, C. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive computing, 7*(4), 36–43.

GSM Association. (2009). *GSM world statistics*. http://www.gsmworld.com/newsroom/market-data/market_data_summary.htm. Accessed July 29, 2009.

Hansapalangkul, T., Keeratiwintakorn, P., & Pattara-Atikom, W. (2007). Detection and estimation of road congestion using cellular phones. In *Proceedings from 7th International conference on intelligent transport systems telecommunications* (pp. 143–146).

Herrera, J. C., Work, D. B., Herring, R., Ban, X. Jacobson, Q., & Bayen, A. M. (2010). Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies*. Corrected Proof (in Press) doi:10.1016/j.trc.2009.10.006.

Hongsakham, W., Pattara-Atikom, W., & Peachavanish, R. (2008). Estimating road traffic congestion from cellular handoff information using cell-based neural networks and K-means clustering. In *Proceedings from 5th conference on electronical engineering/electronics, computer, telecommunications and information technology* (Vol. 1, pp. 13–16).

Kummala, J. (2002). *Travel time service utilising mobile phones* (Vol. 55, p. 67). Finnish Road Administration Report, Helsinki.

Liu, H. X., Danczyk, A., Brewer, R., & Starr, R. (2008). *Evaluation of cell phone traffic data in Minnesota*. TRB 87th annual meeting compendium of papers CD-ROM.

Maerivoet, S., & Logghe, S. (2007). Validation of travel times based on cellular floating vehicle data. In *Proceedings from 6th European congress and exhibition on intelligent transport systems and services*. Aalborg, Denmark.

Ministry of Transportation and Water Management. (1999). *Policy rules incident management Rijkswaterstaat Netherlands*.

Ministry of Transportation and Water Management. (2008). *Bereikbaarheidsmonitor hoofdwegennet*, 7 July 2008, Rijkswaterstaat, Dienst Verkeer en Scheepvaart, Netherlands.

Papageorgiou, P. (1991). *Concise encyclopedia of traffic and transportation systems*. Oxford: Pergamon Press.

Pattara-Atikom, W., & Peachavanish, R. (2007). Estimating road traffic congestion from cell dwell time using neural network. In *Proceedings from telecommunications, 7th international conference on ITS* (pp. 1–6).

Pattara-Atikom, W., Peachavanish, R., & Luckana, R. (2007). Estimating road traffic congestion using cell dwell time with simple threshold and fuzzy logic techniques. In *Proceedings from IEEE intelligent transportation systems conference* (pp. 956–961).

PR Newswire. (2009). *Two billion GSM customers worldwide—3G Americas*. http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=109&STORY=/www/story/06-13-2006/0004379206&EDATE=. Accessed July 29, 2009.

Promnoi, S., Tangamchit, P., & Pattara-Atikom, W. (2008). Road traffic estimation based on position and velocity of a cellular phone. In *Proceedings from IEEE 8th international conference on ITS telecommunications* (pp. 108–111).

Punzo, V., Borzacchiello, M. T., Ciuffo, B. (2009). *Estimation of vehicle trajectories from observed discrete positions and Next-Generation Simulation Program (NGSIM) data*. In TRB 88° annual meeting. Washington DC, USA, 11–15 January 2009.

Ratti, C., Pulselli, R. M., Williams, S., & Frenchman, D. (2006). Mobile landscapes: Using location data from cellphones for urban analysis. *Environment and Planning B—Planning and Design, 33*(5), 727–748.

Ratti, C., Sevtsuk, A., Huang, S., & Pailer, R. (2007). Mobile landscapes: Graz in real time. In G. Gartner, W. Cartwright, & M. P. Peterson (Eds.), *Location based services and telecartography* (pp. 443–444). Berlin: Springer.

Reades, J., Calabrese, F., Sevtsuk, A., & Ratti, C. (2007). Cellular census: Explorations in urban data collection. *IEEE Pervasive Computing, 6*(3), 30–38.

Reades, J., Calabrese, F., & Ratti, C. (2009). Eigenplaces: Analyzing cities using the space-time structure of the mobile phone network. *Environment & Planning B, 36*(5), 824–836.

Rose, G. (2006). Mobile phones as traffic probes: Practices, prospects and issues. *Transport Reviews, 26*(3), 275–291.

Rutten, B., Vlist, M., & Wolff, P. (2004). GSM as a source for traffic information. In *Proceedings of European transport conference 2004*, Strasbourg, France. www.etcproceedings. org/paper/download/1262. Accessed July 29, 2009.

Smith, B. (2006). *Wireless location technology-based traffic monitoring demonstration and evaluation project—final evaluation report*. Smart Travel Laboratory Center for Transportation Studies University of Virginia, Blacksburg, VA.

Smith, B. L., Pack, M. L., Lovell, D. J., & Sermons, M. W. (2001). Transportation management applications of anonymous mobile call sampling. In *TRB 80th annual meeting compendium of papers CD-ROM*.

Sohn, K., & Kim, D. (2008). Dynamic origin-destination flow estimation using cellular communication system. *IEEE Transactions on Vehicular Technology, 57*(5), 2703–2713.

Spinak, A., Chiu, D., & Casalegno, F. (2009). *"Mobile century" traffic monitoring study (Berkeley, CA)—Connected Urban Development project*. http://www.connectedurban development.org/pdf/sust_ii/mobile.pdf. Accessed July 29, 2009.

Thiessenhusen, K. U., Schafer, R. P., & Lang, T. (2003). *Traffic data from cell phones: A comparison with loops and probe vehicle data*. Germany: Institute of Transport Research German Aerospace Center.

University of Maryland Transportation Studies Center. (1997). *Final evaluation report for the CAPITAL-ITS operational test and demonstration program*. University of Maryland, College Park: US.

University of Virginia Center for Transportation Studies. (2006). *Wireless Location technology-based traffic monitoring demonstration and evaluation project*. Charlottesville, VA.

US Department of Transportation—FHWA. (2008). *NGSIM—Next Generation SIMulation*. http://www.ngsim.fhwa.dot. gov/. Accessed July 29, 2009.

Vaccari, A., Dal Fiore, F., Beinat, E., Biderman, A., & Ratti, C. (2009). Current Amsterdam: studying social dynamics through mobile phones network Data. In *Imagining Amsterdam*. Amsterdam, Netherlands, 19–21 November 2009.

Virtanen, J. (2002). *Mobile Phones as probes in travel time monitoring*. Technical report, finnish road administration report, Helsinki, Finland.

Westin, A. (1970). *Privacy and freedom*. London: The Bodley Head Ltd.

White, J., & Wells, I. (2002). Extracting origin destination information from mobile phone data. In *Proceedings from 11th international conference on road transport information and control* (Vol. 486, pp. 30–34).

Ygnace, J.-L (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.

Ygnace, J.-L., & Drane, C. (2001). Cellular telecommunication and transportation convergence. In *IEEE intelligent transportation systems conference*. Oakland, August, 2001.

Yim, Y. B. Y., & Cayford, R. (2001). *Investigation of vehicles as probes using global positioning system and cellular phone tracking: field operational test*. Report UCB-ITS-PWP-2001-9. California PATH program, Institute of Transportation Studies, University of California, Berkeley.