POLITECNICO DI MILANO

DIPARTIMENTO DI
ARCHITETTURA E
STUDI URBANI

MOBILE PHONE DATA FOR MAPPING URBAN DYNAMICS

Paola Pucci, Fabio Manfredini, Paolo Tagliolato

DAStU
Working Papers
n. 3/2013
ISSN 2281-6283

This working paper is hosted by Dipartimento di Architettura e Studi Urbani, www.dastu.polimi.it/pubblicazioni/working_papers

Pucci, Manfredini, Tagliolato | {*paola.pucci, fabio.manfredini, paolo.tagliolato}@polimi.it*

Abstract
During the last few years, new tools for estimating people's density in cities have emerged through mobile network data. As opposed to the more traditional methods of urban surveys, the use of aggregated and anonymous mobile phone network log files has shown promise for large-scale surveys with notably smaller efforts and costs. Moreover, a frequent data feed from the mobile network has been argued to demonstrate fine grain over-time variation in urban movements, lacking from the traditional prediction methods. Despite the positivist approach to the new methodology, additional evidence is needed to show how mobile network data correlate with the actual presence of people, and how they can be used to map different urban domains. We try to address this shortcoming presenting the results of a research carried out in Lombardy Region, using mobile phone data provided by Telecom Italia, as a promising approach to assist the traditional database and analysis of urban dynamics as new challenges for urban planning.

1. Introduction

New tools for estimating people's density and collecting mobility measurements from mobile phone traffic analysis have been introduced and studied during recent years by research centres and private companies.
Compared to the traditional methods of urban surveys, the use of aggregated and anonymous mobile phone network log files has shown promise for large-scale surveys with notably smaller efforts and costs (Reades, Calabrese, Sevtsuk and Ratti, 2007).
Traditional data collection sources for urban planning are mainly based on statistical surveys and are not able to catch the variability in the intensity of urban space usage by present population.
Traditionally, traffic flows are automatically collected from a subset of links in a network using fixed detectors located along the road, such as inductive loops, infrared, visual cameras, radars, lasers, etc.
These measurements can be collected with little effort but they provide only localized information. Unfortunately, the extension or the modification of these types of instruments tends to be expensive due to the cost of installation and maintenance of detectors.
With the ubiquity and ever-increasing capabilities of mobile devices, the localization of mobile phones could potentially become a powerful source to demonstrate fine grain over-time variations in urban movements, which are lacking from traditional prediction methods (Arai, 2006; Ratti, Pulselli, Williams, Frenchman, 2006).
Within this approach, mobile phone users can be considered as sensors of a network, distributed in the space, who are able to provide information about forms and modes of usage of urban spaces, that traditional data sources (like census, interviews or the deployment of sensor networks), can hardly return, being extremely expensive and scarcely updated and updatable.
Indeed, many authors indicate passive and anonymous monitoring of mobile phone traffic as a valid alternative or complementary to traditional methods, these being able to resolve both traditional survey's limitations of latency (mobile network information can be easily retrieved in real time) and pervasivity (huge diffusion of mobile phones) at once. In particular, the use of mobile phone traffic data shows many advantages:

- large data samples, proportional to the high penetration rates of mobile phones in modern societies, with many countries surpassing the 90 percent penetration rate (CIA World Factbook, 2009).
- low implementation cost and the need for no additional hardware;
- monitoring of any covered area around an antenna, given the extent of mobile phone network coverage;

Pucci, Manfredini, Tagliolato | *{paola.pucci, fabio.manfredini, paolo.tagliolato}@polimi.it*

- generation of data almost in real time;
- high spatial and temporal resolution, compared to traditional and institutional data sources.

However, despite the positivist approach to the new methodology, additional evidence is needed to show how mobile network signals correlate with the actual presence of people in the city, and how they can be used to characterize and map different urban domains and their occupants and how this tool could support urban planning and urban policies.

The possibility to establish a correlation between mobile telephone traffic and presence of people in an urban context represents a relevant and necessary condition for exploiting mobile phone network data also for surveying the density of people's presence, which would be extremely useful for the implementation of policies concerning supply and services, that are more effective in responding to different needs.

The paper presents the results of a research carried out using mobile phone network data, aimed at verifying the validity of telephone traffic data in describing the density of use of the city and its spatial differences at the urban block scale.

This paper provides some insights gained from our research, outlining the potential of this technology in the urban planning community.

In order to analyze the complex temporal and spatial patterns of mobile phone activity, we used mobile phone traffic data provided by Telecom Italia, the main Italian phone operator. The data represent the measure of Erlang in each pixel covering the whole Lombardy Region (Northern Italy), which describes the number of active mobile phone calls as a function of position and time at a spatial resolution of about 250 meters, every 15 minutes, for one year period.

In the first step of our research, we focused on the reliability of Erlang data with respect to traditional data sources. We therefore compared mobile phone data of the main cities of Lombardy Region and their population dynamics[1] during a typical weekday and we found a strong correlation between these variables proving the potentialities of Erlang data for describing the variability of daily changes of urban population at the municipality scale.

The second step was the analysis of mobile phone activity trends for selected urban districts, characterized by comparable and similar land-use patterns of population and activities, densities and socio-economical profiles.

Analyzing Erlang trends during a typical week for each of these urban situations, the results show that mobile phone activity patterns can provide useful information for interpreting the specific dynamics of different urban situations such as mono-functional residential areas, railway stations, urban sprawl areas, social housing districts, university districts, industrial ambits.

This condition makes the data useful for a spatial functional clustering of the territory, as a relevant information - now available from surveys - for the urban planning.

We also worked on mobile phone traffic data at the Milan urban region scale[2], during the 2009 and 2010 International Design Week, a leading event which concentrates its activities in the Fair area and in tens of places within and outside the Milan city centre.

We mapped and interpreted the spatial configuration of mobile phone activity in order to represent a large event with significant concentration of traffic, from the beginning until its conclusion, as an important information for managing the spatial and temporal impacts on the urban system (in term of mobility monitoring, mobility and services demand).

[1] The latter was obtained from a traditional mobility database, the 2002 Lombardy Region Origin-Destination survey, which provides information on the population trips in a working day (see section 4.1).

[2] The Milan urban region is defined as an extended region that go far beyond the traditional administrative structure. Milan is the main center of the urban region which is characterized by a stratum of dense urbanization stretched over the ancient framework while the bordering provinces have been incorporated in the strongly urbanized and enlarged urban region.

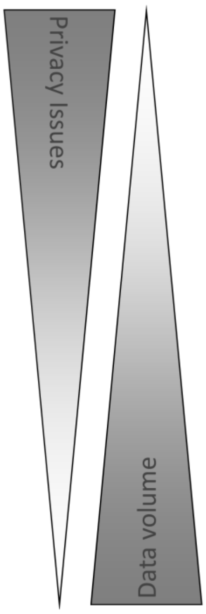
Further analysis focused on the correlation between the intensity of telephone calls at certain times of the day with the spatial configuration of residents and workers in the Milan urban region. The outcomes showed that telephone traffic data could effectively help to represent and to describe, dynamically over time, the intensity of activities and the presence of temporary populations at the urban scale.
Because of its spatial and temporal resolution, mobile phone data constitute an interesting and unique source of information on urban uses. Indeed, if we consider observed and aggregated telephone traffic as the result of individual behaviours and habits, we conclude that mobile phone data can provide information, which changes over time in urban contexts and can bring about new interpretations of urban and spatial dynamics.
This study therefore suggests that mobile phone-network data have the potential to drastically change the way we view and understand the urban environment.
Secondly, it explores whether mobile network data can reveal the significant time-dependent variation, which is missing from traditional analysis and can thus describe cities dynamically over time.
A further conclusion is that urban planning competences, with specific knowledge on urban dynamics, are needed to correctly interpret mobile phone data and to characterize and map urban contexts and their occupants, as emerged from interviews with different stakeholders, belonging to private and public sectors, with which also future applications have been discussed (event management, civil protection, mobility monitoring, urban rhythms analysis and mapping).

2. State of the art

Emerging disciplines (mobile positioning, space-time movement studies, life-map geography) are involved in the analysis, visualization and interpretation of people's presence and movements in urban spaces through mobile phone traffic.
The technology for mobile phones geographic location and other hand-held devices are becoming increasingly available. It is opening the way to a wide range of applications, collectively referred to as Location Based Services (LBS) that are primarily aimed at individual users. However, if deployed to retrieve aggregated data in cities, LBS could become a powerful tool for urban analysis (Ratti et al., 2006).
Indeed, in recent years several research projects aimed at understanding whether and how the telephone traffic data can be used as tools of analysis and representation of urban dynamics and people’s movements, have been developed.
These interdisciplinary studies in the fields of urban geography, social studies, computing and interaction design recognize anonymous and passive monitoring telephone traffic as a valuable alternative to traditional methods, because it can simultaneously overcome the limitations of the detection of latency time typical of traditional data sources and can take advantage of the pervasiveness of the detection area due to the ubiquity of mobile phone networks.
These studies proved that mobile phone data are relevant to estimate and establish the quality of urban spaces: people's presence is indeed an indication of urban vitality and of different uses of the urban spaces in time (temporal extension of urban activities) and in its functional patterns (Ahas and Mark, 2005; Ahas, Aasa, Silm and Tiru, 2010).
Many indices of urban quality can be developed from passive mobile positioning data (space-time coordinates), by a reliable, quick and not much expensive tool. These indices can be useful both for spatial policies and for commercial evaluations of private actors (real estate, marketing).
There are basically three main types of survey methodology, emerging from literature (Figure 1):

- Individual traces detected with tracking technologies (such as GPS, SMS, ...) of a sample, useful to study the mobility behaviour of specific groups of people;
- Individual trajectories, previously anonymized and collected by mobile phone carriers, useful to study geometric patterns of individual mobility, without geographical references (Network science);
- Mapping geo-referenced and aggregated mobile phone data useful, to study land use density (Mobile landscapes); network measurement results related to active calls allow tracking of all active handsets.

The first technique - individual traces of a selected sample - certainly offers a more precise result because it is possible to record the origin and destination track of individual moves. On the other hand this means a greater cost for data processing and the necessity to build a statistical sample of users. Moreover, problems related to individual privacy raised several ethical questions for this type of research.
This technique is based on active mobile positioning (tracing) that occurs through a specific location request, both as network based positioning and as handset positioning method.
Instead, the use of aggregated data collected from the network (mainly cell towers) allows moving from the individual level, focusing the interest on the emergence of complex urban dynamics related to the places that people use and frequent (Song, Qu, Blumm and Barabasi, 2010; Gonzalez et al., 2008).



Data	Research area	Purpose
Tracks of displacements of individuals of a small selected sample (case study)	Tracking technologies (GPS, SMS, ...)	Study of mobility behaviour of specific group/population category
Tracks of calls of individual phone users (active mobile positioning)	Network science	Study of geometrical patterns of the mobility of individuals (deprived of the specific geographical reference)
Spatial distribution of cell phone activity (derived from cellular network log files)	Mobile landscapes	Study of the density of use of a territory

Figure 1 – The research fields on active and passive mobile positioning

Within the third technique, the focus becomes the use of urban space by the people, considered as time and space dependent variable. The scale of investigation widens and this type of synthetic data allows the representation of urban density also through the mapping in real time (named soft real-time because the data are usually returned with a delay of 15 minutes).

Unlike origin destination matrices or individual mobile phone tracking, aggregate data do not indicate where a caller comes from or goes to, but they simply estimate the amount of call volume in a given network cell at a given time. Although we lose the traces of the origins and destinations of individual movements, this limitation does not appear relevant if, by using the volume of cell activity in mobile network cells, we can estimate the distribution patterns of the population in different time slots considered for the survey (hourly, weekly, seasonal).
From a technical point of view, it is based on the analysis of aggregated data and traffic volume detected on towers network. Among the methods proposed in literature, we mention the social positioning method (SPM) of Positium LBS (Ahas and Mark, 2005; Ahas et al., 2010) based on active and passive positioning systems, and mobile census (MIT Senseable City Lab) which is instead a totally passive tracking system.
The opportunities offered by the use of mobile phone data compared to traditional data sources are:

- A more regular distribution of data in time and space;
- A finer network of detection;
- The precision of information (the accuracy of locational data, the frequency of data availability);
- The time required for calculating the position;
- The availability of service coverage, especially in urban areas;
- The characteristics of aggregated and anonymous data do not infringe the privacy of mobile phone users;
- The implementation of integrated solutions that enhance the information obtained from mobile data, combining information from the identification code of the telephone prefixes for outgoing and incoming calls (ID of incoming and outgoing calls), but also on user profiles (social identification).

Since locational data (GPS, A-GPS, SMS) are becoming increasingly available and their applications are currently a hot topic in the mobile phone industry, aggregated locational data have not yet been widely used to describe urban systems.
Research efforts on this topic are emerging: in particular on the mapping of the mobile phone activity in cities or on the visualization of urban metabolism (Wolman, 1965; Acebillo and Martinelli, 2012; Brunner, 2007) based on handsets' movements.
Because our research is focused on the analysis of aggregated mobile phone data provided by Telecom Italia, the studies selected from the bibliographical review are referred to mobile phone traffic density data integrated with information on user profiles or with tracking analysis data.
Using mobile phones as a device for monitoring urban practices shows that phone calls are closely related to population density in urban areas: the intensity of activity in a cell (the covered area around an antenna) is proportional to the presence of mobile phone users (Sevtsuk and Ratti, 2010; Reades et al., 2007; Ratti et al., 2006; Ahas and Mark, 2005).
Most of the researches focus on the Erlang data, using the volume of call activity in mobile network cells as the spatial unit of analysis (Sevtsuk and Ratti, 2010), in order to describe the correlation between mobile phone data and people's daily activities (that are not the density of people, but the density of the phone activities) (Ratti et al., 2006).
The aim is to understand patterns of daily life in the city, using a variety of sensing systems (Erlang measures as telecommunication traffic intensity, Location-based data as GPS devices, wireless sensor network) and to illustrate and confirm that there are significant differences in the urban activity distribution at different hours, days and weeks.
Graphic representations of the intensity of urban activities and their evolution through space and time, based on the geographical mapping of mobile phone usage at different times of the day (Ratti et al., 2006) are the main output of the Mobile Landscape approach.

Pucci, Manfredini, Tagliolato | *{paola.pucci, fabio.manfredini, paolo.tagliolato}@polimi.it*

The main question is how to correlate the Erlang trends - a measure of the density of phone activities - with the density of people.
Because the data can be used to map different urban domains and their occupants, some studies focused on integrating available mobile phone data with traditional data sources and surveys (Pucci, Tagliolato, Manfredini, 2010; Manfredini, Pucci, Tagliolato, 2012).
The aim is to identify a relevant correlation between mobile phone intensity and people density and to propose "social profiles" of mobile phone users, which influence the intensity of calls, emerging from Erlang data.
In this direction, statistical models supported by empirical data (traditional sources or surveys) are proposed to investigate:

- the relationships between location coordinates of mobile phones and the social identification of the people carrying them (as Social Positioning Method and its possible applications in the organization and planning of public life proposed by Rein Ahas and Ülar Mark, 2005);
- the relationships between mobile phone measures (the volume of call activity in mobile network cells as Erlang) and population distribution in cities (Sevtsuk and Ratti, 2010).

The state of the art for using mobile phones as traffic monitoring tools can be found in Caceres, Widenerg and Benitez, 2007; Qiu and Cheng, 2007; Fontaine and Smith, 2005. In terms of volume data, the concept of using mobile phones as probes has been explored by various researchers working on simulated frameworks (Fontaine and Smith, 2005), as well as in field tests (Bekhor, Hirsh, Nimre and Feldman 2008; Höpfner, Lemmer and Ehrenpfordt, 2007; Thiessenhusenk, Schafer and Lang, 2006).
In all cases, volume traffic data would be associated with cell phone transit through a boundary area using processes related to mobility management, detecting boundary crossing rates either at inter-cell boundary level (handover) or at location area boundary level (LU procedure).
Most of them focus on the handover event detecting phone transit through boundaries between two cells. However, these studies concluded that accurate vehicle flows couldn't be obtained directly from mobile phone data due to the characteristics of this data source.
The main question is how to correlate the number of crossing phones with the real number of crossing vehicles. Volume data on inter-cell boundaries provided by mobile phones does not yield information on the complete set of vehicles crossing a boundary, but only a statistical sample of all travelling vehicles.
According to these findings, more accurate estimates for the number of vehicles can be obtained by means of adequate treatments to relate both measures (phone counts and vehicle counts) (Caceres, Castillo, Romero, Benitez and Del Castillo, 2010).
Although studies show promising results, the attempt to establish a "direct" link between phone calls and the number of people or trips, encounters some major limitations.
To begin with, the use of the mobile phone depends on age, gender, profession, time, activities (Aguiléra, Guillot, Bonin, 2009) and it is difficult to take into account the possible cross effects: the large number of situations means that it is almost impossible to reach a conclusion on a purely quantitative basis derived solely by the mobile phone data.
This is particularly important if we want to use mobile phone data for urban investigations aimed at planning the provision of personal services, for which statistical data are needed.
Next, long term effects may diverge from short term effects, in particular because when individuals gain familiarity with these technologies, they may start to combine them, or because the equipment rate increases and the available functionalities change rapidly, as has been the case of mobile phones. Moreover, the correct measurement scale is not necessary that of the individual.

2.1. Some possible applications

Although these approaches are experimental, some experiences - often interdisciplinary pilot-studies - have evaluated the impact of applications of mobile phone data analysis, especially on urban and territorial marketing, mobility management, urban land-use planning, tourism management and social network profiling. These experiences underline the following benefits retrievable from mobile phone data:

- Marketing and territorial management applications (Positium LBS ; www.positium.ee; Ülar Mark, Heikki Kalle, Rein Ahas),
- Real-time views of people fluxes in urban areas (Mobile Scape Graz, 2005, Real Time Rome, 2006; Wiki City Rome, 2007) and views of traffic flows of IP mobile phone traffic (Senseable City Lab; http://senseable.mit.edu; Carlo Ratti); Current City (http://currentcity.org; Euro Beinat, Assaf Biderman, Francesco Calabrese, Filippo Dal Fiore, Carlo Ratti, Andrea Vaccari);
- Definition, analysis and assessment of feasible models to infer traffic volume: using anonymous call data, collected automatically by mobile phone network operators (Caceres et al., 2007; Caceres et al., 2010); using mobile phone positions derived by signal strength measured by serving antennas (Ramm and Schwieger, 2007;
- Creating a platform for analyzing mobile database in real time, and applications to provide tracks for localizing people (social navigation) (Sense Networks; http://www.sensenetworks.com; Greg Skibiski, Alex Pentland);
- Passive location of mobile users in the interior spaces, through a radio frequency detection system installed at the sites and to provide marketing surveys for the location of shops and advertising (Path Intelligence; http://www.pathintelligence.com);
- Environmental data collection from digital traces to study the practices of social networking and to model the behaviour of people through stochastic methods (Reality Mining MIT MediaLab; http://reality.media.mit.edu; Nathan Eagle).

3. The analysis framework

Starting from the need to verify direct correlation between mobile phone data and density of people, the research has compared mobile phone traffic data with institutional and traditional statistical variables.
This comparison is intended to evaluate possible correlations between the variability of mobile network calls' intensity and the variability of urban activities (density, presences).
Our research was aimed at:

- Evaluating the use of telephone traffic data for providing statistically relevant description of urban activities;
- Describing the intensity of the use of the city (during the day, weekdays / holidays, seasons) linked to the differences in the urban activities distribution at different hours, days and weeks, as a tool to define urban policies oriented toward the offer of services for the individual behaviours;
- Managing large and special events (inflow, outflow, monitoring), also estimating population density;
- Describing time-dependent variation that is missing from traditional analysis and could thus describe cities dynamically over time.

Mobile phone data, provided by Telecom Italia, are mobile phone traffic matrices, expressed in Erlang, namely the average number of concurrent contacts in a time unit.

In the present case, the data represent the telephone traffic density every 15 minutes and was supplied by Telecom Italia in a spatialized form. From the telephone traffic recorded by each cell of the network, Telecom Italia distributed the measurements, by means of weighted interpolations, throughout a tessellation of the territory in 250 meters x 250 meters squared areas (pixels). Figure 2 shows the mobile phone traffic data spatial resolution superimposed to building blocks, the more detailed available statistical unit, for the central area of Milan.
In the present work we do not consider the Erlang data directly. Raw data describe the amount of the absolute mobile phone traffic in one pixel at a certain moment. We are interested in something that can be related with the presence of people. If we consider the raw data this is unlikely: there are obviously less telephone calls in the early morning (let us say around 5 am) than in the afternoon, and this does not depend on the timely variation in the amount of people in a given spatial region. We will then consider a different measure derived from raw data: given a set of pixels (i.e. a spatial region) at a certain moment, we will take into account the ratio between the Erlang of traffic in those pixels and the total amount of traffic, at that same moment, in the "universe" (i.e. the sum of traffic on all the pixels of the whole matrix, representing the Lombardy Region in our case)[3]. This is a relative measure that tells us the amount of telephone traffic in a certain spatial region with respect to the total telephone traffic, and this is more likely to tell us something about the variation in the amount of people.

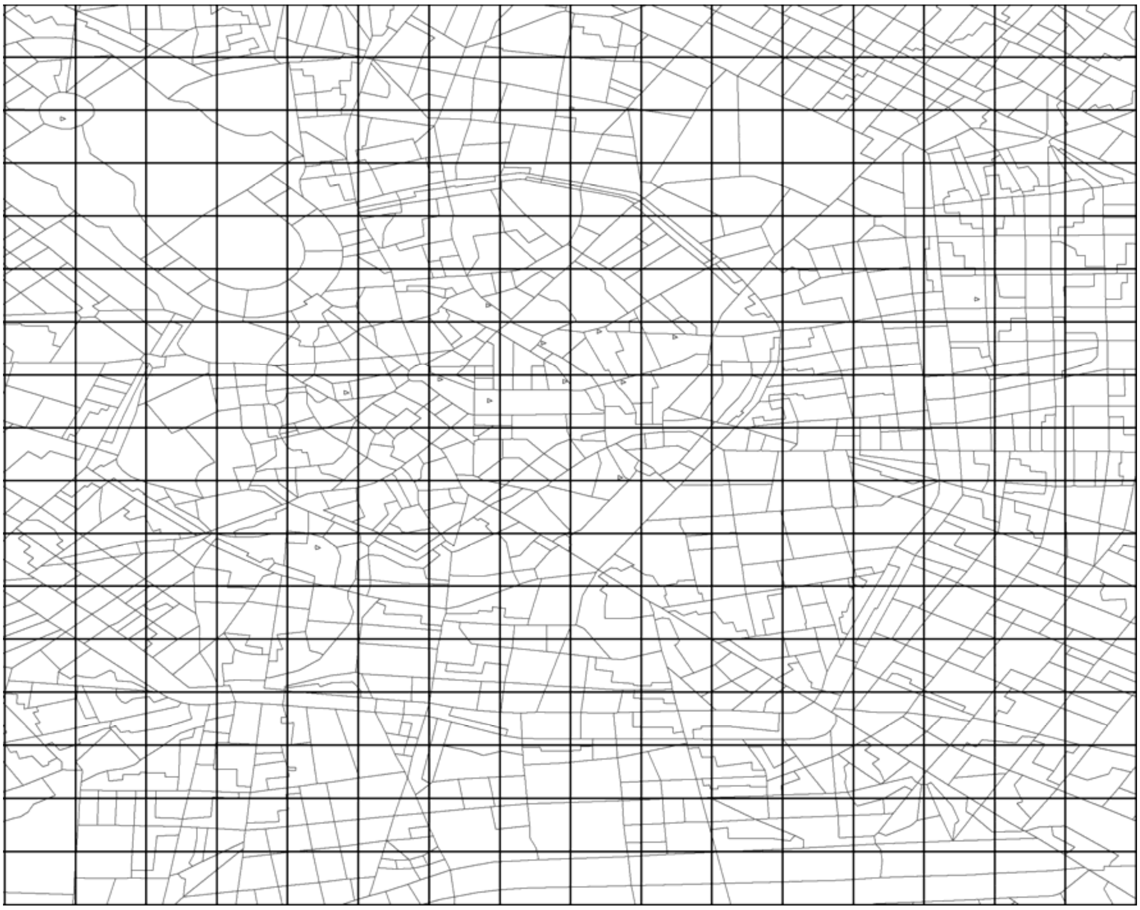


Figure 2 – The telephone traffic matrix (in black) and the city building blocks (in grey): a central area of Milan.

[3] We considered in the present work also a slightly modification of this measure: the ratio between the Erlang per pixel instead of the Erlang tout court.

The data are characterized by a high spatial and temporal resolution and they show the intensity of telephone traffic per unit of time, but they do not allow a direct correlation between traffic intensity and number of users that generate it.
It is known that the modes of mobile phone use vary, consistently, with the age of the users and their socio-professional condition.
In addition, the Telecom Italia cell-traffic is a sample, as Telecom users are about 40% of the market on a national scale, and of these, only those that generate traffic are recorded by Erlang data. Therefore these precautions oriented the data analysis approach.

4. The approach

Telephone traffic data at our disposal have been compared with statistical variables derivable from traditional and well-established data sources.
This comparison aimed at the evaluation of possible correlations between the variability of mobile network calls' intensity and the variability of urban activities at two scales: macro and micro scales.
At the macro scale, we explored whether the mobile phone data in Erlang (aggregated in the main cities of the Lombardy Region) are correlated with the variation of density of people at a rate of 60 minutes in the same cities. The time series of people's density in Lombardy Region was obtained by an elaboration of a traditional database, the Origin-Destination matrix (see section 4.1), with difficult updating, leading to an estimate of the population reported in any municipality in any hour of a typical weekday.
Our intent was to inquire whether cell phone activity correlates with people's presence in the main cities of Lombardy Region and, therefore, this analysis was used to support the evaluation of the significance of the phone calls data to return the density of people at different times of the day.
At the micro scale, we selected similar urban context types[4] – e.g social housing districts, railway stations, university districts, factory areas, residential districts, business districts, areas of urban sprawl – in order to check possible correlations between land-use patterns and cell phone activities, for any different selected urban zones.
Our analysis aimed at evaluating if similar urban context types, in term of land-use patterns and socio-economic profiles, have a recognizable mobile phone attitude.
In this case, the aim is to propose a functional clustering of the Lombardy Region, starting from the Erlang trends. At the same scale we examined also the variabilities of density before and during a special event like the "Salone del Mobile - International Design Week" in Milan city town.
For both scales of analysis, statistical correlations between Erlang and (few) comparable data (land use, density of population, socio-economic trends, data on commuters, surveys…) show that the cell phone activity describes people's daily activities well.

[4] The urban district is one or more building blocks with similar land-use and activity densities. For the railway stations, we have taken into account the catchment area of each station. So, the urban districts are smaller than municipality.

4.1 The macro scale analysis. Mobile phone activity: an indicator of people presence?

In the macro scale analysis, we tested the hypothesis that a proportionality between the changeability of the cell phone activity and the variability of the present population in one typical weekday exists.
This part of the work was intended as a first evaluation of nature of the data at our disposal. We were inclined to believe, in accordance with other, already cited, works, that Erlang data were able to describe the variability of people's density in space and time, but we wanted to trace some more precise limit of this capacity, and we tried to do it, at a certain degree, in a quantitative way.
We were prepared to weak results, that is to obtain a very low upper bound for the limits we tried to outline.
The preliminary results, on the contrary, revealed a high level of correlation between the considered variables (which were restricted to the main municipalities).
This outcome can help us henceforward to better interpret the meaning of Erlang data.
Testing the conjecture is rather difficult, for lack of comparable traditional data sources. For the sake of comparison we considered as a benchmark an elaboration of the Lombardy Region Origin-Destination matrix (2002), which consists of a set of time-series of people's presence in the region, one for each municipality, with a rate of 1 hour.
The Origin – Destination Survey provides information on the population movements for the whole Lombardy Region during a working day in the 2002 year. In particular, the region has been divided in 1456 zones, corresponding to the administrative municipalities or to their aggregation. The survey conducted via CATI interviews, involved a sample of about 296.000 families, for an amount of more than 750.000 residents in the region: the average sampling rate for the whole region was 9,1%[5] (Regione Lombardia, 2002).
Results from the sample has been generalized to speak for the entire population using the most updated available sources and a unique source (Istat, 2001) for the Lombardy Region overall population.
This survey provides information on daily mobility of Lombardy Region residents during 24 hours of a typical working day, for whatever reason. In fact the reasons why people move daily are manifold: commuting, studying, meeting persons, shopping, business, tourism, visiting friends and relatives, doing personal activities.
Usually, mobility census counts only job and study trips. The O/D survey gave us the possibility to include in our analysis overall daily mobility (i.e. systematic and non-systematic).
Unlike traffic counts, which are a count of traffic along a particular road and which can be used for generating and updating origin destination matrix within a traffic model, the O/D survey concerns the mobility demand of people living in the Lombardy Region. It is therefore a source of data which is not easy to update because of high time and money costs necessary to obtain direct information from inhabitants via CATI interviews. Lombardy Region has not yet released a new version of the survey since 2002 and at the moment it is not expected an update for the next years at all. For these reasons, we decided to use this source of data, notwithstanding its temporal distance with available mobile phone data (2002 vs 2009).
We considered an elaboration (Manfredini, 2009) that obtained from the O/D Survey data, for each municipality and for each hour of the day, the number of exits (out - flows) and the number of entrances (in – flows): assuming that the population is equal to the residents during the night (in particular at 4 a.m.), it has been calculated, for each hour of the day, the quantity of present

[5] The sample was obtained through a geographical stratified cluster sampling, where clusters were the families and the 1456 geographical zones defined the strata (subgroups of the population). Sampling rates ranged from 3% to 10%, based on the number of residents of each zone.

population by adding the in-flows populations and subtracting the out-flows population. In this way it has been rebuilt the population size of each municipality for all the hours of the day.
Depending on the amount of trips leaving or entering the cities, it is possible to find on a hourly basis situations where the resident population decreases or where the present population is higher than the resident population.
We evaluated the correlations between mobile phone traffic[6] and people's presence, taking into account some further operation on the time-series, as reported in Table 1. The corresponding scatter plots of the two variables for all the municipalities are depicted in Figure 3.

Table 1 - Linear correlation between the time-series of cell phone traffic and those of people's presence in the provincial capitals of Lombardy Region municipality

municipality	ρ (Pearson's correlation coefficient) - original Time Series	ρ - moving average (3 terms) smoothing of the Time Series	ρ - moving average (3 terms) smoothing of the Time Series, time range 07:00-23:00
BERGAMO	0.82	0.92	0.94
BRESCIA	0.85	0.94	0.97
COMO	0.67	0.90	0.93
CREMONA	0.78	0.97	0.99
LECCO	0.76	0.89	0.92
LODI	0.83	0.93	0.98
MANTOVA	0.48	0.77	0.94
MILANO	-0.27	0.45	0.88
MONZA	-0.32	0.94	0.98
PAVIA	0.29	0.84	0.87
SONDRIO	0.74	0.86	0.92
VARESE	0.90	0.97	0.99

[6] More precisely we considered here, for each window of 15 minutes, the ratio between the number of Erlang per pixel of each urban context and that of the whole region. We then constructed the typical curve of the typical weekday considering, 15 minutes per 15 minutes, the average of these ratios. We finally aggregated these values hour by hour obtaining the same time granularity of the people's presence time series.

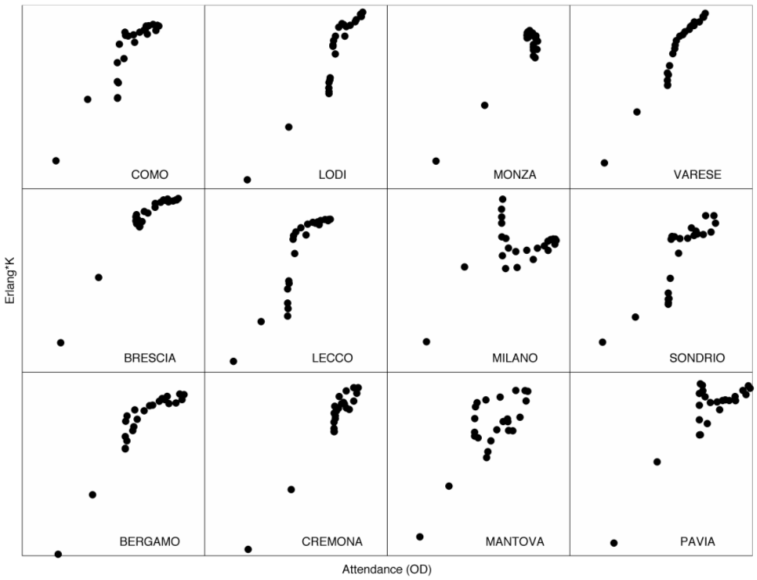


Figure 3 - Scatter plot of the time-series (ratio Erlang municipality/region VS presence of people). Note that isolated points mostly represent night measurements

Given the municipalities in the first column, the second column contains the Pearson's correlation coefficient[7] (ρ) of the original time-series. In the third column the correlation is calculated on a smoothing of the original series (moving average of three terms: past, present, future[8]). In the fourth column the same smoothing was considered, but restricting the series to the time range from 7:00 to 23:00, where the denominator of the telephonic series, the Erlang for the whole region, is more continuous than during the night, when lower values, near to zero, and the not homogeneous use of the mobile phone network across the region makes the curves less reliable.

Preliminary results seem to confirm a high level of correlation between these variables: correlation values are indeed globally very interesting. While the first comparison leads to high ρ just for few municipalities, the second presents values mostly around 0.8 and always more than 0.5, with the only exception of Milan.

Filtering the series reduces their local variability, and lets a more reliable evaluation of the trends, as we can clearly observe in Figure 4, where the curves of Milan (ρ=-0.2) and their smoothing (ρ=0.4) are plotted. Finally, the third comparison of the daytime part of the time series shows very high correlations for all the municipalities.

These results corroborate the hypothesis that cellular network data can be, at least at the municipality scale, an interesting indicator of the variability of people's presence.

[7] The Pearson's correlation coefficient is the covariance of two variables divided by the product of their standard deviations. It measures the degree of linear association between two variables. It is a dimensionless quantity and can take values between -1 and +1. It assumes positive values when the values of variables grow together and negative values when the values of a variable decrease while the values of the other increase. The correlation is "0" if the two variables are independent. It is important to note that the correlation coefficient is able to show whether two variables are related or not but not if a variable depends on another.

[8] Given a time-series F(t), $t=t_1 \ldots t_n$ the moving average gives us a new time-series G(t) defined as follows: $G(t_i)=(F(t_{i-i})+F(t_i)+F(t_{i+1}))/3$. Note that the drops of curves in Figure 4, right side, are due to the moving average filtering method: the first and the last terms are computed as the average of three terms, but just two are available at the beginning and the end.

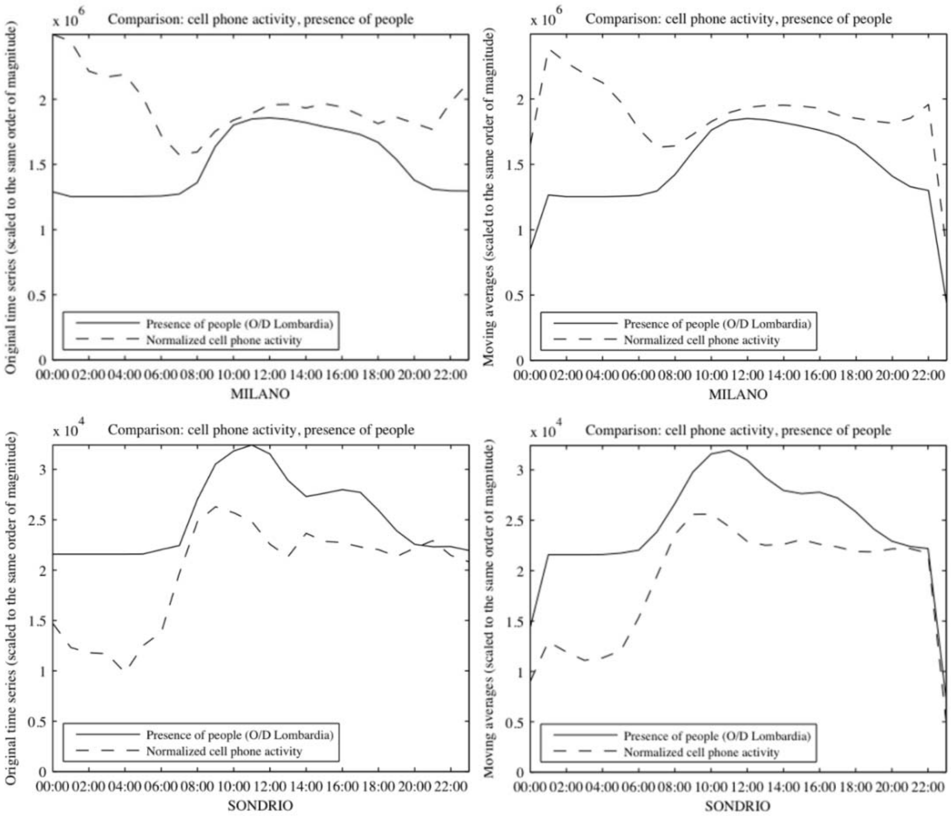


Figure 4- Plot of the time-series for two limit cases: the municipalities of Milan (top) and Sondrio (bottom). Telephonic activity (dashed) is compared to the presence of people (solid). The left side represents the original time-series, the right side the moving average smoothing. For the sake of comparison the cell phone activity trend is scaled to the same order of magnitude of the presence of people trend.

To conclude we want to point out some critical aspects of the experiment, concerning the data taken for the comparison. As was said, we took into account, as a benchmark, the presence of people derived from the origin-destination matrix of Lombardy Region.
The advantage of this choice is that the two datasets have corresponding time-spatial granularity. The disadvantage is the fact that they concern two very different periods (2002 vs 2009). In this sense we must highlight that the restriction of the analysis to the provincial capitals plays a key role in the experiment. In fact the observed features in these territories are stable enough, with respect to time, to make the comparison possible even with this time-lapse.
On the contrary, if we extend the analysis to minor municipalities, we encounter different problems, due to many factors: on one hand, on the front of cellular network, we have a narrower range of telephonic cells in the territory, which determines a reduced reliability of the data, on the other hand, concerning the time gap, it is more probable to come across situations in which transformations of the territory could have influenced its profile.
This extension to minor municipalities is a work in progress, and here we are trying to turn the problem inside out, testing whether the comparison of actual telephonic data with older traditional data sources, can help us to individuate changed situations of the territory. A possibility that, we think, could be very useful to many applications.

Pucci, Manfredini, Tagliolato | {*paola.pucci, fabio.manfredini, paolo.tagliolato*}*@polimi.it*

4.2 The micro scale analysis and its relevance to spatial analysis

Differently from other researches (e.g. Ratti et al., 2006) that analyze the Erlang trends at the municipality scale to describe correlation between cell phone data and people's daily activities (that are not the density of people, but the density of phone activities), we have instead planned an analysis at the micro scale.
This analysis starts from the identification of similar urban conditions (urban context type) – as social housing districts, railway stations, university districts, factory ambits, residential districts, business districts, areas of urban sprawl - selected on the basis of recognizable and well known land-use patterns, density of population or activities or socio-economic profiles.
The tests carried out have allowed us to verify that, for different urban context types, specific traffic curves correspond, and the variability in behaviors among different hours of the day, including weekdays and holidays within the same urban context type, corresponds. The comparison with demographic and socio-economic data has allowed then to analyze even the "anomalies" in the Erlang curves, recorded in the urban districts of the same urban context type and to recognize the significant socio-economic indicators useful inexplaining the correlation between statistical data and telephone trends.
We chose then a set of urban districts of each urban context type, and we studied the Erlang trends of these examples, in order to check for analogies in the cell phone activity of similar urban districts in term of density and land-use patterns. The aim is to identify, for each urban pattern, typical Erlang trends for a functional clustering of the region.
The first step of this analysis was to elaborate the typical daily curves (Figure 5) of mobile phone activity (Erlang from January to October 2009) to verify if urban districts with similar socio-economic profiles had similar cell phone activities trends. In addition we examined the variability of trends in working days and during holidays, for the same urban context type.
Micro-scale elaboration, by considering the correlations among Erlang curves (from 9.00 a.m. to 10.15 pm) of the different urban context types (Figure 6), let to the identification of:

- Homogeneous cell phone traffic trends within the majority of the considered urban context types (in particular within the industrial ambits, type 5, and the social housing districts, type 7);
- Similar cell phone traffic trends in different urban context types, suggesting affinities in their cell phone use practices (industrial ambits, type 5, and university districts, type 1);
- Different cell phone traffic trends in similar urban district types, as railway stations (type 3) and urban sprawl areas (type 4).

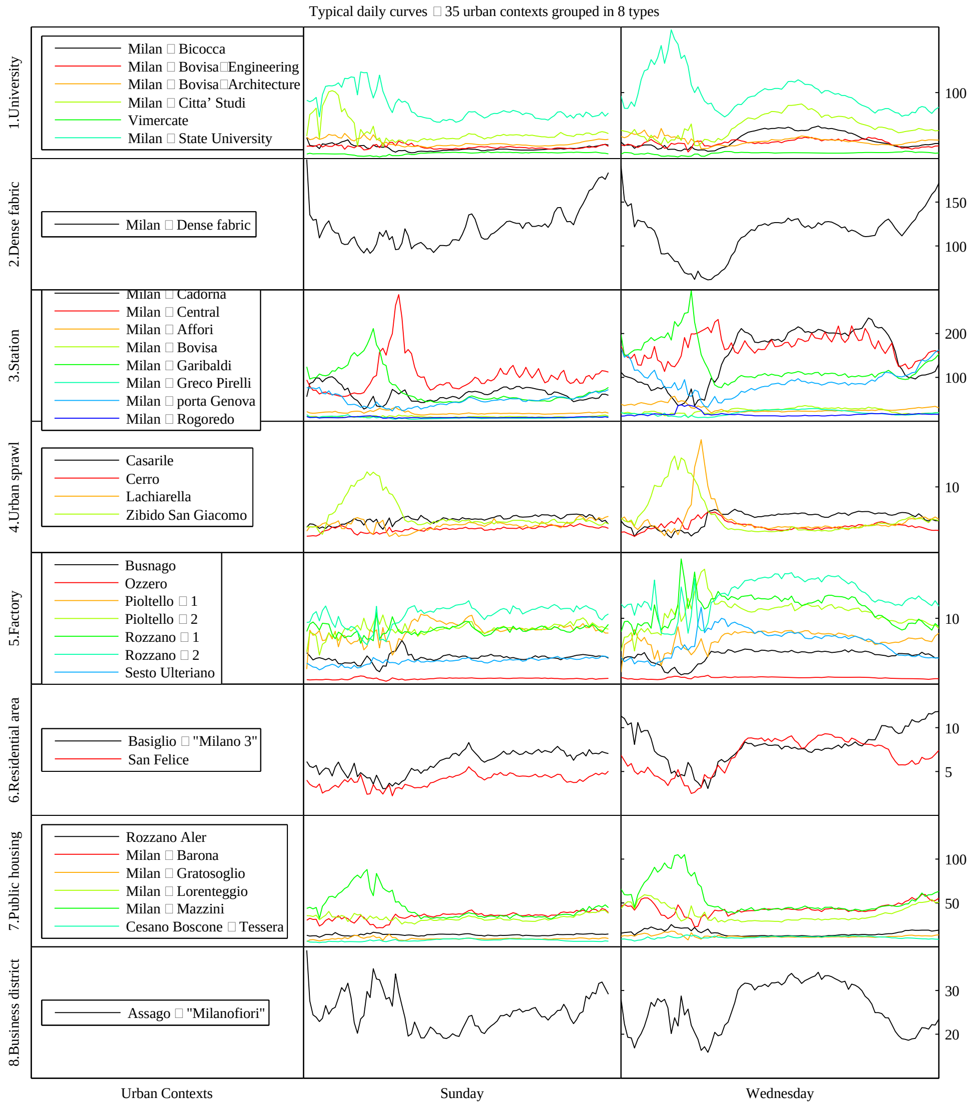


Figure 5 - Typical daily curves (Sundays on the left, Wednesday on the right) for 35 urban districts grouped in 8 urban context types.

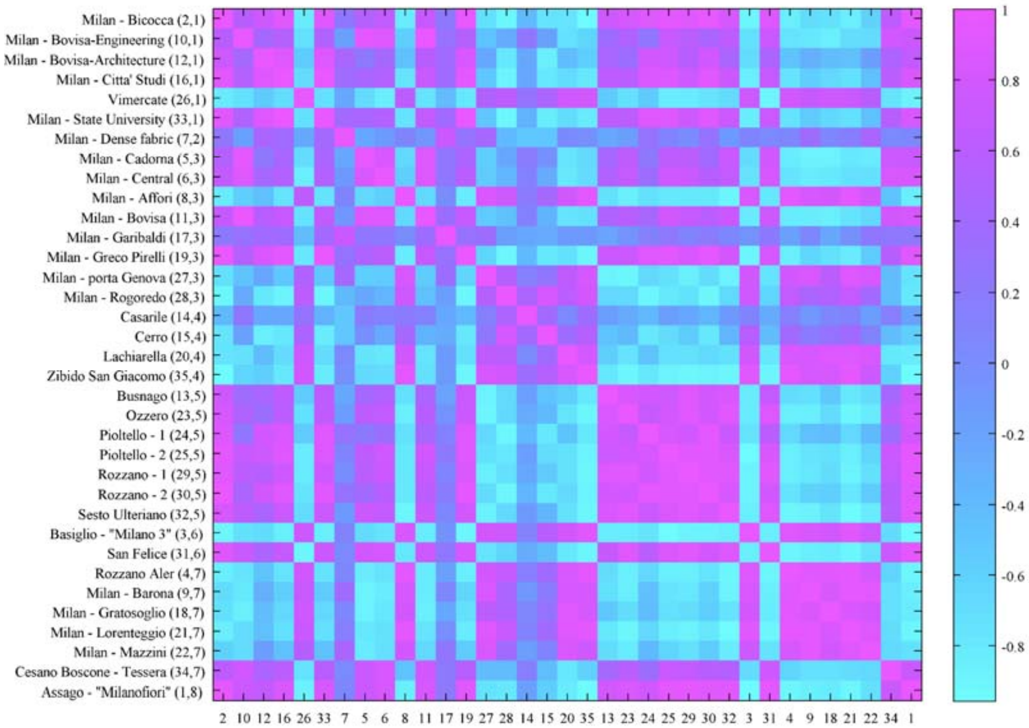


Figure 6 - Correlation matrix among different urban context types. Row labels represent: urban district's name, id of the district (from 1 to 35, the same used for columns), id of the context 's type (1: university; 2: dense urban fabric; 3: railway station; 4: areas of urban sprawl; 5: factories; 6: residential districts; 7: social housing districts; 8: business district). The analyzed 35 urban districts are ordered by type. If each type had its own pattern, we would see the emergence of pink (correlation values near 1.0) squared blocks along the principal diagonal of the matrix. The presence of blue pixels (correlation values near -1.0) along the diagonal reveals that some context does not behave the same way as the other contexts of the same type. Pink blocks far from the diagonal on the contrary indicate some "telephonic" analogy among urban districts of different type.

These situations (different mobile phone traffic trends in similar urban context types) were analyzed in depth, with demographic and socio-economic data, in order to recognize the significant variables, able to explain the anomalies.

In particular we observed for railway stations[9] a linear correlation between mobile phone traffic intensity and the quantity of people gotten on and off during the whole day. This is due to the different profile and role played by each station for the local passenger traffic in the Milanese Region.

Considering, as an example, two different profiles of stations as Milan Central Station – the main station for national and international railway connections - and Milan Cadorna, the commuter train station with an important role for the regional railway links (Figure 7), we can observe

[9] These stations, located in Milan, play both the roles of regional stations (for commuters flow) and nodes of national relations.

Pucci, Manfredini, Tagliolato | ***{paola.pucci, fabio.manfredini, paolo.tagliolato}@polimi.it***

that during weekdays the Central Station (grey line in figure) presents less telephone traffic than Cadorna Station (black line), while the opposite situation characterizes holidays[10].

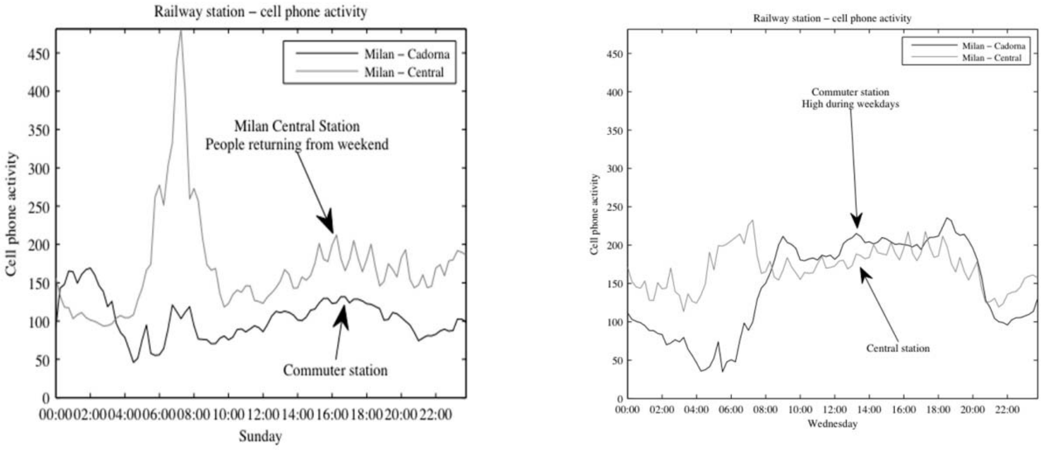


Figure 7 - Milan railway stations on Sunday (left) and Wednesday (right).

This confirms the prevailing role of Milan Central Station: rather a station for middle/long journeys, than one used by commuters. On the contrary, the considerable raise of the telephone traffic registered for Cadorna Station during working days, is consistent with data about people getting on and off, and corresponds to the key role of this station for the commuters' flow. For what concerns the other stations the distribution of telephone traffic is more regular and it is characterized by a greater intensity in weekdays than in holidays.
For the areas of urban sprawl (Figure 8), characterized by a strong functional specialization (residency), often accompanied by few or no services and by a standardized time of use, the comparison of the Erlang data between weekdays and holidays shows different behaviors of Sunday relative to Wednesday, in line with the profile of these areas, almost exclusively residential and with a high percentage of working population11 (Figure 9).

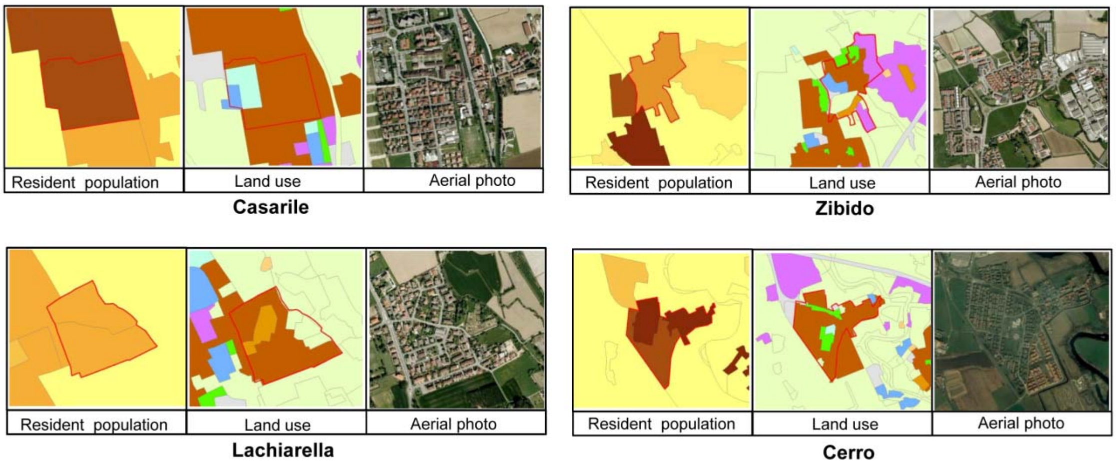


Figure 8 – Areas of urban sprawl

[10] The Wednesday peak in the Central Station from 5:00 to 7:00 can be explained by the corresponding low value of the denominator of the ratio station Erlang / Lombardy Erlang. When the total amount of telephone traffic is low (0:00-6:00), Erlang Lombardy curves are not very reliable.
11 During the day, the people generally move in the workplace or school, to come back in different time, according to age.

In the Sunday curves higher values of cell phone traffic are found, distributed fairly regularly throughout the day, although not related to the density of settlement. In fact, from the comparison between Erlang and statistical data[12], we can notice two interesting and "extreme" situations.

The least populous area among those considered (Zibido), is one that has a density of mobile phone traffic between the most relevant on Sunday, which significantly reduces on Wednesday. This trend may be related to the profile of this area, characterized more significantly than other areas, by the presence of young population in an active phase of its life cycle, by an equally high percentage of the population under 14 years , compared to residents over 55 years under-represented. The commuters flow data (working out) are the most significant among the considered areas (79.7% of all trips are out of the district).

The opposite situation - Cerro - is represented by the highest residential density in the considered urban context type, but with intensive cell phone traffic between Sunday and less significant, similar in intensity to other areas, on Wednesday.

These trends can also be traced back to the age profiles of residents (high density of population over 55 years) and to the component of outflows (69.2%), conditions that explain the working day Erlang trend, similar to that of less populated areas.

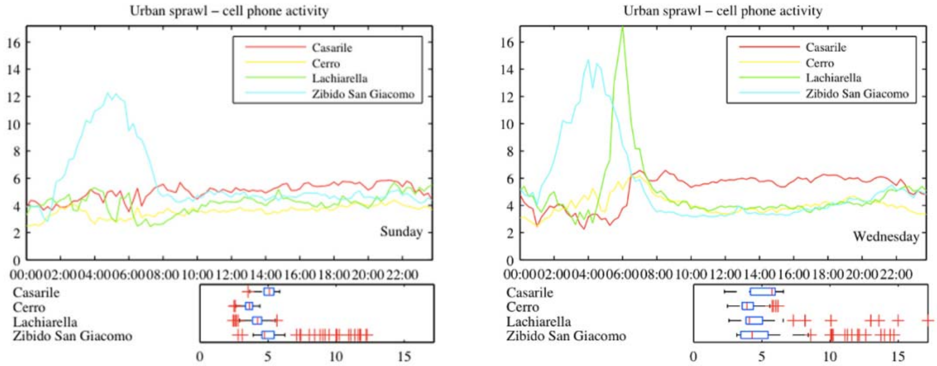


Figure 9 – Areas of urban sprawl on Sunday (left) and Wednesday (right).

The results lead us to consider two relevant insights, starting from two different perspectives:

- to propose a land-use classification, at the large scale, starting from the same kind of Erlang curves as a "signature" of each urban context type, chosen from the specific functional profile; in this case, since the mobile phone data would allow a functional description of the space, without the burden of a direct survey;
- to experiment a spatial clustering of Erlang data, finalized to identify some urban contexts characterized by homogeneous telephonic patterns, considered proxies of areas that are used at the same way by the population.

4.2.1. An international event transforming the use of the city : the Milan International Design Week

We also worked at the Milan urban scale, analyzing mobile phone data during the 2009 and 2010 International Design Week, a leading international event that lasts 5 days and concentrates its activities in the Rho-Pero Exhibition District and in hundreds of places within and outside the city (Fuori Salone).

[12] The socio-economic data for comparison with Erlang data, are: inhabitants by sex and age; socio-professional profile of population; foreign population; commuter flows; employees and economical activities.

These activities (presentations, exhibitions, pre-opening previews) modified even substantially the ways and the time in which some Milan urban districts were used by visitors and temporary populations.

Using Erlang data we made different elaborations for the period of the Milan International Design Week, aimed at mapping the spatial configuration of mobile phone traffic during the event and aimed at identifying which parts of the city show a significant concentration of activity, comparing the days when the event occurred with events-free days.

This information can be useful to quantify the consequences of a specific event on the whole urban system, to handle the event itself, to assess its impact on mobility, tourism and temporary population attractiveness and to eventually guide future decisions concerning the provision of new urban services.

We therefore performed several analyses to evaluate the potential contribution of Erlang data to describe, to represent and to manage an event, from the beginning until its conclusion.

We defined a set of significant spatial operations between Erlang matrices aimed at underlining the territorial effects of the event at a wider scale and at different temporal intervals.

The maps therefore represent the spatial configuration of cell phone traffic at different hours of the day and can be divided into two main categories:

- "Snapshots" maps that show cell phone traffic intensity at different hours of the day in which the event occurs. They represent the variation in urban uses by temporary population over time;
- "Complex" maps that display the relationship between two matrices of cell phone traffic, such as ratio between night-time and day-time mobile phone activity, ratio between weekdays and holidays, ratio between days with event and days without events.

Figure 10 therefore represents at a high level of spatial disaggregation and at the wider urban scale, the regional effects of an event that does not end within the perimeter of the Exhibition District, but rather affects significant portions of the urban region.

Pucci, Manfredini, Tagliolato | {*paola.pucci, fabio.manfredini, paolo.tagliolato}@polimi.it*

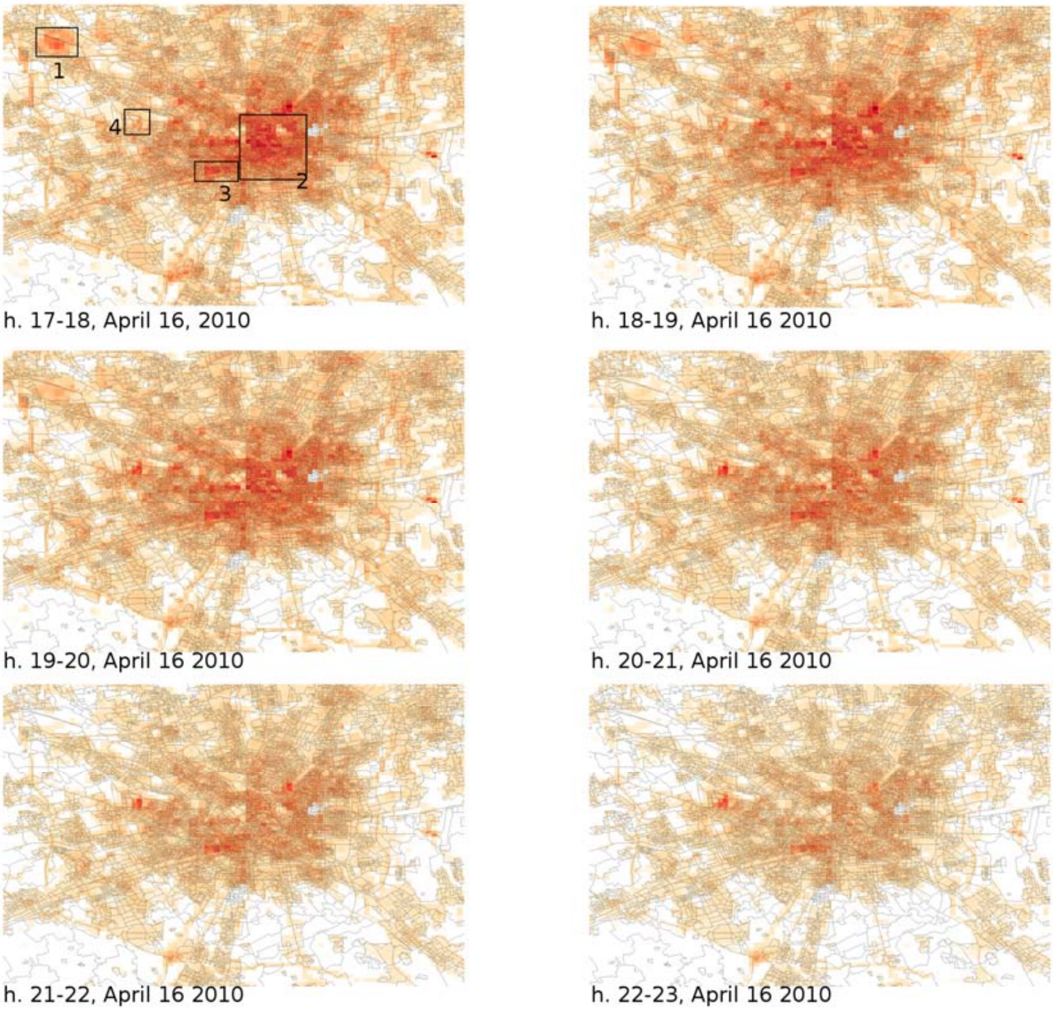


Figure 10 - The sequence of maps shows the cell phone traffic on 16 April 2010 from h. 17 to h. 23. Increasing intensity of red colour corresponds to higher cell phone traffic. It is possible to observe the parts of the city where phone activity is high, largely due to the International Design Week activities. Numbered boxes in the top-left frame highlight areas of particular interest: 1 = Exhibition District; 2 = Milan city centre; 3 = Tortona District; 4 = San Siro Stadium.

Thus, analyzing in detail the change of cell phone traffic activity during Friday, April 16, 2010, the day when, in addition to the exhibition in the Rho-Pero District in the Northern-Western side of Milan (box 1), dozens of activities (presentations, inaugurations, etc.) occurred in several areas of the municipality, it clearly appears that a substantial modification in density of use takes place. It can be observed that the centre of Milan empties gradually during the afternoon (box 2), while the Tortona district, in the southwest side (box 3), where the main "Fuorisalone" activities have been organized, remains crowded until late evening. During the evening, the area around the San Siro stadium (box 4) enlightens because of the major soccer league match between Inter-Juventus, which was held that day. The city centre, together with other districts, is therefore subject to major transformation in the manner and time of use.
We can deduce this phenomenon by observing the variability of cell phone traffic (Figure 11), which increases compared to a typical weekday. We can relate this dynamic with the presence of visitors and tourist, arrived in the city during the period of the International Design Week.

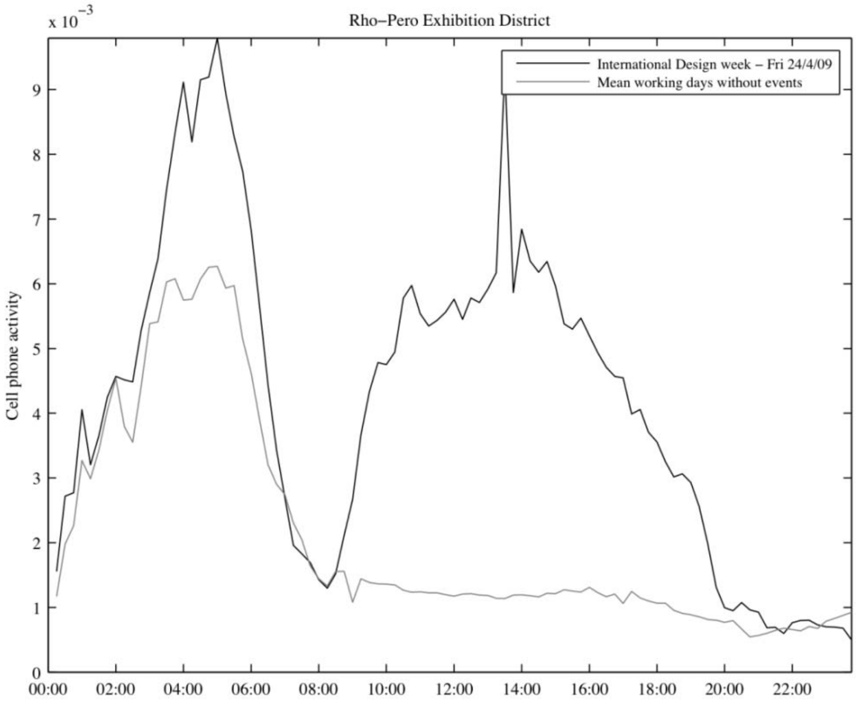


Figure 11 - Variability of cell phone traffic (Erlang) normalized by the total Lombardy value at the Rho-Pero Exhibition District during the 2009 International Design Week (black) and in a typical ferial day, without events (red).

The interpretation of maps, like those previously presented, is facilitated by specific knowledge on the activities established in the territory and by their integration with traditional databases, used in urban and regional studies (land use/land cover maps, transport networks, etc.).

These maps can provide new knowledge on urban dynamics at a spatial and temporal resolution not comparable with maps achievable by traditional data sources.

Further synthetic representations concerning the 2009 International Design Week, focused on the relationship between mobile phone activities during different days (Figure 12: ratio between night-time and day-time cell phone traffic; Figure 13: ratio between weekdays and holidays, ratio between days with event and days without events).

Data have proven to be effective in describing the temporal profile of the city, pointing out the areas with increased activity at night, as the Navigli area and the area around the Milan Central Station where there are pubs and nightclubs.

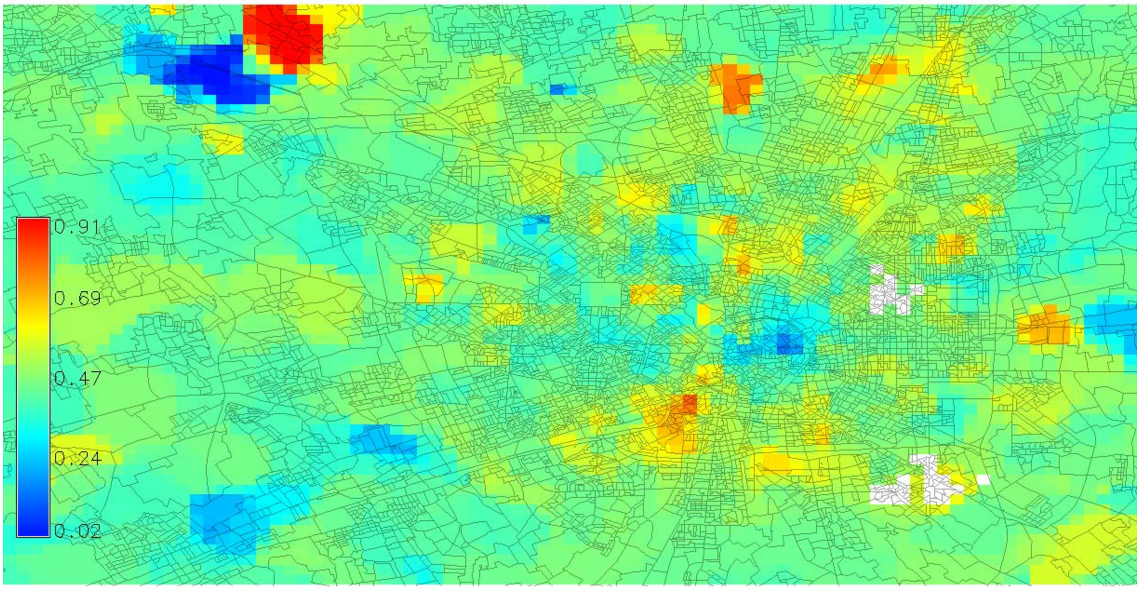


Figure 12 - Night / day cell phone traffic activity ratio during weekend in the period of the 2009 International Design week

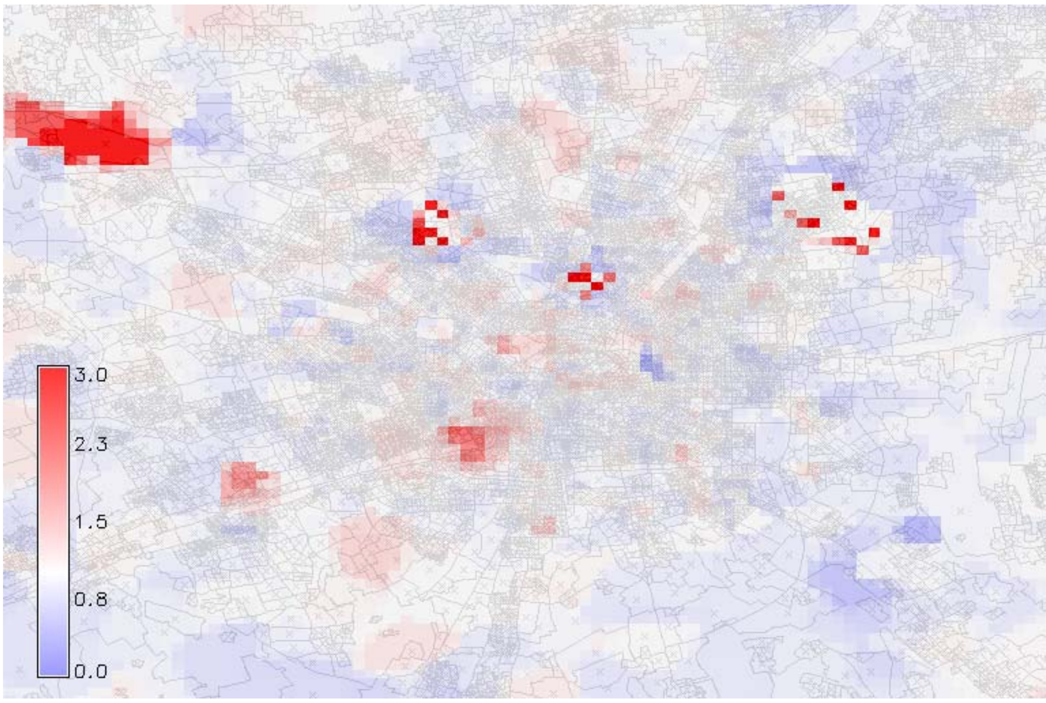


Figure 13 - Ratio between April 24, 2009 and April 29, 2009 cell phone traffic. The map highlights the parts of the city where the activity, during the International Design Week (April 24), was greater than during a typical ferial day (April 29).

Although some caution is needed in the interpretation, the maps confirmed a correspondence between the variability of cell phone traffic and the distribution of activities and initiatives in the area.

These data can therefore help to build a descriptive and an interpretative model of events through the analysis and mapping of cell phone traffic matrices. This model can take into account what happens when the event occurs, even in real time, comparing it to typical spatial patterns.
Urban planning traditional data sources are mainly based on static statistical surveys and are not able to catch the variability in the intensity of urban space's use by present population.
This type of information, if suitably interpreted, may be useful to assess the consequences of a specific event on the entire urban system in its spatial and temporal patterns and to evaluate its impacts on the whole urban system (mobility, congestion, tourism).

4.2.2. Integrating cell phone traffic data with statistical data

Further investigation focused on the correlation between the intensity of telephone calls at certain times of the day with the spatial configuration of residents and workers in the Milan area.
With the aim to correlate traditional indicators, typically used in urban studies, with cell phone traffic data, we calculated two simple indicators related to population density (residents per square km) and to employees density (employees per square km). These indicators, although are dated as they relate to the year 2001, represent the variability of the concentration of population and activities as it can be drawn from traditional data sources (Italian Census).
The correlation matrix was then calculated on an hourly basis, for April 29 2009, considering two different geographical areas: a large area, which includes Milan, the Rho-Pero exhibition district and the first ring of municipalities around Milan and a restricted area that coincides with the centre of Milan.
The results are shown in Table 2 and in Figure 14. During the working hours, approximately from 8:00 to 20:00, the correlation between cell phone traffic and employees density is much higher than that with population density. The trend highlights a rapid increase since early morning until it reaches the maximum (0.51) around h. 12:00. On the contrary, focusing on population density, we observe a weak decrease of the correlation, after h. 8:00 and a subsequent rise in the afternoon (h. 16:00) when people is supposed to start returning home after work. The maximum is around 21:00 when the value reaches 0.42.

Table 2 - correlation matrix between traditional indicators (population density and employees density) and cell phone traffic (Erlang) at different hours at the Milan urban region scale (left) and at the Milan city scale (right) (April 29, 2009).

	Milan Urban Region		Milan core city	
	Population density (residents per sq km)	**Employees density (employees per sq km)**	**Population density (residents per sq km)**	**Employees density (employees per sq km)**
00:00	0.35	0.28	0.15	0.12
01:00	0.30	0.27	0.13	0.14
02:00	0.19	0.16	0.10	0.06
03:00	0.12	0.12	0.08	0.09
04:00	0.10	0.06	0.07	-0.03
05:00	0.07	0.05	0.04	-0.03
06:00	0.10	0.10	0.00	0.01
07:00	0.26	0.28	0.00	0.15

08:00	0.31	0.40	0.01	0.29
09:00	0.29	0.46	-0.02	0.36
10:00	0.28	0.48	-0.03	0.39
11:00	0.27	0.50	-0.03	0.42
12:00	0.27	0.51	-0.04	0.44
13:00	0.27	0.51	-0.04	0.44
14:00	0.27	0.51	-0.03	0.44
15:00	0.29	0.50	-0.01	0.42
16:00	0.29	0.49	-0.01	0.41
17:00	0.30	0.47	0.01	0.38
18:00	0.32	0.45	0.03	0.35
19:00	0.35	0.42	0.07	0.30
20:00	0.38	0.40	0.11	0.27
21:00	0.42	0.33	0.19	0.17
22:00	0.41	0.32	0.17	0.15
23:00	0.40	0.32	0.18	0.13

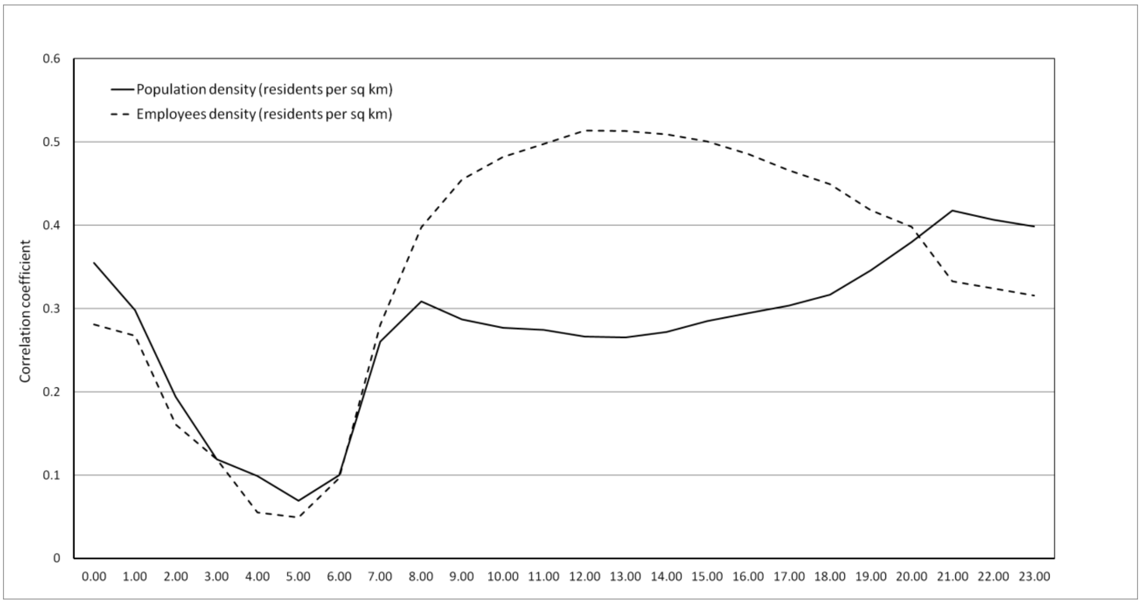


Figure 14 – Trends of correlation coefficient between cell phone traffic and population density / employees density on an hourly basis for the April 29 2009 at the Milan Urban Region scale.

A first result of the analysis, showed, therefore a good correlation between the matrices of cell phone traffic, the density of employees (0.57) and the density of residents (0.43) as well as a trend of the correlation which is consistent with the spatial distribution of jobs and homes in the Milan area.

The study on the central area of Milan gives us new insights to assess the utility of integrating traditional urban data with cell phone traffic data. In fact, in this case, the correlation coefficient with the density of residents is very low. It is always around 0 and only during late afternoon it reaches higher values (0.17). This result may be interpreted by the changes that occurred in the city centre of Milan, where a gradual replacement of the residents with activities mainly related to the services sector, happened in last decades.

The correlation index between cell phone activity and employees density is always lower than that at the larger urban scale, despite in the central hours of the day (from 11:00 to 16:00), when

it achieves values over 0.4. The Milan core city area is characterized by a distribution of activities and attractors (i.e. shops, museums, tourist sites, and workplaces), that cannot be detected by the employees density indicator.
Preliminary results showed therefore that cell phone traffic data could effectively help to represent and to describe, dynamically over time, the intensity of activities and of presences at the urban scale.

5. Conclusions and further research

Traditional data sources for urban and mobility investigations (i.e. surveys, census) have some known limitations, including the high cost of surveys, the difficulty of data updating, the difficulty of describing city dynamics and time dependent variations in intensity of urban spaces usage by temporary populations[13] at different scales.
In sum we can refer to three issues for which data normally used in urban analysis and planning have significant limitations.
A first consideration regards the traditional sources for the analysis of daily mobility, as an information through which it is possible to interpret the daily urban practices.
The second issue refers to the definition of the geographical scale, the most relevant to analyze urban transformations. The re-scaling of data sources requires more flexible data and tools able to intercept urban phenomena in their correct spatial dimension.
Thirdly, the temporal dimension is virtually absent from traditional surveys and investigations, despite the fact that it is one of the most significant components of contemporary cities.
In reference to those limitations in traditional data, the preliminary findings of our study suggest that cell phone data constitute an interesting and unique source of information on urban uses.
Because of its spatial and temporal resolution, mobile phone data can describe time-dependent variation that is missing from traditional analysis and could thus describe cities dynamically over time. If we consider the observed and aggregated telephone traffic as the result of individual behaviours and habits, we conclude that mobile phone data can provide useful information on temporary populations, which are difficult to intercept by traditional data sources, but which, at the same time, increasingly affect urban practices both quantitatively and qualitatively.
In fact, the diffusion of individual and especially portable communication tools, like cell phones, has recently given a new impetus to the strongly debated topic of the relationships between the diffusion of communication tools and the physical mobility of individuals in the fields of transportation economy, geography and sociology.
Many researchers recognized that the use of mobile technologies "may lead to changes in location, timing and duration of people activities [...] and will be associated with new patterns of activity and travel in space time" (Kwan, 2007).
In order to describe these changes, several studies exploited tracking technologies, used also to provide information about the users and to develop commercial applications[14]. The review demonstrates that the majority of this works is concerned with explorative empirical research or with the investigation of alternative theories to model such behaviours. These works are mainly based on the analysis of individual traces of a sample of users (active mobile positioning).

13 Time-related changes in the city can be described through the concept of "temporary populations". The urban population can no longer be regarded as a collection of residents tied to move within the boundaries of the city.

14 Google has published Google Transit and Google Maps Navigation. Companies including Tomtom and Navman are planning to provide real-time personal navigation devices and Garmin and Polar both provide such devices for a number of sports activities.

Considerably less is known about the links between ICTs, individuals' activity, schedules and their deployment in time and in space (Schwanen and Kwan, 2008).

Other researches, more relevant to the present study, concern the way in which the anonymous mobile network data can represent a potential tool in reference to a wide variety of applications in the monitoring of urban dynamics.

In this field, the applications are focused on two different products:

- some studies deal with aspects of data representation, emphasizing the aspects most directly evocative and highlighting how these data may represent "Mobile landscapes" (Ratti et al., 2006);
- other studies focus on the development of data-mining methods to manage large amounts of data, to create new tools capable of deriving summary information and relevant data on cell phone (Ahas and Mark, 2005).

Instead, in this research we concentrated on the potential impact of application of these data, exploring the more technical features, and then suggesting different uses, depending on a variety of applications that currently do not have appropriate responses in traditional database used in urban studies.

The first conclusion is that the different types of cell-phone data used, integrated with each other and inserted into a urban information system, can help to provide knowledge on new urban dynamics.

This also emerged from the interviews with some institutional players and researchers, with whom we were able to confront both at the beginning and at the final phase of our research.

The most interesting aspects are summarized in table 3.

Each of the experiments carried out, has opened new questions and research perspectives.

We are currently working on the topic of spatial clustering. The aim of this line of research is to profile the region, described in terms of cell phone data, in order to study the usage of the city, by means of synthetic Erlang trends, able to characterize prototypical temporal patterns.

Because of its spatial and temporal resolution, cell phone data constitute an interesting and unique source of information on urban uses. Indeed, if we consider the observed and aggregated cell phone traffic as the result of individual behaviours and habits, we conclude that cell phone data can provide information, which changes over time, on urban contexts, and can bring to new interpretation of urban dynamics.

Table 3 - Possible applications of cell-phone data

Experiments	**Description of Analysis**	**Application**
Erlang: Micro Scale	Variability in cell phone traffic in different parts of the city during a typical weekday and a typical day of a weekend.	Validation of the relevance of the data in Erlang expressing the variability of the uses in different urban sections of a city. Identification of the most significant variables able to explain the Erlang trends. Using profiles for defining the different urban typologies, for mapping the entire urban region and for proposing a clustering of urban districts.
Erlang: International Design Week	Maps of cell phone traffic during a big event and during days without events, correlation of these data with urban indicators.	Monitoring and management of big events, variation during days/nights, concentration of activity. Provision of services to the users.

Erlang: Macro	Correlation of the variability of population presence with cell phone activity.	Validation of the aggregated data from Erlang, for monitoring presences at the municipality scale.

Acknowledgements
This project has been financed by Telecom Italia during the year 2009. The authors acknowledge the collaboration of Paolo Beria, Paolo Dilda, researchers at the Department of Architecture and Planning, Politecnico di Milano.

References

Acebillo, J., Martinelli A., 2012. New urban metabolism, Actar, Barcelona.

Aguilera, A., Guillot, C. and Bonin, O., 2009. Mobile phone use and the management of individual reachability. Cost 298 conference, Copenhagen. Available at: http://www.abs-center.si/gbccd/papers/P062.pdf

Ahas, R., Aasa, A., Silm, S. and Tiru, M., 2010. Daily rhythms of suburban commuters' movements in the Tallin metropolitan area: case study with mobile positioning data. Transportation Research Part C: Emerging Technologies 18(1), 45–54. DOI: 10.1016/j.trc.2009.04.011

Ahas, R., Mark, Ü., 2005. Location based services–new challenges for planning and public administration?. Futures 37(6), 547–561.

Arai, Y., 2006. Geolocation technology and local information in mobile telephony. Networks and Communication Studies 20(1-2), 9–25.

Bekhor, S., Hirsh, M., Nimre, S. and Feldman, I., 2008. Identifying spatial and temporal congestion characteristics using passive mobile phone data. Transportation Research Board 87th Annual Meeting 100(08-1534), 22–25. Available at: http://trid.trb.org/view.aspx?id=848079

Brunner, P. H,, 2007. Reshaping Urban Metabolism. Journal of Industrial Ecology, 11(2), 11-13.

Caceres, N., Castillo, Romero, L. M., Benitez, F. G. and Del Castillo, J. M., 2010. Traffic flow estimates inferred from mobile phone networks. 12th World Conference for Transportation Research. World Conference on Transport Research Society. Lisbon. pp. 1–24.

Caceres, N., Wideberg, J. and Benitez, F., 2007. Deriving origin destination data from a mobile phone network. Intelligent Transport Systems, IET 1(1), 15–26. DOI: 10.1049/iet-its:20060020

Central Intelligence Agency, 2013. The World Factbook 2009-10. Washington, DC. Available at:https://www.cia.gov/library/publications/the-world-factbook/index.html

Fontaine, M. and Smith, B., 2005. Part 1: Freeway operations: Probe-based traffic monitoring systems with wireless location technology: An investigation of the relationship between system design and effectiveness. Transportation Research Record: Journal of the Transportation Research Board 1925(-1), 2–11. DOI: 10.3141/1925-01

Gonzalez, M. C., Hidalgo, C. A. and Barabási, A.-L., 2008. Understanding individual human mobility patterns. Nature 453(7196), 779–782. DOI: 10.1038/nature06958

Höpfner, M., Lemmer, K. and Ehrenpfordt, I., 2007. Cellular data for traffic management—first results of a field test. ITS Europe Conference. number 2407.

ISTAT, 2001. 14° Censimento Generale della Popolazione e delle Abitazioni. Available at: http://dawinci.istat.it/

Kwan, M.-P., Dijst, M. and Schwanen, T., 2007. The interaction between ict and human activity-travel behavior. Transportation Research Part A: Policy and Practice 41(2), 121–124. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0965856406000255

Manfredini, F., 2009. Spatial and temporal distribution of present population in the Milan area. in M. Caglioni and F. Scarlatti (eds), Representation of Geographical Information in Planning. Società Editrice Esculapio. Bologna. pp. 119–127.

Manfredini, F., Pucci, P. and Tagliolato, P., 2012. Mobile phone network data. new sources for urban studies?. In: G. Borruso, S. Bertazzon, A. Favretto, B. Murgante and C. M. Torre (eds), Geographic Information Analysis for Sustainable Development and Economic Planning: New Technologies. IGI Global. Hershey PA, USA.

Pucci, P., Tagliolato, P. and Manfredini, F., 2010. Uncovering urban dynamics from mobile phone data. Proc. of 24th AESOP annual conference. Aalto University School of Science and Technology. Aalto, Finland.

Qiu, Z. and Cheng, P., 2007. State of the art and practice: cellular probe technology applied in advanced traveler information system. 86th Annual Meeting of the Transportation Research Board, Washington, DC. number 0223.

Ramm, K. and Schwieger, V., 2007. Mobile positioning for traffic state acquisition. Journal of Location Based Services 1(2), 133–144. Available at: http://www.cirgeo.unipd.it/cirgeo/convegni/mmt2007/proceedings/papers/ramm_katrin.pdf

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

Reades, J., Calabrese, F., Sevtsuk, A. and Ratti, C., 2007. Cellular census: Explorations in urban data collection. IEEE Pervasive Computing 6(3), 30– 38. DOI: 10.1109/MPRV.2007.53

Regione Lombardia, Direzione Generale Infrastrutture e mobilità, 2002. Indagine origine/destinazione regionale 2002 - sintesi. Technical report. Regione Lombardia.

Schwanen, T. and Kwan, M., 2008. The internet, mobile phone and space-time constraints. Geoforum 39(3), 1362–1377.

Sevtsuk, A. and Ratti, C., 2010. Does urban mobility have a daily routine? learning from the aggregate data of mobile networks. Journal of Urban Technology 17(1), 41–60.

Song, C., Qu, Z., Blumm, N. and Barabási, A.-L., 2010. Limits of predictability in human mobility. Science 327(5968), 1018.

Thiessenhusenk, K. U., Schafer, R. P. and Lang, T., 2006. Traffic data from cell phones: a comparison with loops and probe vehicle data. ITS World Congress. number 1550. UMTS World. UMTS Network Coverage Planning. Available at: http://www.umtsworld.com/technology/coverage.htm

Wolman, A. 1965. The metabolism of cities. Scientific American 213(3), 179–190.