The use of GSM data for transport safety management:
An exploratory review

Research Memorandum 2011-32

John Steenbruggen
Maria Teresa Borzacchiello
Peter Nijkamp
Henk Scholten
THE USE OF GSM DATA FOR TRANSPORT SAFETY MANAGEMENT: AN EXPLORATORY REVIEW

John Steenbruggen  
VU University  
Department of Spatial Economics  
De Boelelaan 1105, 1081 HV  
Amsterdam, The Netherlands  
Phone: +31 152757301  
john.steenbruggen@rws.nl

Maria Teresa Borzacchiello  
European Commission - DG Joint Research Centre  
Institute for Environment and Sustainability  
Spatial Data Infrastructure Unit  
Via Enrico Fermi, 2749 - T.P. 262  
I - 21027 Ispra (VA) Italy  
Phone: + 39 0332 78 5506  
e-mail: maria-teresa.borzacchiello@jrc.ec.europa.eu

Peter Nijkamp,  
VU University  
Department of Spatial Economics  
De Boelelaan 1105, 1081 HV  
Amsterdam, The Netherlands  
Phone: +31 205986090  
p.nijkamp@vu.nl

Henk Scholten  
VU University  
Department of Spatial Economics  
De Boelelaan 1105, 1081 HV  
Amsterdam, The Netherlands  
Phone: +31 205986125  
h.j.scholten@vu.nl

Abstract

The consequences of traffic accidents are not only limited to road travellers (congestion, delays), but may also affect the area surrounding the incident, for instance, in the case of incidents involving release of chemical substances. A lack of a real-time assessment of the mobility consequences of an incident as well as of its consequences for the surrounding area hampers the decision makers’ ability to respond effectively to an incident and to manage its consequences. The main purpose of this paper is to provide a full picture of the mobility and area consequences of an incident in near real time by creating situation awareness for Incident Management actors. In particular, the study focuses on situation awareness for (1) mobility and how it is affected by an incident, (2) for the area surrounding the incident, and (3) for the site accessibility. The paper addresses information obtained from location-based services and pays particular attention to GSM data obtained from cellular phone networks. The advantages and limitations of the process of retrieving location information and proper transportation parameters from cellular phone networks are examined.
network use are highlighted. After an overview of new opportunities of GSM data for incident and safety management, some illustrations of applications are briefly presented and discussed. The paper ends with concluding remarks.

**Keywords:** incident management, mobile phones, GSM network, situation awareness, context
1. INTRODUCTION

Traffic incident management is a complex undertaking. It may be defined as a planned and coordinated process to detect, respond to and remove traffic incidents as well as restore traffic capacity as safely and quickly as possible (US Federal Highway Administration, 2000). In the Netherlands, Incident Management (IM) is conceived of as 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 and control the damage (Dutch Ministry of Transportation and Water Management, 1999). Clearly, road networks are part of a country’s transport infrastructure and are therefore subject to general transport policies and measures.

Road traffic injuries in the European region are a major public health issue, and are responsible for about 127 thousand lives per year (World Health Organization, 2004). Next to this intolerably high number of lives lost, about 2.4 million people every year are injured in road traffic crashes. Over 1.2 million people die each year on the world’s roads, and between 20 and 50 million people suffer from non-fatal injuries. The WHO predicts that road traffic injuries will rise to become the fifth leading cause of death by 2030 (World Health Organization, 2009). Therefore, the topic of traffic incident management is one of the key concerns in transportation safety policy.

It is worth noting that there is some division within the literature concerning the definitions of ‘accident’ and ‘incident’. Road traffic accident can be defined as ‘the product of an unwelcome interaction between two or more moving objects, or a fixed and moving object’ (Whitelegg, 1987). An “incident” is defined as any non-recurring event that causes a reduction of roadway capacity or an abnormal increase in demand. Such events include traffic crashes, disabled vehicles, spilled cargo, highway maintenance and reconstruction projects, and special non-emergency events (e.g., football games, concerts, or any other event that significantly affects roadway operations) (US Federal Highway Administration, 2000).

Everywhere in the developed world, steadily growing traffic volumes and traffic intensity have since the early seventies led to enormous congestion and mobility problems, especially during the rush hours. Traffic jams can be structural (recurrent) or incidental (non-recurrent). Structural (recurrent) means that they regularly recur on the same roads at the same times because the traffic demand is greater than the available road capacity. Incidental (non-recurrent) traffic jams are the result of accidents, road works and major events (Marchesini and Weijermars, 2010). These can be predictable events such as in the case of road works and planned events or unpredictable accidents. On a yearly basis there are in the Netherlands about 100,000 incidents (Berenschot, 2009), varying from small accidents to major multi-vehicle incidents causing casualties and vast damages to the road and its supporting structures.

Besides its direct impacts in terms of property damage, injuries, fatalities and other road safety effects for road users, it is also relevant for the mobility (Knibbe et al. 2006, 2007). Incidents can quickly lead to congestion and associated travel delay, wasted fuel, increased pollutant emissions and higher risks of secondary incidents. They are an important cause of congestion and increase the total cost of traffic congestion. According to information available in The Netherlands, approximately 30 to 44 million hours have been lost in 1990 due to traffic congestion. Translated into monetary costs, is equivalent to a loss of 700 €million (Dutch Ministry of Transportation and Water Management, 2007). After the introduction of IM policy in the country, in 1999 the average time of incident handling was reduced by 25% (Gronmij, 2004, Berenschot, 2009). Between 1985 and 2004, approximately 10 to 13% of all traffic jams on Dutch roads have been the result of incidents such as crashes and vehicles shedding their loads (McKinsey & Company, 1995,
Knibbe et al. 2004, TNO, 2006). In 2008 incidents accounted for 21% of lost vehicle hours (Dutch Ministry of Transportation and Water Management, 2008), so they are among the most important causes of the presence of traffic jams.

IM calls for real-time use of accessible traffic information. There are many sources of traffic information, but in our study we will pay attention in particular to the potential offered by traffic information from telecom networks. The aim of this paper is therefore, to analyze the use of telecom data – especially GSM data - to improve the situational interface to aid incident response managers in a complex decision making process. Situation awareness is essential to reach almost any other objective of IM improvement. The main purpose of the study is therefore 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 or managers.

The present paper will thus offer an overview of the potential of electronic data use in traffic management, with particular attention for the applicability of GSM data for incident management. First, we will show the relevance of GSM data from the perspective of space-time geography. Then we will address the wealth of research possibilities offered by Location-Based Services (LBS), inter alia GPS and GSM data. We will pay particular attention to the information needs of incident management in the transportation sector, where we will focus on the relevance of situational awareness. We will next offer some empirical illustrations and will conclude the paper with some research perspectives.

2. ELECTRONIC FOOTPRINTS IN TRAFFIC MANAGEMENT

We will now introduce the IM policy issues of the present paper by a general discussion of electronic data. In recent years, we have seen a rapid acceptance and use of individually based space-time data in the transportation sector. In particular, the use of electronic data for traffic management, logistic operations and incident management is remarkable.

Spatio-temporal tracking and tracing data are not only useful for general traffic management purposes, but also for incident management (e.g., non-recurrent traffic jams as a result of an accident). Incident management is one of the main ingredients of a dynamic traffic management system that aims to provide efficient, smooth, sage, reliable and sustainable transportation flows. To this end, a traffic management centre uses normally various information sources, such as induction loops, cameras and other monitors (including human observers). To control traffic flows, the centre may use VMS (Variable Message Signs), speed limitation signs, ramp metering and other measures for effective incident management. It should be noted that traffic incidents – in all forms ranging from traffic jams to traffic accidents – lead to high socio-economic costs for society, especially if one also adds the costs of emergency services to this. Traffic management aims to offer a wide range of prevention and abatement measures in order to minimize the costs of traffic incidents by optimizing detection time, warning time for emergency services, travel and operation time for emergency vehicles, normalization and flow recovery time. Clearly, reliable data have to be used to organize effectively all support measures in case of incidents. In this context, geo-science technology may play an essential role. In addition to the use of GPS data, the awareness is growing that GSM technology may be another source of rich space-time information on individuals or objects (see also Hale, 1997; Papageorgiou, 1991; Punzo et al., 2009; Turoff et al., 2004).

The use of wireless location and cell phone data appears to offer a broad range of new opportunities to create sophisticated and less expensive applications for traffic management. There are indeed many advantages compared to other technologies, but there are still some important issues that need to be resolved, in particular factors that influence
accuracy, reliability and data quality as well as techniques used for validation. Another important issue is the effective use of electronic data. This concerns performance requirements of transportation agencies, privacy issues, road safety implications of mobile phone use, data ownership, business models and public-private partnership to deploy, test, improve and implement new technologies. Understanding the accuracy, reliability and timeliness of the required data is necessary for any data collection system to create value added services (Richardson et al., 1995; Hearn, 1995; Paterson et al., 1999). These quality indicators are directly related to the usefulness of the information derived and depend also on the way technologies are deployed. There are differences in data quality requirements associated with different end uses of data (Caceres et al., 2008). Data may well be collected in real time, but the timeliness with which those data are made available may place limitations on potential applications. For example, accuracy and timeliness are very relevant to determine whether the data support real-time information systems such as incident detection or near real-time applications to provide travel time information to road users (Rose, 2006).

The acceptance of electronic data for traffic and incident management is rapidly increasing, although many applications are still in an experimental stage. The majority of field tests have attempted to monitor freeways in urban areas. These sites were often selected, since there is a clear need to monitor traffic in congested urban areas and they have a dense network of point detectors that can be used to calibrate and validate the performance of wireless location technology monitoring systems. In addition, they tend to have more robust cell coverage, a higher traffic volume, simpler path estimations and higher frequencies of handovers (Fontaine and Smith, 2007). Relatively few applications have attempted to monitor arterial roadways; they face additional complexity such as more path possibilities, correct filtering devices in vehicles, and so forth.

Cellular phone carriers have the ability to collect and store aggregated location data on hundreds of millions of subscribers. This means that the size of cell phone samples is very high. However, obtaining a sufficient sample size of probe phones is important for reliable estimations of traffic data and for calibrating models based on aggregated measures. Rose (2006), for example, argues that probe fleet size has a direct implication for the sample size available. Different studies show that a probe fleet size between 3000 to 5000 vehicles could provide reasonable travel times in denser areas (Boyce et al., 1991; Longfoot, 1991; Srinivasan and Jovanis, 1996). Of course, reliable real-time spatio-temporal information data are essential for precise geographical positioning of people and objects. This will be put in the perspective of space-time geography in the next section.

3. SPACE-TIME GEOGRAPHY AND DIGITAL DATA

Since the 1970s there has been a plethora of statistical models applied to the understanding of road accidents. However these models have had a tendency to neglect the spatial patterns of road accidents (Anderson, 2006). Historically, it has been argued by Whitelegg (1987) that geographers have not paid enough attention to the geography of road traffic incidents. For many years, it was considered by police, local authorities and road engineers, that road engineering, road layout and vehicle manufacturing faults were the main causes for road incidents. However, it has become evident through increasing police awareness and related research that road incidents need to be seen with a broader geographical perspective (both spatial and temporal).

Whitelegg (1987), in his paper ‘The Geography of Road Accidents’ seeks to understand road traffic incidents with reference to the scale of analysis and the importance of focusing on neighbourhood and community scale for an answer to the reduction of incidents (see below). Whitelegg outlines then the strong links between road traffic incident
analysis and other geographical dimensions such as population density distribution and spatial designs of neighbourhoods.

The history of quantitative data analysis in transport geography spans now already several decades. The need for a more appropriate behavioural underpinning of spatial interaction models led in the 1980s to the emergence and popularity of discrete utility (or choice) models, in particular multinomial logit and probit models, later on followed by conjoint analysis modelling. Such individually based models were proven consistent with aggregate-oriented spatial interaction models and were widely accepted in the transport research community. They also turned out to be well suitable for actor-based policy simulation experiments, for instance, in the context of micro-simulation models and agent-based models.

All such models were widely used for prediction purposes, evaluation experiments and policy analyses in the planning and transportation science field, for example, to trace the system-wide effects of road pricing on the behaviour of car drivers. With the advent and introduction of ICT, the computing capacity in quantitative research showed a dramatic increase, so that also spatial dynamics could be captured in a statistically more satisfactory way. In recent years, complexity theory has, offered a remarkable contribution to a better understanding of the sensitivity of spatial systems’ evolution to endogenous non-linear space-time behaviour. Space-time dynamics (e.g. in the cellular automata domain) became an important ingredient of advanced transportation research and spatial analysis, and prompted a new departure, viz. the use of data mining methods for large data sets. The current use of computational neural networks and genetic algorithms demonstrates convincingly the great potential of more sophisticated data collection techniques. The real essence of space as highlighted in Tobler’s (1970) law (“all things in space are related to each other, but nearby things are more related than distant things”) was taken up in a new strand of literature addressing spatial – and spatio-temporal – autocorrelation, either as testing devises or as design mechanisms for spatial (dynamic) models (see also Tobler, 2004). Cellular automata, spatial filtering techniques and self-organized mapping procedures (‘Kohonen maps’) for spatial interaction analysis were a logical follow-up and complement to the above mentioned trends (see e.g. Arribas et al., 2010; Codd, 1968; Couclelis, 1997; Kohonen, 2000; Patuelli et al., 2010 and Kulkarni et al. 2002).

In recent years, we have witnessed an increasing popularity of location-based services (LBS) and data using various kinds of electronic identification systems, so that at an individual level (a traveller, a container, a truck, or a taxi) the geographic position of a unit can be traced with great precision. Many applications are available for purchase and for free to cell phone and other wireless device users. For example, Japanese parents are using location-based tracking devices to monitor the spatial movement of their kids. This new approach will certainly prompt many new applications in space-time geography.

An interesting source of individually based information on the space-time position and behaviour of persons is in principle available from mobile (or cell) phone data, derived from the GSM network. The penetration rate of mobile phones is rapidly reaching a full saturation level in most OECD countries, so that a system-wide coverage does in principle exit, almost in continuous space-time format. Such data – as very accurate representations of the individual space-time location – are in principle available with telephone operators. If such data – in anonymous form – could be made available to the research community, an unprecedented source of information on the space-time geography of individuals could be used in applied research (see for an overview Steenbruggen et al., 2010).

It is noteworthy that this idea of a continuous space-time map at an individual scale was already put forward by the late Swedish geographer Torsten Hägerstrand in 1967. He introduced the ‘space-time cylinder’ and its related time-space model to offer a description...
of both individual space-time patterns and the resulting spatial interactions if many individuals were ‘en route’ at the same time and place, a situation caused by the universal limited supply of daily time resources. His work was regarded as a new perspective in social-behavioural geography, as it highlighted so clearly the essence of interaction and congestion phenomena in space (see Pred, 1977). Three constraints appear to act as constraints on the daily mobility pattern of individuals, viz. capability constraints, coupling constraints and authority constraints. It also laid the foundation for activity-based transport geography, but, unfortunately, lack of data and the technology available to implement the framework precluded often a full operational application of his path-breaking ideas. Now with the potential availability of large-scale continuous space-time information bases on spatial movements of individuals, a really interesting novel approach might be developed, which may have great implications for spatial modelling. Two such approaches can be found in the literature. The first incorporates elements of cognition by considering individuals’ preferences via the theory of affordances proposed by Gibson (1979). Cognitive constraints, e.g. choice behaviour, were not given explicit attention in the original time-geography framework. These constraints can help personalize LBS, allowing the possibility to collect more detailed information about the choices individuals make. The second adjusts the space-time prism concept to support interactions and activities between the physical and virtual spaces (Yu and Shaw, 2008). This approach would help model and understand how, in the age of mobile computing, where a variety of activities and services can be carried out on the go, individuals are allocating their space and time resources.

4. BACKGROUND OF LOCATION BASED SERVICES

Location is historically recognized as a strategic asset of cellular network providers (Teckinay, 1998). One of the most powerful ways to personalize mobile services is based on location. The use of this information enables users to experience value-added services and the cellular network provider to offer differentiation and incremental profitability by increasing its subscriber base (Drane et al., 1998). The concept of ‘Location Based Services’ (LBS) has become increasingly important to create intelligent information service in a broad range of domains. LBS are information and entertainment services, accessible with mobile devices through the mobile network and utilizing the ability to make use of the geographical position of the mobile device. In 1991, the European Union established 112 as the universal emergency number for all its member states. E-112 is a location-enhanced version of 112. The telecom operator transmits the location information to the emergency centre. The EU Directive E-112 (2003) requires mobile phone networks to provide emergency services with whatever information they have about the location from where a mobile call was made. This directive is based on the FCC’s Enhanced 911 ruling in 2001. The E-112 (E-911) was the driving force to invest in location-based technology (Yuan and Zhang, 2003; Kumar, 2004).

Location is the most important element to create context-aware services. Definitions of context-aware computing date back to the early nineties, the period in which Olivetti developed several context-aware applications based on so called ‘active badges’ (Want and Hopper, 1992). Context plays a central role in human behaviour and activity, and in the way we use technology. Context has historically been ignored in computer science (see Broens, 2004). Computers were designed as much as possible as black boxes to enhance their abstracting power, overcome complexity and safeguard their reliability. An important reason that context awareness of computer devices has never been a major issue is that most interaction between computer and human has, for a long time, taken place in office-like situations with immobile desktop computers.
In the last fifteen years, we have also witnessed a clear trend towards embedding computers in a variety of devices. Computers are becoming smaller, lighter, cheaper and at the same time more powerful. We can find microprocessors in a wide range of devices, from mobile phones, handheld computers or cameras to electronic postcards, toys for children or home appliances. A standard car has dozens of microprocessors, while the premium car can have more than a hundred. We can also see embedded computing beyond the personal devices. Motion sensors, electronic tags or surveillance video cameras are all examples of equipment supporting collaborative work among people. The embedded computers, being part of our everyday, physical world, bring new requirements on Human-Computer Interaction (HCI). Research on this matter has been framed using names such as “ubiquitous computing” (Weiser, 1991), “pervasive computing” (Ark and Selker, 1999), “embodied interaction” (Dourish, 2001a, 2001b), “tangible interfaces” (Ishii and Ullmer 1997), and others. These research directions may contribute on slightly different topics, but they all address human-computer interaction beyond the traditional desktop environment, where computing is embedded into the fabric of the world around us.

Context-aware computing is often associated with the term ubiquitous or pervasive computing, coined by Mark Weiser (Weisser, 1996). The idea behind ubiquitous computing is that it provides people all kind of support and automatic services, anytime, anywhere, based on their personal needs and preferences and on their current context. It includes concepts of context-aware computing and all kinds of possible sensors to detect the relevant aspects of the user environment. Striving for a definition, Lyytinen and Yoo (2002) describe mobile and pervasive computing as conceptually different. While mobile computing is about “increasing our capability to physically move computing services with us”, pervasive computing is focused on the integration aspects. It is described as computers having the “capability to obtain information from the environment in which it is embedded and utilize it to dynamically build models of computing”. Ubiquitous computing is, according to Lyytinen and Yoo, the domain where mobile and pervasive computing meet. It means that “any computing device, while moving with us, can build incrementally dynamic models of its various environments and configure its services accordingly. From a system point of view, ubiquitous computing implies interconnected, communicating networks of numerous, casually accessible, often invisible or very tiny computing devices, either mobile or embedded in almost any type of object imaginable, including cars, tools, appliances, clothing and various consumer goods. A field with extensive literature related to ubiquitous computing and context-aware computing is ambient intelligence (see for instance Weber et al., 2005). Ambient intelligence is seen as a convergence of several computing areas including ubiquitous or pervasive computing, context awareness and intelligent systems research.

Obviously one of the great challenges within ubiquitous computing research is to understand the relation between computing and the context in which it is embedded. What is the HCI impact of continuously changing social, spatial, temporal or technical settings? An even greater research challenge (see Abowd and Mynatt, 2000) is to explore how computation can be made sensitive to the setting and thus provide appropriate, tailored services to the user. The theme is recognized as context-aware applications.

Context awareness and ubiquitous computing go hand in hand with the new paradigm of “access anytime, anywhere”. They entail a new research question on the merit of the activities carried out: does the type of context influence the type of tasks and, vice versa, do different tasks require different contexts? On this issue, Wiberg and Ljungberg (2001) propose the matrix shown in the following figure, to categorize work tasks along their geographical and temporal components. This matrix is based on the assumption that many tasks require a specific place and time to be done, so it decouples the concept of
ubiquitous computing from the paradigm of “access anytime, anywhere”. In addition, it justifies the need for shedding new light on which tasks are more or less place and time dependent. Vice versa, different locations and times may also determine the relevance of tasks, as well as the needs of those who carried them out. Time and place, in other words, are here considered as essential elements of classifying computing types, and form four main context classifications.

![Figure 1: Categorisation of tasks along place and time (source: Wiberg and Ljungberg, 2001)](image)

5. THE USE OF GSM NETWORK DATA IN SUPPORT OF INCIDENT MANAGEMENT

Transportation operations and public safety operations are intertwined in many respects. Public safety providers (by law enforcement), fire and rescue services, and emergency medical services ensure safe and reliable transportation operations by helping to prevent crashes and rescuing crash victims. Conversely, the transportation network enables access to emergency incidents and, increasingly, provides real-time information about roadway and traffic conditions.

The key obstacle to effective emergency response is the communication needed to access relevant data or expertise and piece together an accurate understandable picture of reality (Hale, 1997). In recent years data deriving from mobile phones has attracted the attention of researchers, in particular in the emergency field to improve the situational interface. Examples are:

- Location information from mobiles has been utilized to provide detailed maps of the density of people in different cities, such as Graz (Ratti et al., 2005, 2006, 2007) Real-time Rome (Reades et al, 2007) and Current City Amsterdam (Steenbruggen et al, 2010);
- Agent-Based Modelling simulations to provide emergency responders with timely information on the status of a city or region, as well as the capability to detect, follow and possibly predict crisis events (WIPER project) (Schoenharl et al., 2006a; Schoenharl et al., 2006b; Madey et al., 2006; Madey et al, 2007);
- Simulation system that models the evacuation base on autonomous intelligent agents which are used to represent various types of actors and study the effect of different disaster scenarios and agent behaviours (Filippoupolitis et al., 2008);
- Providing a visualization interface for Emergency Responders for detection, warning, response and mitigation based on cell phone data (CAVIAR project) (Vaidyanathan et al., 2008a; Vaidyanathan et al., 2008b)
- Telecom sensors used in the case of emergency evacuations (Inoue et al., 2008)
- Issuing and distribute early warnings using cell-broadcasting on GSM phones to citizens (Wood, 2005; Pries et al., 2006; Sillem et al., 2006).
- Integrated mobile information and communication system for emergency response operations (MIKoBOS project) (Meissner et al., 2006)

Traffic incident management is a much more limited domain than general emergency management; however, it can be seen as a special case of emergency response. To a certain extent, the examples above can also be applied to design a traffic incident management system based on telecom data. This will be the subject matter of the next section.

6. INCIDENT MANAGEMENT AND SAFETY ISSUES

Avoiding incidents is even a more important issue than to develop strategies and technology for efficient incident response. The emergence of an incident is determined by a combination of causes. In many cases, human or technical failure plays an important role, e.g. driver distraction, alcohol abuse or motor breakdown (Wegman, 2007). However, these factors are more likely to cause an incident if external conditions complicate the driver task like critical traffic, road or weather conditions. Therefore, preventing incident occurrence is possible for at least a fraction of all incidents for which critical traffic or weather conditions are dominant causes. Hence, strategies to prevent incidents are of course preferable, in terms of safety and mobility, to strategies designed to respond to incidents already occurred. Whenever an incident occurs on the freeway, this also has an effect on the safety of the people near the incident.

Models for incident prediction and forecasting seem to gain increasing importance in the literature, where may be found several attempts to get a better understanding of the nature and consequence of incidents and of how they affect the environment in terms of safety and mobility. For example, in Valenti et al. (2010), a comparative analysis between five models (Multiple Linear Regression, Prediction/decision tree, Artificial Neural Network, Support/Relevance Vector Machine, K-Nearest-Neighbour) to predict the incident duration is presented. The datasets used include 237 incident events, with information about the incident characteristics, the personnel and equipment involved to clear the incident, response times, beginning and ending time of the incident. The authors present a classification of the recorded information on the incidents as well, distinguishing among: (i) incident details (ii) operational details (iii) infrastructure and traffic variables. After accounting for account of the state of the art on incident duration prediction, presenting an interesting diagram on the phases that is possible to recognize since the moment that an incident occurs to the instant in which the traffic flow is restored at its normal conditions, they perform statistical analysis to select the explanatory variables of the model. Good prediction rates with duration less than 90 minutes were found. They propose moreover an incident classification scheme in order to select the most appropriate prediction model. Of course, quality of input data is crucial for the good predictability of the model.

Ji et al. (2009) also present a study on incident duration, in particular on the portion of incident duration constituted by “recovery time”, using a “cell transmission model” (Daganzo, 1994). The diagram on incident duration by Valenti et al. (2010) is clearer, but it effectively does not take into account the “recovery time” which instead is the focus for
the case study on which the proposed prediction model has been analysed made use of data from a stretch of an elevated freeway of Shangai, and the model’s results perform rather well, according to the authors.

Most of the research efforts in the field of incident management have been recorded in the USA from the beginning of the nineties, as witnessed by a series of early reports by Cambridge Systematics (Cambridge Systematics, 1990) and the Federal Highway Administration (US Federal Highway Administration, 2000), and as well explained in Ozbay et al. (2009). The later studies presents a good definition of incident management, as “a combination of policies and strategies that effectively coordinates the available resources to reduce incident durations”. They classify incident management operations in “network related”, concerning the operations to be performed to prepare all the units to the possibility of an incident occurrence, and “incident related”, concerning the actions and the means to be deployed during an incident. In particular, the paper describe RIMS, a software aimed at evaluating the benefits and the costs of the deployment of different incident management strategies and technologies, using a traffic simulation model based on Daganzo’s cell transmission model (Daganzo, 1994). The incident management process, according to the authors, who have a long experience in the incident management field (see Ozbay and Kachroo, 1999) is subdivided in different activities:

1. **Incident detection**: identification of the incident, by means of different technologies;
2. **Incident verification**, “usually completed with the arrival of the first respondent on the scene”
3. **Incident response**, corresponds to the deployment of the planned strategy to deal with the occurred incident. It is carried out by means of several forms of technologies
4. **Incident clearance**: the activity of removing from the road all the residual things disturbing the normal evolution of the traffic flow

The diagram presented, describing these phases and the incident time phases as well, is similar to the other ones found in the previous references, and includes the “recovery phase” (from the effective clearance of the road to the restoration of the normal traffic flow). The paper aims at testing the RIMS simulation tool by changing different variables and using field data coming from the Southern New Jersey highway network database. In particular, interesting for our purposes is the use of different scenario to see which the effects are, on incident detection time, of changing the percentage of cellular phone owners, the threshold number of cellular phone calls fixing the traffic flow (i.e. the number of vehicles per hour). Some cost-benefit considerations are also provided, including actual investment and annual costs and time savings, showing that high benefit-cost ratio may be achieved.

In the US, a National Incident Management System (NIMS\(^1\)) is also present, whose aim is to provide a collaborative framework among all the stakeholders involved in the incident management process in order not only to prevent and adequately react to incidents but also to mitigate their effects, on both people and the environment.

There are still few examples in the academic literature that provide the use of wireless positioning technology in support of incident management, one of them is Wen et al. (2008). Their paper presents an incident detection method, which uses GPS and CPS (Cellular Phone Positioning System) technology to gather data and then sending the information to the GSM network, from which the Traffic Management Centre can take it in order to locate the incident. A simulation model has been used to validate this detection method. The authors first present a review and first evaluation of existing incident management tools.

---

\(^1\) [http://www.fema.gov/emergency/nims/]
detection methods, subdivided in manual detection techniques and automatic incident detection methods, the latter classified in direct and indirect ones, if abnormal changes in traffic flow are recorded using image processing or traffic flow parameters monitoring. The incident detection system proposed by the authors is based on three components: the Traffic Management Centre (TMC), the in-car terminal and the GSM network. The TMC is described as the core of the incident management system: its role is to receive the information on the incident occurrence and to send a proper and timely response. In the system imagined by the authors, the in-car terminal is supported by a GPS, a GSM and eventually a Child Passengers Safety (CPS) module. Thanks to the GSM module, it is possible to exchange information with the Traffic Management Centre. The initiative is given to the driver who is supposed to send a message to the TMC using the in-car terminal system, if it happens that he is involved in an incident or in an incident-like situation. The incident detection algorithm presented is based on the two possible scenarios of single or multiple incident report. In the latter case, of course, the incident verification is facilitated.

As the authors recognize that there are no examples in the literature of traffic incident detection methods based on wireless positioning technology, they use a simulation model to test the incident detection method proposed, using as main explanatory variables the proportion of vehicles with in-car terminals, the severity of the incident, and the minimum number of reports.

Other earlier and interesting examples of incident management systems may be found in Huang et al. (2007) where a tool for incident response management in which the traffic simulation and optimization model runs on a GIS platform is described. Moreover, in Dia et al. (2008), it is recognized that incident management strategies have not a neutral impact on the traffic flow, but they have influence on it, and in particular, they are network dependent, and have the possibility to provide benefits in terms of travel times and safety conditions. Even this work is based on a simulation model, using real traffic data, testing different incident management strategies with different scenarios.

Looking closely at Europe, there are many European initiatives to use new in-car technology and cooperative systems to increase safety for road users. A good example, which is directly related to the Incident Management situation interface is the project “e-Call”, which stands for ‘Pan-European in-vehicle emergency call system’. e-Call gives the precise coordinates of an accident’s location to the emergency services, which are responsible for the follow-up assistance. eCall is part of the eSafety² initiative led by the European Commission. According to an analysis conducted by the European Commission supported by the wider European project E-MERGE, eCall is supposed to allow for a reduction in accident response time of about 50% in rural areas and up to 40% in urban areas. As a consequence, the time period between the occurrence of an accident and its clearance is reduced. eCall also aims to decrease congestion by 15%, which will lead to a reduction in fuel consumption and thus less harm on the environment. Another good example, although not very recent (2000-2002), is the European project PRIME³ (Prediction of congestion and incidents in Real time, for intelligent Incident Management and Emergency traffic management). The objectives of PRIME were to develop innovative methods to increase the effectiveness of incident management policies and increase road safety. These include Estimation of Incident Probability, Incident Detection, Incident Verification and Integrated Incident Management Strategies.

Victims of the primary incident, emergency workers, road users (upstream of the incident and on the other side of the road) and people living near the freeway around the

² http://www.esafetysupport.org/en/welcome.htm
³ http://www.trg.soton.ac.uk/prime/index.htm
incident are the most important risk groups who are exposed to additional risks when an incident occurs (Knibbe et al., 2006). By the term “risk” we mean the risk of being in a secondary incident.

Incidents have a direct effect on the safety of a number of the mentioned risk groups. These risk groups are used to determine the “group specific” effect and the overall effect of an incident on safety, as well as the impact of IM measures on safety. The victims of the primary incident and the emergency workers are staying in the middle of the road during incidents, which increases the probability of a secondary crash. Road users on both sides of the road often have to deal with congestion around incidents (Knibbe and Wismans, 2007). The difference in speed, at the interface between the tail of the traffic jam and uncongested traffic, is a factor for the probability of a secondary incident (Aarts, 2004).

7. ANALYSIS TOOLS

Mobile telecommunication networks are complex distributed systems producing huge amounts of data continuously. The data are collected into a database and used by network monitoring and management tools. The data include, for example, performance measurements from the radio interface and log data from application servers. A daily data set from an operational GSM network may consist of several gigabytes of data.

A major issue for emergency response managers is the problem of “information overload”. Studies have shown a correlation between abundant available data and bad decision making in crisis situations (Smart et al, 1977). Good emergency response systems provide access to the large amount of available data in such a way that the emergency response manager can use the data effectively to reach good decisions (Belardo, 1984; Jennex, 2007). Situational awareness, that is mandatory for the successful monitoring and decision-making in many scenarios, is one of the founding characteristics of intelligent software agents. An ‘agent’ can be defined as “a computer system that is capable of flexible autonomous action in dynamic, unpredictable, typically multi-agent domains” (Luck et al., 2005). A characterizing feature of agents is situatedness: the agent receives sensory input from its environment and it can perform actions which change it in some way (Kowalski, 1999). When combined with reactivity, situatedness may lead to the early detection of anomalies and to the formulation of a suitable plan for solving them. In (Florez-Larrahondo, 2006), the integration of intelligent anomaly agents and traditional monitoring systems for high-performance distributed systems is discussed.

In this paragraph, we explore some theoretical concepts and tools from literature and similar projects (e.g. Real-time Rome, WIPER, CAVIAR) to examine if telecom data could be a useful source to improve the Incident Management process. We will focus on four different research directions, which are strongly related to situational awareness.

7.1. Incident detection

There currently exist several ways to detect an incident: road users involved who calls emergency number, alerts from human observers from emergency organisations like police, road-inspectors and employers from towing services. Technical detection is also possible by (automatic) camera detection and induction loops or by means of new ways of detection, like the above-mentioned European e-call project. Quick incident detection is crucial for mobility consequences in terms of reducing traffic jams and vehicle lost hours.

An interesting challenge is to analyse whether telecom data could be useful for incident detection. Changes or anomalies in (the use of) the telecom network could be potential indicators. Models of anomaly detection were first used for the development of intrusion detection systems in a telecom network. Detecting anomalous behaviour is also one of the main tasks in telecom network operation (Kumpulainen and Hätönen, 2008).
Anomalies may result from faults, misbehaviour or unauthorized intrusion. It is essential to detect such situations as soon as possible. In the existing literature, there is still no consensus on the terminology, when it comes to the classification of various types of intrusion and anomaly detection systems. Debar et al. (1999a) have provided one of the most cited intrusion detection taxonomies up to date. Revised versions of this classification and categories can be found in (Debar et al., 1999b; Debar et al., 2000; Axelsson, 2000; Arvidson and Carl bark, 2003, and Burbeck, 2006).

There are two complementary trends in intrusion detection (1) the search for evidence of attack, based on knowledge accumulated from known attacks and (2) the search for deviations from a model of unusual behaviour, based on observations of a system during a known normal state. The first trend is often referred to as misuse detection (Jagannathan et al. 1993, Kumar and Spafford, 1994) or detection by appearance (Spirakis et al., 1994). The second trend is referred to as anomaly detection (Jagannathan et al., 1993) or detection by behaviour (Spirakis et al., 1994).

The problem of anomaly detection in data networks has been extensively studied. Anomaly detection consists of identifying patterns that deviate from the normal traffic behaviour, so it is closely related to traffic modelling of data in the mobile network. The anomaly detection literature treats the detection of different kinds of anomalous behaviours: network failures (Hood et al., 1997; Katzela and Schwarz, 1995; Ward et al., 1998), flash crowd events (Jung et al., 2002; Xie et al., 2008) and network attacks (Cheng et al. 2002, Zou et al., 2005, Wang et al., 2002; Lakhina et al., 2005; Tartakovsky et al., 2006).

There is abundant literature on the anomaly detection problem, which describes a variety of approaches, including statistical, neural network, and machine learning methods. In statistical-based techniques, the network traffic activity is captured and a profile representing its stochastic behaviour is created. This profile is based on metrics such as the traffic rate, the number of packets for each protocol, the rate of connections, the number of different IP addresses, etc. Two datasets of network traffic are considered during the anomaly detection process: one corresponds to the currently observed profile over time, and the other is for the previously trained statistical profile. As the network events occur, the current profile is determined and an anomaly score estimated by comparison of the two behaviours. The score normally indicates the degree of irregularity for a specific event, such that the intrusion detection system will flag the occurrence of an anomaly when the score surpasses a certain threshold.

The main challenge in developing anomaly detection algorithms, in the field of incident management, is that the “ground truth” about the real traffic conditions is generally unknown. The idea is that abrupt changes in some network signalling events might be the symptom of a road anomaly (accident/congestion), for example: (a) drop in the handover rate; (b) abrupt change in the Location Register update; (c) increase in the number of calls/SMS; (d) drastic change in the number of road users. 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 Home Location Register and Visit Register Location network databases. The most basic indicator of anomalous behaviour in a cell phone network is an increase or a decrease in cell phone call activity within a given geographical area. This type of anomaly can be detected by monitoring a time series consisting of the number of calls made in disjoint time intervals of a fixed size, e.g. the number of calls made every 15 minutes.
A nice example of the use of statistical techniques with telecom data can be found in the project Real time Rome where they analyse spatiotemporal signatures based on k-means clustering technique (see Reades et al. 2007). These spatiotemporal signatures can be the basis approach for anomaly detection. Clustering allows the usage of local thresholds, taking the local variance of the data into account. There are several ways to use clustering for anomaly detecting (Tan et al., 2005). In a basic clustered distance based method, the data are clustered and the distances to the nearest cluster centroids are calculated for each data sample. Samples that are very far away from the centroids are considered anomalies. The threshold can be for example the 95% percentile of the distances, thus assuming that 5% of the data are anomalous.

Clustering is an appealing non-parametric method because it allows capturing various classes of “normal” and “abnormal” behaviour. This may be quite useful since, in addition to detecting anomalies caused by events that have never been seen before, knowing various types of “abnormality” would allow us to identify interesting events that have already been seen. The goal of clustering is to group similar data items together. There are three major types of clustering algorithms: partitional, hierarchical, and incremental (Jennex, 2003). A few methods have been developed for clustering data streams (Guha et al., 2003; Aggarwal et al., 2003). Stream clustering algorithms are similar to incremental algorithms. Hybrid clustering combines two clustering algorithms (Cheu et al., 2004; Chipman and Tibshiran, 2006; Surdeanu et al., 2005). Partitional and hierarchical clustering algorithms may also incrementally incorporate new data into the cluster model (Jain et al., 1999).

A nice example of anomaly detection for streaming data to detect crisis events can be found in the project WIPER (see Schoenharl et al., 2006a, Schoenharl et al., 2006b, Madey et al. 2006, Madey et al., 2007). The techniques they analyzed for detecting anomalous patterns of spatial activity are a hybrid clustering algorithm that combines k-means clustering and statistical process control (Pawling et al., 2006) and the Markov-Modulated Poisson Processes (Yan et al., 2007).

The Markov modulated Poisson process, which uses a Poisson process in conjunction with a hidden Markov model to identify anomalies in the data, is described by Ihler et al. (2006, 2007). A Poisson process models the number of random events that occur during a sequence of time intervals, and can be used to model the baseline behaviour of such a time series.

An important issue is the timeline in which the telecom network provides data to the Current City application. At this moment we use a time interval of 15 minutes. Interesting research questions are which type of incidents can be detected by an anomaly, Next to this its interesting to see how other types of anomalies (e.g. big events) can be filtered out to be sure we deal with traffic incident anomaly detection. It is also relevant to see which anomaly detection method and which GSM telecom data is the much more suitable for incident detection.

### 7.2. Safety risks for surrounding areas

Present societies are vulnerable to natural and industrial disasters as well as daily traffic accidents. Major incidents (e.g. with dangerous substance) on freeways have a direct effect on the safety of the people passing close to the incident location. Currently used risk maps\(^4\) give a static overview of the potential risks for specific areas.

Responding to incidents that involve chemical or dangerous substances requires a high level of preparedness and the precise knowledge of how many people could be

---

\(^4\) [http://www.risicokaart.nl/](http://www.risicokaart.nl/)
exposed in the incident area. Risk maps give a static picture of the risks, based on information from census or other regular surveys. Emergency operations need more than this. The current number of people affected, rather than the expected number of people affected, dictates the extent, the size and the organization of the rescue operations. Interesting research question are:

- What if we knew the number of people in the incident area and its surroundings at the time of the incident?
- Imagine we knew how many people are affected by an incident and how many are in the surrounding area. How could this affect the organization of the emergency and of the support infrastructures such as hospitals?
- How many lives could be saved by a more effective directed intervention?

**Figure 2: How many people are at the (incident) site** (Source: Currentcity.org, 2010)

Real time forecast models for the prediction of potential risks in case of major incidents with dangerous goods for emergency services, road users and people living near the freeway are currently not available. A forecast of the combination of Incident hotspots and the real-time presence of people based on normality maps (spread over day – night, week – weekend pattern) could give useful insights on which areas are affected by the highest potential risks and which people could be potential affected.

Near real-time counts of people represent the number of people in a certain area at a certain time. The areas can be entire urban agglomerate, a specified zone or an arbitrarily defined grid element.

The research questions outlined above appear very simple but underline an extremely complex modelling exercise. The scope of this exercise is to formulate which data can be realistically derived from GSM networks, which cellular network data at which spatial resolution and with which level of confidence may be taken into account. Calibration is the process of creating and validating models to derive useful data from the telecom network input.

At a very high level, this can be described as the process of modelling and fitting estimates of the desired output variables based on (1) a certain number of real-time inputs (the telecom data stream); (2) static and dynamic spatial data (land use, road network); (3) telecom users’ data (caller profile, average SMS per day, etc.) and (4) ground truth (e.g. counts of people or cars). The following interesting research questions can be therefore raised:

- how much people are in a specific area (which can be defined by a circle with radius x) at a certain time (i.e. day – night, week – weekend, at a certain time with intervals of 15 minutes)?
• what will be the best location area to define ground truth for calculating presence of people (i.e. street level, telecom cell id…)?
• which telecom data gives best estimates for the presence of people?
• which known census data can be used for calibration to derive the presence of people?

7.3. Prediction of flows and site accessibility of emergency services
Traffic counts measures the amount of vehicles that travel past a certain point on the roadway (Caceres et al., 2007). Traffic volume (flow) is the number of vehicles that passing trough a point or section of a lane or roadway during a specific time period (Thiesenhusen et al., 2003), (Caceres et al., 2008). Traffic congestion is originated by the existence of some event that causes change in typical average speed values and travel times associated with a section. (Ygnance et al., 2001). Traffic density is closely related to congestion. It is defined as the number of vehicles occupying a given length of a lane or roadway at a certain instant. Its expressed as vehicles per kilometre. This is a key parameter to measure the quality of service on a road section (Ratti et al., 2006; Hansapalangkul et al., 2007; Pattara-Atikom et al., 2007).

OD matrices are used to quantify and synthesize mobility associated with persons or goods. They provide information about the number of trips performed between origin and a destination area during a given period of time. These matrices can be produced with different levels of aggregation depending on the level of detail desired or type of information required. By combining data on the cells and LA’s in which a mobile phone has been registered, it is possible to construct journey’s from which OD data may be inferred (Caceres et al., 2007; White and Wells, 2002, Sohn and Kim, 2008). Two recent simulation projects on the OD matrix focusing on the generation of traffic flows can be found in Spain (Caceres et al., 2007) and Korea (Sohn and Kim, 2008). 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.

7.4. Incident forecasting
Forecasting is the process of making statements about events whose actual outcomes (typically) have not yet been observed. Prediction is a similar but more general term and is a statement made about the future. Incident prediction in terms of a real accident is not possible; however, there can be made some historical correlations between incidents and telecom data in the surroundings, in order to be able to say "Watch out! This telecom situation means a higher probability of incidents!" and incident duration.

A key issue in traffic safety analysis is determining the reason for a site to be hazardous, also referred to as hot spot analysis. The road accident literature provides no universally accepted definition of a road traffic accident ‘hotspot’ (Brimicombe, 2004, Hauer, 1996). Areas of concentrated crime are often also referred to as Hot Spots (Gundogdu, 2008). Researchers and police use the term in many different ways. Some refer to Hot Spot addresses (Eck and Weisburd, 1995; Sherman et al., 1989), others refer to Hot Spot blocks (Taylor et al., 1984; Weisburd and Green, 1994), and others examine clusters of blocks (Block and Block, 1995).
Traffic accidents tend to be concentrated in clusters in geographic space (see e.g. Yamada and Thill, 2004); accidents are more likely to occur at dangerous locations. Concentrations of traffic accident occurrences suggest spatial dependence between accidents and common causes. These “black zones”, or zones with significantly high accident numbers, can be detected by several geostatistical techniques. This identification and analysis of locations producing more accident than the average, is hence an important step in traffic accident prevention (Aerts et al., 2006).

Figure 3: Overview of incident black spots in the Netherlands

The first step is to identify the dangerous locations. Road accident hotspot analysis has traditionally centred on road segments or specific junctions (Thomas, 1996; Cook et al., 2001) while area wide hotspots are often neglected. Hauer (1996) describes how some researchers rank locations according to accident rate (accidents per driven vehicle kilometre), while other researchers use accident frequencies (accident per road kilometre). Sometimes the number of accidents per vehicle-kilometre driven (VKD) or per number of vehicles is used to identify hot spots, other researchers use an absolute figure (accidents per km/year or per year), and some use a combination of both.

The key challenge is to produce a tool that can be implemented in the traffic management centres that will allow them to use traffic detector information to make real-time, short-term predictions of when and where incidents and congestion are likely to occur on the infrastructure. The focus is using telecom data and traffic flow conditions as predictors of incident conditions. A precursor is a variable derived from traffic stream data whose variations can indicate or point to a desirable pattern in traffic flow behaviour. Recent research in incident prediction has widely used the concept of precursors in its models for predictions. Traffic volume has traditionally been a precursor of interest to many researchers for statistically relating to crash frequency (Gribe, 2003). Hourly flow, which is a shorter time aggregation of volume, is another precursor that has been used by a couple of researchers to predict accident rate (Cedar, 1982; Hiselius, 2004). Research has found that the coefficient of variation in speed along the lane to be sensitive in predicting accidents (Oh et al., 2005; Lee et al. 2002; Kockelman et al, 2004). Traffic density is another parameter that has a good correlation in explaining incidents and is usually used in conjunction with the coefficient of variation (Lee et al., 2003; Abdel-Aty et al., 2004).
8. THE CURRENT CITY PROJECT

8.1 Introduction

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). An early attempt to construct a dynamic map of Amsterdam is ‘The Amsterdam Real Time project’ (Polak, 2002) and is based solely on the movement of a selected number of people carrying GPS receivers and being tracked in real time.

More in detail, the project uses anonymized data of the KPN Mobile network. In the study area over 1200 cells were identified, grouped in 8 LACs. It involves the city of Amsterdam and its surroundings, for an area of about 1000 km². Data included, for each sector:
1. Number of New Calls;
2. Number of SMS;
3. Number of Handovers;
4. Erlang, a standard unit of measurement of traffic volumes, equivalent to 60 minutes of voice calls (initiated, received or handed-over from other radio cells);
5. Number of Location Updates.

![Fig 4: Current City test area](image)

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.

The Current City approach is strongly related to the concept of Dynamic Data Driven Application Systems (DDDAS). It entails the ability to dynamically incorporate additional data into an executing application and creates a rich set of new challenges for applications, algorithms, systems, software and measurement methods. A nice example of the use of the DDDAS concept similar to Current City is WIPER (Madey et al., 2007). The concept is characterized by dynamic data, the first two Ds in DDDAS. In the Current City system the dynamic data is cell phone activity. Another nice example of the DDDAS concept is ‘Firegrid’. This is an initiative to create a next generation real-time emergency

---

5. [www.currentcity.org](http://www.currentcity.org)
response system using GRID technology. Challenges are sensing, modelling, forecast, feedback and response.

The key obstacle to effective emergency response is the communication needed to access relevant data or expertise and piece together an accurate understandable picture of reality (Hale, 1997). To improve situational awareness IM organizations need to have real time access to different kind of information:

- Alerts – has something happened?
- Location - Where is congestion located? Where are the road users? Where are the Incident managers / Emergency services?
- Cause - What causes congestion: incidents, events, weather?
- Numbers – how many people are involved?
- Movement – is the incident stationary or in movement?
- Mobility - How far the consequences of an incident reverberate on the road network? What is the site accessibility of emergency services?
- Safety - Which risks are there for the surrounding areas (e.g. chemical releases)?
- Command and Control - How should we respond? Which traffic management strategies do incident managers have at their disposal?
- Prediction – how to anticipate the incident? Are there special events with a higher risk for traffic jams and possible incidents?

8.2 First results Current City project Amsterdam

The design and operation of the Current City prototype application is shown schematically in figure 5. The system has three main components: 1) Data sources, storage streaming telecom data and pre-processing, 2) Data management, modelling and support system and 3) Web based applications for users.

The GSM cellular network is built from radio cells based on two best serving cell maps generated by antennas with two different frequencies overlaid each other, namely 900 MHz and 1800 MHz. In order to obtain the real mobile phone usage pattern of certain place in the city, the two Best Serving Area maps have been integrated. The telecom operator applied special scripts to extract the necessary data for the project. All received data is cell based with the duration interval of one hour. One raw of data contains aggregated information from one hour in a certain cell. Each cell contains a large number of grids. The spreading of cell traffic to grids is not simply a division but takes into account that land-use changes the spread: spreading thus depends on the land-use type that is present on a certain cell.
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. Figure 6 shows the increase of SMS activity during New Year’s Eve at 31 December 2008. Each dot on the map represents one sent SMS which captures an increase human activity during New Year’s Eve.

![Image](source: www.currentcity.org)

A more detailed data analysis 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 figure 7. The diagram shows for each day the average traffic (Erlang) over a period of five months (1 January – 30 May 2008). The data are normalized on the averages for comparison (Steenbruggen et al., 2010).

![Image](source: http://www.currentcity.org/)

A next step is anomaly detection. After exclusion of known extreme events (e.g. New Year’s Eve and Queen’s Day) the normality shows, for each grid cell the average traffic and standard deviation. The average value of SMS/Terminating Call/Originated Call, for each cell in our sample for each hour of the day of day of the week is calculated (based on the last 52 weeks of data). This is a dynamic process. Then we compare the real-time values (now, this hour of this day of this week) with the average values, for each cell. We produce the standard Deviations, for each real-time value from the average value, for each cell. We rank the standard deviations, and pick up the top 7 and bottom 4 associated cells, excluding the ones which are based on real-time value=0. The eleven selected cells with an anomaly are then showed with in the web-platform screen, disguising their actual
boundaries to protect the telecom operator confidentiality. With each new dataset from the telecom network this process is automatically repeated, the indicators are then recalculated and updated. This is straight forward approach. Next research steps are 1) to see if there are more appropriate anomaly detection methods to filter out anomalies that are specially related to incidents and traffic jams, 2) calculating real-time presence of people nearby an incident and 3) prediction of flows based on OD matrices.

9. RETROSPECT AND PROSPECT

Smart management of incidents in traffic systems has become a major challenge to road traffic operators and incident managers. In case of distributions of traffic flows, it is necessary to assess direct on-site and wider areal consequences in order to mitigate the road traffic externalities and restore a smooth traffic flow. To that end, a broad information base – including access to the site and the risks for surrounding areas – is needed. Electronic data may be instrumental in building up an efficient and effective IM information system, in particular using LBS data. A particularly interesting case of LBS data is offered by GPS and GSM data. The use of GSM data from cellular networks is very recent and has great potential. This paper has demonstrated – through a broad overview and the many examples from first pilot studies – that GSM data embody a great opportunity for smart IM. Clearly more work needs to be done to ensure the benefits of such systems but the first test cases appear to be highly promising.

The following general conclusions can be drawn:

- the topic of traffic incident management is one of the key concerns in transportation safety policy regarding the number of road traffic casualties per year and the huge impact on mobility;
- the use of telecom data has huge potential in providing new types of information services to create an accurate understandable picture of reality to improve situational awareness for the daily IM work processes;
- there are limited examples in the literature where telecom data are analysed for incident management, most projects focus on general traffic and emergency management;
- timeliness seems to be an important issue for incident detection, thus further research is needed to see if the use of telecom data can improve existing methods (e.g. cameras, detection loops and human observers);
- normality maps based on telecom data in combination with land use is a promising approach to improve incident detection and calculating presence of people;
- the use of telecom data for incident detection is strongly related to the magnitude of an incident or an event;
- telecom data could provide a real-time overview of the current number of people which are affected in case of incidents with dangerous goods; however, more research is required in terms of modelling and validation with ground truth;
- telecom data could be a useful source for prediction of flows and site accessibility of emergency services; however more research is needed in terms of modelling and integration of different data sources like loop detection data;
- telecom data has potential for incident prediction in terms of calculating traffic volume and traffic density in combination with other data sources.


TNO (2006) *Effecten van de landelijke invoering van incidentmanagement maatregelen op de voertuigverliestijd in het netwerk*, Jeroen Schrijver, Ben Immers, Maaike Snelder and Robert de Jong, 11 mei 2006, Delft 06.34.15/N079/034.65116/JJS/LK


<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007-2</td>
<td>Daniel Leliefeld, Evgenia Motchenkova</td>
<td>To protect in order to serve, adverse effects of leniency programs in view of industry asymmetry</td>
<td>29 p.</td>
</tr>
<tr>
<td>2007-5</td>
<td>Bas P. Singer, Bart A.G. Bossink, Herman J.M. Vande Putte</td>
<td>Corporate Real estate and competitive strategy</td>
<td>27 p.</td>
</tr>
<tr>
<td>2007-7</td>
<td>Stella Flytzani, Peter Nijkamp</td>
<td>Locus of control and cross-cultural adjustment of expatriate managers</td>
<td>16 p.</td>
</tr>
<tr>
<td>2007-13</td>
<td>Miranda Cuffaro, Maria Francesca Cracolici, Peter Nijkamp</td>
<td>Measuring the performance of Italian regions on social and economic dimensions</td>
<td>20 p.</td>
</tr>
</tbody>
</table>
2007-14  Tüzin Baycan-Levent
          Peter Nijkamp  Characteristics of migrant entrepreneurship in Europe, 14 p.

2007-15  Maria Teresa Borzacchiello
          Peter Nijkamp

2007-16  Tibert Verhagen
          Selmar Meents  A framework for developing semantic differentials in IS research: Assessing the meaning of electronic marketplace quality (EMQ), 64 p.

2007-17  Aliye Ahu Gülümser
          Tüzin Baycan
          Levent

2007-18  Laura de Dominicis
          Raymond J.G.M.
          Florax

2007-19  E. Dijkgraaf
<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-3</td>
<td>Marlous Agterberg, Bart van den Hooff, Marleen Huysman, Maura Soekijad</td>
<td>Keeping the wheels turning: Multi-level dynamics in organizing networks of practice, 47 p.</td>
</tr>
<tr>
<td>2008-5</td>
<td>Bernd Heidergott, Haralambie Leahu</td>
<td>Differentiability of product measures, 35 p.</td>
</tr>
<tr>
<td>2008-6</td>
<td>Tibert Verhagen, Frans Feldberg, Bart van den Hooff, Selmar Meents</td>
<td>Explaining user adoption of virtual worlds: towards a multipurpose motivational model, 37 p.</td>
</tr>
<tr>
<td>2008-9</td>
<td>Junbo Yu, Peter Nijkamp</td>
<td>China’s prospects as an innovative country: An industrial economics perspective, 27 p</td>
</tr>
<tr>
<td>2008-10</td>
<td>Junbo Yu, Peter Nijkamp</td>
<td>Ownership, r&amp;d and productivity change: Assessing the catch-up in China’s high-tech industries, 31 p</td>
</tr>
<tr>
<td>2008-11</td>
<td>Elbert Dijkgraaf, Raymond Gradus</td>
<td>Environmental activism and dynamics of unit-based pricing systems, 18 p.</td>
</tr>
<tr>
<td>2008-12</td>
<td>Mark J. Koetse, Jan Rouwendal</td>
<td>Transport and welfare consequences of infrastructure investment: A case study for the Betuweroute, 24 p</td>
</tr>
<tr>
<td>2008-13</td>
<td>Marc D. Bahlmann, Marleen H. Huysman, Tom Elfring, Peter Groenewegen</td>
<td>Clusters as vehicles for entrepreneurial innovation and new idea generation – a critical assessment</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>2008-16</td>
<td>Peter Nijkamp, Mediha Sahin</td>
<td>A study on Amsterdam, 33 p.</td>
</tr>
<tr>
<td>2008-17</td>
<td>Enno Masurel</td>
<td>Local shopkeepers’ associations and ethnic minority entrepreneurs, 21 p.</td>
</tr>
<tr>
<td>2008-23</td>
<td>Patricia van Hemert, Peter Nijkamp</td>
<td>Thematic research prioritization in the EU and the Netherlands: an assessment on the basis of content analysis, 30 p.</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>2009-3</td>
<td>Julie E. Ferguson, Marleen H. Huysman</td>
<td>Between ambition and approach: Towards sustainable knowledge management in development organizations, 33 p.</td>
</tr>
<tr>
<td>2009-6</td>
<td>Eric de Noronha Vaz, Mário Caetano, Peter Nijkamp</td>
<td>Trapped between antiquity and urbanism. A multi-criteria assessment model of the greater Cairo metropolitan area, 22 p.</td>
</tr>
<tr>
<td>2009-7</td>
<td>Eric de Noronha Vaz, Teresa de Noronha Vaz, Peter Nijkamp</td>
<td>Spatial analysis for policy evaluation of the rural world: Portuguese agriculture in the last decade, 16 p.</td>
</tr>
<tr>
<td>2009-8</td>
<td>Teresa de Noronha Vaz, Peter Nijkamp</td>
<td>Multitasking in the rural world: Technological change and sustainability, 20 p.</td>
</tr>
<tr>
<td>2009-9</td>
<td>Maria Teresa Borzacchiello, Vincenzo Torrieri, Peter Nijkamp</td>
<td>An operational information systems architecture for assessing sustainable transportation planning: Principles and design, 17 p.</td>
</tr>
<tr>
<td>2009-10</td>
<td>Vincenzo Del Giudice, Pierfrancesco De Paola, Francesca Torrieri, Francesca Pagliari, Peter Nijkamp</td>
<td>A decision support system for real estate investment choice, 16 p.</td>
</tr>
<tr>
<td>2009-11</td>
<td>Miruna Mazurencu Marinescu, Peter Nijkamp</td>
<td>IT companies in rough seas: Predictive factors for bankruptcy risk in Romania, 13 p.</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>2009-15</td>
<td>André de Waal, Karima Kourtit, Peter Nijkamp</td>
<td>The relationship between the level of completeness of a strategic performance management system and perceived advantages and disadvantages, 19 p.</td>
</tr>
<tr>
<td>2009-20</td>
<td>Karima Kourtit, André de Waal, Peter Nijkamp</td>
<td>Strategic performance management and creative industry, 17 p.</td>
</tr>
<tr>
<td>2009-22</td>
<td>Roberta Capello, Peter Nijkamp</td>
<td>Regional growth and development theories revisited, 19 p.</td>
</tr>
<tr>
<td>2009-23</td>
<td>M. Francesca Cracolici, Miranda Cuffaro, Peter Nijkamp</td>
<td>Tourism sustainability and economic efficiency – a statistical analysis of Italian provinces, 14 p.</td>
</tr>
<tr>
<td>Year-Customer</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>2009-28</td>
<td>Marco Percoco, Peter Nijkamp</td>
<td>Individual time preferences and social discounting in environmental projects, 24 p.</td>
</tr>
<tr>
<td>2009-29</td>
<td>Peter Nijkamp, Maria Abreu</td>
<td>Regional development theory, 12 p.</td>
</tr>
<tr>
<td>2009-30</td>
<td>Tüzin Baycan-Levent, Peter Nijkamp</td>
<td>7 FAQs in urban planning, 22 p.</td>
</tr>
<tr>
<td>2009-31</td>
<td>Aliye Ahu Gülümser, Tüzin Baycan-Levent, Peter Nijkamp</td>
<td>Turkey’s rurality: A comparative analysis at the EU level, 22 p.</td>
</tr>
<tr>
<td>2009-33</td>
<td>Mediha Sahin, Peter Nijkamp, Marius Rietdijk</td>
<td>Cultural diversity and urban innovativeness: Personal and business characteristics of urban migrant entrepreneurs, 27 p.</td>
</tr>
<tr>
<td>2009-34</td>
<td>Peter Nijkamp, Mediha Sahin</td>
<td>Performance indicators of urban migrant entrepreneurship in the Netherlands, 28 p.</td>
</tr>
<tr>
<td>2009-35</td>
<td>Manfred M. Fischer, Peter Nijkamp</td>
<td>Entrepreneurship and regional development, 23 p.</td>
</tr>
<tr>
<td>2009-37</td>
<td>Enno Masurel, Peter Nijkamp</td>
<td>Bridging the gap between institutions of higher education and small and medium-size enterprises, 32 p.</td>
</tr>
<tr>
<td>2009-42</td>
<td>Teresa de Noronha Vaz, Peter Nijkamp</td>
<td>Knowledge and innovation: The strings between global and local dimensions of sustainable growth, 16 p.</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>2009-48</td>
<td>Andrea Caragliu, Chiara Del Bo, Peter Nijkamp</td>
<td>Smart cities in Europe</td>
</tr>
<tr>
<td>2009-50</td>
<td>Eric de Noronha Vaz, João Pedro Bernardes, Peter Nijkamp</td>
<td>Past landscapes for the reconstruction of Roman land use: Eco-history tourism in the Algarve</td>
</tr>
<tr>
<td>2009-51</td>
<td>Eveline van Leeuwen, Peter Nijkamp, Teresa de Noronha Vaz</td>
<td>The Multi-functional use of urban green space</td>
</tr>
<tr>
<td>2009-52</td>
<td>Peter Bakker, Carl Koopmans, Peter Nijkamp</td>
<td>Appraisal of integrated transport policies</td>
</tr>
<tr>
<td>2009-53</td>
<td>Luca De Angelis, Leonard J. Paas</td>
<td>The dynamics analysis and prediction of stock markets through the latent Markov model</td>
</tr>
<tr>
<td>2009-55</td>
<td>Bas Straathof, Gert-Jan Linders</td>
<td>Europe’s internal market at fifty: Over the hill?</td>
</tr>
<tr>
<td>2009-58</td>
<td>René Sitters</td>
<td>Efficient algorithms for average completion time scheduling</td>
</tr>
</tbody>
</table>
Masood Gheasi
Peter Nijkamp
Piet Rietveld

Migration and tourist flows, 20 p.
<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-1</td>
<td>Roberto Patuelli, Norbert Schanne, Daniel A. Griffith, Peter Nijkamp</td>
<td>Persistent disparities in regional unemployment: Application of a spatial filtering approach to local labour markets in Germany, 28 p.</td>
</tr>
<tr>
<td>2010-3</td>
<td>John Steenbruggen, Maria Teresa Borzacchiello, Peter Nijkamp, Henk Scholten</td>
<td>Real-time data from mobile phone networks for urban incidence and traffic management – a review of application and opportunities, 23 p.</td>
</tr>
<tr>
<td>2010-6</td>
<td>Mark G. Lijesen</td>
<td>Empirical applications of spatial competition; an interpretative literature review, 16 p.</td>
</tr>
<tr>
<td>2010-7</td>
<td>Carmen Lee, Roman Kraeussl, Leo Paas</td>
<td>Personality and investment: Personality differences affect investors’ adaptation to losses, 28 p.</td>
</tr>
<tr>
<td>2010-8</td>
<td>Nahom Ghebrihiwet, Evgenia Motchenkova</td>
<td>Leniency programs in the presence of judicial errors, 21 p.</td>
</tr>
<tr>
<td>2010-10</td>
<td>Jani Merikivi, Tibert Verhagen, Frans Feldberg</td>
<td>Having belief(s) in social virtual worlds: A decomposed approach, 37 p.</td>
</tr>
<tr>
<td>2010-11</td>
<td>Umut Kilinç</td>
<td>Price-cost markups and productivity dynamics of entrant plants, 34 p.</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>2011-3</td>
<td>Eric de Noronha Vaz Doan Nainggolan Peter Nijkamp Marco Painho</td>
<td>A complex spatial systems analysis of tourism and urban sprawl in the Algarve, 23 p.</td>
</tr>
<tr>
<td>2011-7</td>
<td>Amitrajeet A. Batabyal Peter Nijkamp</td>
<td>A Schumpeterian model of entrepreneurship, innovation, and regional economic growth, 30 p.</td>
</tr>
<tr>
<td>2011-10</td>
<td>Tıbert Verhagen Jaap van Nes Frans Feldberg Willemijn van Dolen</td>
<td>Virtual customer service agents: Using social presence and personalization to shape online service encounters, 48 p.</td>
</tr>
<tr>
<td>2011-13</td>
<td>Aliye Ahu Akgün Tüzin Baycan Peter Nijkamp</td>
<td>Repositioning rural areas as promising future hot spots, 22 p.</td>
</tr>
</tbody>
</table>
2011-15  Aliye Ahu Gülümser 
Tüzin Baycan-Levent 
Peter Nijkamp

2011-16  Frank Bruinsma 
Karima Kourtit 
Peter Nijkamp
Tourism, culture and e-services: Evaluation of e-services packages, 30 p.

2011-17  Peter Nijkamp 
Frank Bruinsma 
Karima Kourtit 
Eveline van Leeuwen
Supply of and demand for e-services in the cultural sector: Combining top-down and bottom-up perspectives, 16 p.

2011-18  Eveline van Leeuwen 
Piet Rietveld
Climate change: From global concern to regional challenge, 17 p.

2011-19  Eveline van Leeuwen 
Peter Nijkamp
Operational advances in tourism research, 25 p.

2011-20  Aliye Ahu Akgün 
Tüzin Baycan 
Peter Nijkamp
Creative capacity for sustainable development: A comparative analysis of European and Turkish rural regions, 18 p.

2011-21  Aliye Ahu Gülümser 
Tüzin Baycan-Levent 
Peter Nijkamp
Business dynamics as the source of counterurbanisation: An empirical analysis of Turkey, 18 p.

2011-22  Jessie Bakens 
Peter Nijkamp
Lessons from migration impact analysis, 19 p.

2011-23  Peter Nijkamp 
Galit Cohen-blankshtain
Opportunities and pitfalls of local e-democracy, 17 p.

2011-24  Maura Soekijad 
Irene Skovgaard Smith
The ‘lean people’ in hospital change: Identity work as social differentiation, 30 p.

2011-25  Evgenia Motchenkova 
Olgerd Rus
Research joint ventures and price collusion: Joint analysis of the impact of R&D subsidies and antitrust fines, 30 p.

2011-26  Karima Kourtit 
Peter Nijkamp
Strategic choice analysis by expert panels for migration impact assessment, 41 p.

2011-27  Faroek Lazrak 
Peter Nijkamp 
Piet Rietveld 
Jan Rouwendal
The market value of listed heritage: An urban economic application of spatial hedonic pricing, 24 p.

2011-28  Peter Nijkamp
Socio-economic impacts of heterogeneity among foreign migrants: Research and policy challenges, 17 p.

2011-29  Masood Gheasi 
Peter Nijkamp
Migration, tourism and international trade: Evidence from the UK, 8 p.

2011-30  Karima Kourtit