MIT Open Access Articles

Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome

Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome

Francesco Calabrese, *Member, IEEE*, Massimo Colonna, Piero Lovisolo, Dario Parata, and Carlo Ratti

Abstract—This paper describes a new real-time urban monitoring system. The system uses the Localizing and Handling Network Event Systems (LocHNESs) platform developed by Telecom Italia for the real-time evaluation of urban dynamics based on the anonymous monitoring of mobile cellular networks. In addition, data are supplemented based on the instantaneous positioning of buses and taxis to provide information about urban mobility in real time, ranging from traffic conditions to the movements of pedestrians throughout the city. This system was exhibited at the Tenth International Architecture Exhibition of the Venice Biennale. It marks the unprecedented monitoring of a large urban area, which covered most of the city of Rome, in real time using a variety of sensing systems and will hopefully open the way to a new paradigm of understanding and optimizing urban dynamics.

Index Terms—Cell phone network, Global Positioning System (GPS), Global System for Mobile Communications (GSM), intelligent transportation system (ITS), location-based services (LBSs), real-time control systems, wireless locationing.

I. INTRODUCTION

AN INTELLIGENT transportation system (ITS) is a transportation system that makes use of information and communication technology (ICT) to address and alleviate transportation and congestion problems. In general, an ITS relies on location-based information: It monitors and processes the location of a certain number of vehicles (used as probes) to obtain information on estimated travel time, driving conditions, and traffic incidents. Using a relatively large amount of probes, the early stages of bottlenecks can be detected, and traffic can be directed to other routes to mitigate congestion and provide more expedient and efficient itineraries to travelers. A variety of sensors can be used to obtain traffic information. These traditionally fall into two main categories: fixed sensors and Global Positioning System (GPS) receivers.

Fixed sensors, which can be either hard-wired or wireless, are used outdoors on roads and can provide information on the speed at which vehicles are traveling, as well as the capacity of a given road. They can also gather information on the type of

vehicle, distance between vehicles, pollution conditions, road surface conditions, and many others. Fixed sensors that are inductive, piezoelectric, or magnetic can be placed under the road, and those that use radar, ultrasound, or infrared with pyroelectric effect can be placed beside it. Each sensor is equipped with a control unit, battery system, solar panel, and transmission system.

GPS receivers are used inside vehicles and periodically transmit information on a vehicle's location and speed, together with the associated GPS time, to a central system. These receivers are composed of a control unit and a transmission system, which is based on text messages and/or the transmission of data packets. As a result of communication delays and data packet losses, GPS data need to be carefully processed to extract accurate location data that correspond with the coordinated universal time. If vehicle density within a specific area is quite high, an appropriate processing algorithm in the central system is capable of generating an excellent estimation of automobile traffic.

Both fixed sensors and GPS receivers contain limitations. The installation of wired fixed sensors is costly, as they require a cable network to transmit information to the central collecting point. Moreover, they provide only point-based information on traffic conditions, and to gain a realistic and complete view of automobile traffic in a certain area, a large quantity of sensors must be installed.

While GPS receivers provide better data, they also pose significant problems. If each street is to have sufficient sensing capability, a significant percentage of vehicles must be equipped with GPS devices. Mass deployment is expensive, and transmission cost to the central system can become onerous if frequency is high. Furthermore, as a result of signal multipath and urban canyon obstructions, GPSs do not work well in urban areas. Finally, privacy legislation requires careful contractual work, as vehicle owners must explicitly agree to share their location for the purpose of automobile traffic estimation.

All of the above limitations have inhibited the implementation of sufficient monitoring systems, particularly for larger urban areas consisting primarily of smaller streets. Sensor-based systems have mostly been used for the monitoring of highways, major urban streets, and large intersections. One solution to reduce the cost of initial investment and maintenance would be to deploy wireless fixed sensors, which use wireless links to transmit data that they collect to a central processing server. Other solutions involve less conventional approaches.

In the past few years, research activities have begun to explore the usage of data obtained from mobile cellular networks. Mobile phone positioning techniques generally provide

less accuracy than the GPS, but the wide diffusion of mobile equipment (ME) in addition to the widespread installation of radio transmitters, or base transceiver stations (BTSs), in both urban and rural areas makes such positioning techniques very appealing. ME location data are abundant; opportunely aggregated and processed, they can effectively be used to understand driving behaviors, identify areas of congestion, and many others. Mobile phones are, in short, a wide-area sensor network, the data from which could complement and integrate with those coming from traditional sensor networks.

Several projects regarding feasibility and field tests of ME location-based ITSs have been developed in Europe (e.g., France [1] and U.K. [2]) and North America (e.g., California [3], [4] and Washington, DC [5]). Bolla and Davoli [6] presented a model for estimating traffic by means of an algorithm that calculates traffic parameters on the basis of mobile phone location data. Lovell [7] and Smith *et al.* [8] carried out various studies on the basis of anonymous data on the position of mobile phones with a view to transport applications such as journey times or speeds. Cayford and Johnson [9] analyzed the main parameters to be taken into account, namely, precision, metering frequency, and the number of localizations necessary to achieve accurate traffic description. Transportation applications have been studied for freeway [10], arterial traffic [4], and even origin–destination flow estimation [11]. Moreover, several companies worldwide, including ITIS Holdings (U.K.), Delcan (Canada), CellInt (Israel), as well as AirSage and IntelliOne (United States), have begun developing commercial applications. Most of these, however, have focused on and proven the feasibility of an ME location-based monitoring system for applications such as travel-time estimation or incident detection only along highways, where the implementation is simpler because of well-defined traffic paths. The effects of localization error can be at least partially contained when traffic is constrained by the highway network; even very simple location systems based on the position of the closest cell, i.e., cell-ID positioning, can be used. Other studies on arterial traffic [4] have used GPS-equipped smartphones to alleviate this problem, but they rely on people voluntarily downloading and running an application on their phone.

In this paper, we present the Real Time Rome project, which aims to show how ME location-based monitoring could be extended to a whole city: in this case Rome, Italy. A few related projects have been proposed in recent past. The Amsterdam Real Time project[1] aimed to construct a dynamic map of Amsterdam, The Netherlands, based solely on the movement of a selected number of people carrying GPS receivers and being tracked in real time. A real-time cell-phone monitoring system was developed for the "Graz in Real Time" project[2] based on cell-phone traffic intensity, traffic migration (handovers), and traces of registered users as they moved through the city. Unlike previous attempts using solely the GPS or aggregated mobile phone data, this project, which has been developed by the Senseable City Laboratory, Massachusetts Institute of Technology (MIT), for the Tenth International Architecture Exhibition of the Venice Biennale, in collaboration with the Italian mobile phone carrier Telecom Italia, aims to create an integrated approach to urban monitoring through the use of several key components and activities:

- high-resolution and high-definition data over extensive urban areas, the collection of which has been made possible by Telecom Italia's innovative Localizing and Handling Network Event Systems (LocHNESs) software platform;
- the monitoring of a very large portion of Rome with its highly complex network of streets;
- the integration of mobile phone network data with other types of real-time information, e.g., the position of taxis and buses.

The LocHNESs platform, through the use of several probes, extracts all the signaling messages traveling on the A-bis interface and containing the set of measurements made by all active MEs (i.e., MEs involved in a call). It immediately processes each message to derive the actual position of the ME and its speed to produce the traffic map. This process requires only a few seconds of processing time. The information is then merged with the real-time location of buses and taxis and, furthermore, with mobile phone activity in the city that is measured by the BTSs. The platform has potential applications in a variety of areas, including urban management, route planning, travel-time estimation, emergency detection, and general traffic monitoring. The choice of the application allows us to determine an appropriate refresh time to capture the dynamics in the city using the available sensors.

The aim of the project is to make a proof of concept for the Venice Biennale, and as such, it focuses more on artistic visualization than on the derivation of quantitative information and its use in the city in real time. However, this project clearly shows that an integrated approach can open a new way to monitor urban dynamics in real time and, more generally, develop a real-time control system for cities.

This paper is organized as follows: In Section II, we describe the LocHNESs platform developed by Telecom Italia and the data it provides. Section III describes the Real Time Rome application and the location data acquired for the project, and Section IV describes the developed system architecture. In Section V, we provide a detailed description of the developed processing and visual software, and in Section VI, we offer several conclusions.

II. LOCATION DATA FROM THE TELECOM ITALIA PLATFORM

A. LocHNESs Platform

LocHNESs is a software platform developed by Telecom Italia for the evaluation of statistics, such as real-time road traffic estimation, based on the anonymous monitoring of the ME movements.

The LocHNESs platform is based on the localization of events that occur on the mobile network [call in progress, short message service (SMS) sending, handover, etc.] through

---

[1]See http://realtime.waag.org/.
[2]See http://senseable.mit.edu/grazrealtime.

external probes. These probes are installed on the A-bis interfaces, i.e., the interfaces that link the BTS to the base station controller (BSC). These probes analyze all the signaling messages and send a notification of the detected events to the LocHNESs platform. The key data detected by LocHNESs through the A-bis interface are measurement result [12] messages, which are used to report the results of radio channel measurements made by the BTS (uplink measurements) and the measurement reports received by the BTS from the ME (downlink measurements)[3] to the BSC. The measurement result message contains Global System for Mobile Communications parameters such as the average signal quality (RXQUAL) as measured by both ME and BTS, the received signal strength (RXLEV) as measured by the BTS (uplink measurement), the received signal strength on the serving BTS and on the neighboring BTSs as measured by the ME (downlink measurement), and the actual timing advance (TA).[4] The measurement result message related to each active connection (ME in the state "connected") is sent to the BSC every 480 ms, allowing LocHNESs to determine the complete trajectory of the call with the same time resolution. To reduce the computational load of the platform, however, the number of events that are notified to LocHNESs for each call is reduced by the probes according to a fixed sampling ratio (for example, 1 : 10, i.e., with a time resolution of 4.8 s). Using the above data, the LocHNESs platform produces aggregated traffic maps in raster form: The area under analysis is split into a number of contiguous square pixels of varying sizes (typically, 250×250 m in urban areas and 500×500 m in extra urban areas). For each pixel, the platform estimates a number of parameters, such as the average speed in the four quadrants (northwest, northeast, southeast, and southwest) of a Cartesian reference system centered in the center of the pixel, the total average speed, the number of moving users, etc. To have real-time applications for vehicular traffic, these traffic maps are constantly updated with a given periodicity (for instance, every 5 min).

Other network nodes and interfaces allow the extraction of mobile phone position information. For example, by analyzing the signaling messages traveling on the A interface, i.e., the interfaces that link the BSC to the MSC, it is possible to know the cell where a phone call has started or the set of cells where handovers have occurred during a call or the location area (LA)[5] where the mobile phone, both active or idle, is located.[6] These position data could be used to derive information about people movements similar to those contained in the origin–destination matrices but are quite coarse for the traffic monitoring of a whole city because each cell, with a radius of some hundreds
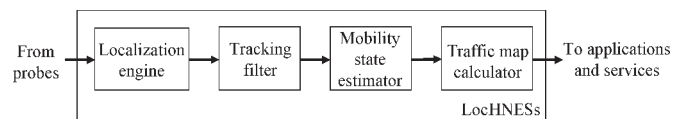

Fig. 1. Functional structure of the LocHNESs platform used for the Real Time Rome project.

of meters, can cover many streets, an LA can cover a great portion of the same city, and to detect ME movements, longer calls are required with at least one handover or one LA change, respectively.

It is important to note that the LocHNESs platform complies with the 2002 Directive by the European Parliament and Council on privacy.[7] Since the managed data are readily anonymized and then aggregate, at no time could individual users be identified based on the collected and analyzed data. We hope that this project might stimulate a dialog on the responsible access to locational data and on how it could provide value-added services, such as traffic monitoring, to local and regional communities.

The functional elements that constitute the LocHNESs are presented in Fig. 1 and will be described in more detail in the following.

B. Localization Engine

The localization engine estimates the instantaneous position of each ME involved in a call using the data extracted from the measurement result messages, which are received from the probes. Location is calculated using an Enhanced Cell ID with TA algorithm (E-CI+TA) [13], named data fit location (DFL); its principal components are the following:

- *Network database*—It is a database that contains all the parameters coming from the planning and dimensioning process of the entire mobile network (cell identifiers, i.e., Cell Global Identity, Base Station Identity Code, and the number of broadcast control channel carriers, BTS latitude, longitude, and height, BTS antenna azimuth and tilt, BTS transmission power and losses, etc.).
- *Antenna database*—It is a database that contains the complete radiation patterns (both in the H and V plane) of all the antennas that are mounted on the BTSs.
- *Propagation model*—It allows calculating the mean received power as a function of different parameters such as the operating frequency, the ME–BTS distance, the ME and BTS heights above the ground, the building density and typology, etc. The localization engine, in particular, uses the COST-Hata propagation model, which is described in [14], which does not require the knowledge of the area morphology and of the building typology with obvious advantages for computational speed.

For each call, the localization engine, through the probes, receives the signal strength level (RXLEV) that is measured by the ME on the serving BTS and on a maximum of six adjacent

---

[3]These measurements form the basic raw data for the handover algorithms in the BSC/mobile switching center (MSC).

[4]The TA is a time interval that represents the round-trip delay between the ME and its serving BTS and can assume a value between 0 and 63, with each step representing an advance of one symbol period (approximately 3.69 μs). With radio waves travelling at about 300 m/μs, each TA step represents a change in a round-trip distance (twice the propagation range) of about 1100 m and, therefore, a change in a distance of about 550 m (TA ring).

[5]The LA is formed by a group of cells and is updated whenever it changes or periodically when a timeout ends. The value of this timeout is network-operator-dependent and can also be 24 h.

[6]The LA can also be obtained by accessing to the visitor location register.

[7]Directive 2002/58/EC of the European Parliament and of the Council of July 12, 2002, concerning the processing of personal data and the protection of privacy in the electronic communication sector (directive on privacy and electronic communications).

BTSs, the cell identifiers (Location Area Code and CI) of the received BTSs, and the actual TA. The DFL algorithm works as follows.

1) Through the received identifiers and using the network database, it obtains the geographic position of the BTSs that are involved in the measurement.
2) Starting from these positions and using the antenna beamwidths that are extracted from the antenna database, it defines an area in which the ME is supposed to be located; this area is represented by the sum of the different areas that are covered by each BTS calculated with simple geometric considerations that take into account just the BTS positions and their antenna radiation pattern and orientation.[8]
3) It further bounds the search area using the intersection with the TA ring whose two circumferences are centered in the BTS position and whose external and internal radii are equal to $R_{ext} = 550 * (TA + 0.5)$ and $R_{int} = 550 * (TA - 0.5)$ for any TA value that is higher than 0.[9]
4) It identifies a grid in this new search area. The step of this grid must be accurately chosen because it affects the localization error and the computation time; in urban areas, for example, a reasonable value is 50 m.
5) For each point p of the grid, it estimated the mean power received by the ME from every BTS involved in the measurement $Pc_i(p)$ using the proper parameters of the network database, the radiation patterns contained in the antenna database, and the propagation model.
6) It estimates the ME position, finding the point that minimizes the mean square error between measured and estimated mean power received by all BTSs. In particular, for each point p of the grid and considering the ith BTS, it calculates the error function $e_i(p) = Pm_i - Pc_i(p)$, where Pm_i is the power measured by the ME on the ith BTS; then, it estimates the ME position p^*, finding the point p that minimizes the mean square error $\|e(p)\|^2 = \sum_i (Pm_i - Pc_i(p))^2$.

C. Tracking Filter

The tracking filter estimates the complete trajectory of the MEs and the related speed, starting from a sequence of punctual localizations received, with the associated timestamps, from the localization engine. It consists of the following blocks.

- *Sampler*—It receives the sequence of ME position estimates, then removes the incorrect localizations (according to an associated numeric code set by the DFL algorithm), and, finally, resamples the remaining ones with a fixed step.
- *Latitude and longitude estimators*—They allow estimating the position (and speed) trajectories along the two directions, attenuating the measurement noises. Regarding the position estimation, two low-pass filters working offline

have been designed. One filter works from the first to the last samples of the location sequence and gives a time-delayed estimation of the position trajectory to take into account delays that are occurring in the data acquisition, processing, and filtering. The other filter works in an analog way but in the opposite direction. A combination algorithm is then used to merge the filtered sequences. Regarding the speed estimation, a filter composed by an ideal differentiator and a second-order low-pass filter has been designed to attenuate the noise and differentiate the location sequence to extract the speed trajectory.

- *Combiner*—It merges the two components to give an estimation of the complete trajectory.

D. Mobility State Estimator

The mobility state estimator separates the set of calls made by "moving ME," i.e., the ME located in cars or on fast means of transport, from the set of calls made by "not moving ME," i.e., the ME used by still users or by pedestrians. The adopted algorithm divides each complete ME trajectory in a sequence of small consecutive trajectories of T_w seconds (T_w is the said evaluation window); it calculates the average ME speed of each small trajectory starting from the instantaneous speeds provided by the tracking filter and compares this speed with a reference threshold v_t: The ME is evaluated as "moving" for each small trajectory if the average speed is greater than the threshold. The values of the two parameters T_w and v_t have been obtained through the following considerations.

- T_w has been obtained minimizing the probability $P(N|M)$ of considering as "not moving" an ME that is "moving."
- Given T_w, v_t is obtained minimizing a linear combination $wP(N|M) + (1 - w)P(M|N)$, $w \in]0, 1[$ of the previous probability and the one of estimating as "moving ME" an ME that is "not moving ME." The minimization of this combination is needed due to the tradeoff between these two probabilities.

Finally, calls lasting less than T_w are discarded, whereas calls lasting more than $2T_w$ are split in different calls, as described before, making it possible to correctly estimate the mobility state, even if it changes during the same call.

Numerically, T_w and v_t have been obtained through an empirical analysis based on 24 car trajectories and on 11 trajectories where the GPS receiver is carried by a pedestrian (five trajectories) or where it is still in an indoor environment (six trajectories). The optimum values for these parameters, for any $w \in [0.3, 0.5]$, i.e., giving more importance to minimize $P(M|N)$, are $T_w = 3.5$ min and $v_t = 7$ km/h.

E. Traffic Map Calculator

The traffic map calculator produces the raster traffic maps for the entire area monitored by the platform. This is accomplished using the set of calls considered by the mobility state estimator as made by "moving ME" in the time interval ΔT that ends when these maps are produced; the value for this interval determines the confidence of the calculated statistics

---

[8]Each point of an area is considered as covered by a given BTS if that BTS is the nearest one and the point is in the antenna radiation pattern.
[9]If the TA value is equal to 0, then $R_{int} = 0$.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

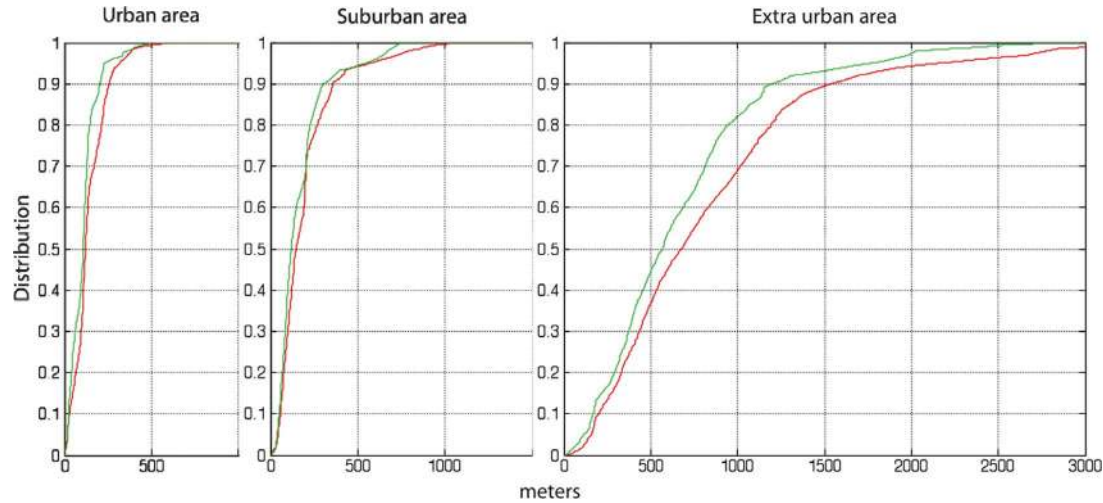CALABRESE *et al.*: REAL-TIME URBAN MONITORING USING CELL PHONES

5

Fig. 2.   Cumulative error distributions of both the (red) DFL positioning algorithm alone and the (green) tracking filter applied to the output of the localization engine in the three different environments.

and, consequently, has to be accurately chosen. Practically, it depends on the number of calls it includes and, finally, on all the parameters that are linked to the telephone traffic density, such as the area type, the time of the day, etc. As previously said, these traffic maps are periodically updated to have real-time estimations. As an example, we describe the algorithm used by this module to produce the traffic maps for the average speed.

- It splits the trajectory of the calls made in the time interval ΔT among the different pixels of the raster map where the trajectory is located.
- It calculates the average speed of an ME on each pixel as the average of the different ME speed estimations in that pixel $v_{ij} = (1/card\, I_{ij}) \sum_{m \in I_{ij}} v_{ijm}$, where i identifies the ith pixel, j identifies the jth ME, and I_{ij} represents the set of different ME_j speed estimations in the ith pixel.
- It calculates the total average speed value for each pixel as the average of the speeds of the MEs that are located in that pixel $v_i = (1/card\, I_i) \sum_{j \in I_i} v_{ij}$, where the set I_i represents the indexes of the ME_j located in the ith pixel.

Similarly, the traffic map calculator obtains the maps of the average speed in each of the four quadrants (northwest, northeast, southeast, and southwest) of a Cartesian reference system centered in the center of the pixel and the analogous maps of the maximum speeds.

The average speed in each quadrant can then be used to calculate the speed and the traveling times on any street segment simply by inverting the process described before and considering the particular orientation of the street segment to select the specific quadrant to be used in the calculation.

F. LocHNESs Performance Evaluation

The different blocks of the LocHNESs platform have been tested with an *ad hoc* measurement campaign to obtain their performance.

The performance of the localization algorithm has been measured in a number of areas being representative of different

TABLE I
STATISTICAL ERRORS OF THE POSITIONING ALGORITHM

	Urban area	Suburban area	Extra urban area
50th percentile	119 m	204 m	659 m
67th percentile	163 m	295 m	1235 m
95th percentile	364 m	665 m	2754 m
Average value	159 m	295 m	1457 m

morphological and land-use Italian environments[10]: They were four large Italian cities, which are representative of both urban and suburban areas with a flat (Milan and Turin) and undulating ground (Naples and Rome), three mountain (Cortina) and sea (Taormina and Cinque Terre) villages, and the Bologna highway, which are all representative of both suburban and extra urban areas. A total of about 2000 measurement points have been selected to cover the eight areas, and four localization tests per point have been carried out to take into account different mobile user positions (two measurements with the user in a car and two measurements with the user on the sidewalk). All the tests have been carried out using an *ad hoc* developed software application installed on the mobile phone that was able to collect all the measurements that are needed and to send them to a remote computer for their elaboration via two SMS messages. The statistics of the error (50th, 67th, and 95th percentile) in the three environments are shown in Fig. 2 and Table I; their average value ranges from 159 m in the urban area to 295 m in the suburban one and up to 1457 m in the extra urban area. These values, particularly those in the urban and suburban areas, are very satisfactory and suggest that this location algorithm can be used to support different location-based services, such as emergency, yellow pages, etc.

The performance of the tracking filter that is applied to the output of the localization engine has been tested on a set of 24 real trajectories consisting of a total of about 55 000 instantaneous positions-speed couples and measured equipping

---

[10]For the environment classification of these areas, we have used the results for Italy of the European Union's Coordination de l'Information sur l' Environnement project (http://stweb.sister.it/itaCorine/corine/progettocorine.htm).

TABLE II
PERFORMANCE IMPROVEMENTS (IN PERCENT) OBTAINED BY APPLYING
THE TRACKING FILTER TO THE OUTPUT OF THE LOCALIZATION
ENGINE WITH RESPECT TO THE PERFORMANCE OF
THE LOCALIZATION ENGINE ALONE

	Urban area	Suburban area	Extra urban area
50th percentile	8.3 %	18.9 %	16.2 %
67th percentile	17.1 %	2.9 %	19.0 %
95th percentile	27.80%	3.90%	18.10%

TABLE III
ERROR PROBABILITIES (IN PERCENT) OF
THE MOBILITY STATE ESTIMATOR

	P("ME moving"|"ME not moving")	Prob("ME not moving"|"ME moving")
GPS on the car	3.2 %	18.7 %
GPS on the pedestrian or indoor	16.0 %	Not applicable
GPS on the bus	Not applicable	12.0%

a test car with a GPS receiver and an odometer moving in urban, suburban, and extra urban environments in the Turin area.[11] Fig. 2 also contains the statistics of the localization error of this filter (green lines), whereas the performance improvement with respect to the performance of the localization engine alone is shown in Table II for completeness.

These 24 trajectories with other 11 different trajectories, where the GPS receiver is carried by a pedestrian (four trajectories), it is carried by a person on a bus (one trajectory[12]), or it is still in an indoor environment (six trajectories), have also been used to evaluate the mobility state estimator error probabilities P("ME moving"|"ME not moving") and P("ME not moving"|"ME moving"), i.e., the probability of considering "moving" an ME that is "not moving" and the probability of considering "not moving" an ME that is "moving," respectively[13]; these probabilities have been calculated by extracting from these trajectories a total of 1011 nonoverlapped evaluation windows. The four error probabilities, as shown in Table III, are all quite low, and particularly remarkable is that of considering as moving a car that is, instead, still (3.2%).

Finally, the performance of the complete system has been tested in the city of Rome comparing the travel times measured with a GPS-equipped car with those provided at the same time by the LocHNESs platform. These tests have been carried out on the Rome's ring road and urban streets, with the latter divided, according to their dimension and importance, into primary and secondary urban streets. In particular, 25 journeys have been identified, with an average length of 5.1 km, and a total of 2421 km has been traveled on them; the mean absolute percentage errors (MAPEs) are shown in Table IV: they are quite low, and consequently, if combined with a frequent update period (i.e., 5 min, as already said before), they suggest positively using the platform for route planning, travel-time estimation, trafficking monitoring, car accident detection, etc.

[11]In six of these 24 trajectories, the car was parked on the side of the street.

[12]For this trajectory, the error probability is P("ME not moving"|"ME moving").

[13]When the ME is on a car, the error probability P("ME moving"|"ME not moving") is related to the six trajectories with the car parked on the side of the street; consequently, P("ME not moving"|"ME moving") is obtained by the remaining 18 trajectories with the car moving.

TABLE IV
MAPEs (IN PERCENT) ON THE TRAVELING TIMES
OF THE LOCHNESs PLATFORM

Type of Street	MAPE
Bypass	14.88 %
Primary urban streets	10.08 %
Secondary urban streets	17.66 %

G. LocHNESs for the Network Operation and Management

The analysis of the protocol messages on A-bis interfaces allows probes to detect network events other than measurement results, such as voice channel request, connection failure, handover detection and failure, and location update. By pairing the measurement result messages and each event message, the ME that originated the event can be localized. This event localization enables two vital applications for a network operator, namely, quality-of-service (QoS) monitoring and real-time network planning. The former application immediately derives from the localization of all events and their subsequent aggregation in "event maps," similar to those previously described for the road traffic, which, in each pixel, contain the number, or equivalently the density, of a given event such as successful handovers, failed handovers, connection failures, etc., allowing the network operator to immediately know the QoS level users are experiencing depending on their positions. Moreover, the RXLEV values contained in the measurement result messages, after their localization, allow the network operator to derive a map of the real coverage of the network, whereas the number of measurement result messages, which, as already mentioned, are transmitted with a period of 480 ms by an ME in the state "connected," gives the offered traffic by all the users in each pixel.

All these maps suggest the areas where the network is not properly working and where it has to be improved, for example, by modifying the BTS parameters (frequencies, transmit power, etc.). In doing so, corrections can be carried out as soon as the problem is evidenced, and their effect can be immediately evaluated with new event maps.

III. REAL TIME ROME

The Real Time Rome project, which was developed for the Tenth International Architecture Exhibition of the Venice Biennale, aimed to create a unique approach to real-time urban monitoring, based on the use of anonymous real-time data gathered from cellular phone and locational data from the public transportation system in the city of Rome (see http://senseable.mit.edu/realtimerome). The data processed by the developed application were coming from different sources, as described in the following.

A. LocHNESs Data

Telecom Italia installed the platform LocHNESs and the related probes on a group of BSCs located in Rome, covering an area of approximately 100 km^2 in the northeast of the city. The area was divided into a grid of 250×250 m squares, and the traffic maps were produced every 5 min, as described in paragraph II E, using for each map the calls that were made by mobile users in the last 30 min (ΔT). Moreover, some *ad hoc*

algorithms were added to obtain maps that are related to the spatial distribution of pedestrians and foreigners. The former was calculated summing the number of MEs that are estimated to be in each pixel of the grid and considered "not moving MEs" by the mobility state estimator,[14] and some performance results are presented in Table III; the latter was calculated considering the trajectories of those MEs whose International Mobile Subscriber Identity (IMSI) numbers were related to foreigner mobile network operators.[15] To do this, the probes just extracted the first three digits of the IMSI that are the mobile country code (MCC) and that identify the nationality of the network operator. The other 12 digits, which identify the network operator and the single user, are, instead, not captured by the probes to preserve the privacy.

As already explained in Section II, the acquired data were related to Telecom Italia subscribers involved in calls, capturing only a sample of the actual population. To understand the size of the sample, we provide the following information.

- In Italy, there are 86 million mobile lines and 45 million mobile customers, which represents 76% of the total population.
- Telecom Italia has a market share of 40.3% with about 34.3 million mobile lines.[16]
- An average of 30 000 calls are made, in the area considered for the project, by users that are estimated to be moving in vehicles (moving MEs) during a weekday—27 000 calls made during the interval 07:00–21:00.

The captured sample of the population can then be considered of statistical relevance to extrapolate useful information about the spatial relative distribution of the whole population in terms of the percentage of people that are present in a certain area compared with the whole monitored area.

The usability of these data in urban planning and tourist management has recently been described and tested in [15] and [16]. In [15], we explored how the analysis of wireless data networks could be used to better understand the "bricks and mortar" of the physical space. In much the same way as a remote-sensing specialist might examine the spectral signature of a distant object and make assumptions about its underlying components and structure, we tried to understand the temporal signatures of mobile phone usage and infer the underlying organization of space itself. Thus, we are able to shed light on how data that are generated as a byproduct of network activity by large populations can drive our understanding of the built environment and its usage. In [16], instead, we developed novel methods and tools to explore the significant mobile phone spatiotemporal data to help uncover the presence and movements of tourists from cell-phone network data.

B. Voice and Data Traffic

A further Telecom Italia server provided the voice and data traffic (expressed in Erlang) served by each of the BTSs located in the urban area of Rome (about 450 directional antennas covering about 47 km^2). The Erlang is the measure of telecommunications traffic density. It is a dimensionless "unit" representing a traffic density of one call-second per second (or one call-hour per hour, etc.). The Erlang honors A. K. Erlang (1878–1929), who was a Danish mathematician who studied the mathematics of telephone networks. These data were localized and collected with a sampling period of 15 min.

C. Location of Atac Buses

The buses run by Atac (the transport company of the City of Rome) carried GPS devices that calculated each vehicle's location based on GPS satellite triangulation and reported these data to an Atac server in real time using User Datagram Protocol (UDP) datagrams sent through a General Packet Radio Service connection. We acquired locations for 7268 vehicles, with a sampling period of 30 s (during operation times).

D. Location of Samarcanda Taxis

The taxis run by Samarcanda (a taxis cooperative working in Rome) carried GPS devices and reported each vehicle's location to a Samarcanda server in real time using a radio transmission. Data for 43 taxis were collected and stored with a sampling period of 5 min.

E. Localized Traffic Noise

A wireless sensor network was developed and implemented to acquire real-time traffic noise from different spots in Rome. An audio-streaming technology, based on MP3 audio content encoding and hypertext transfer protocol/transmission control protocol, was used to send such data to Venice, where it was punctually reproduced using AudioSpotlights [17], which provides a way to play different sounds in different listening areas inside the same exhibition area.

IV. SYSTEM ARCHITECTURE

To meet the real-time requirement of the application, an *ad hoc* software platform was developed to split the data processing between various stages based on the amount of data managed by the nodes and their processing power. Three servers were set up by Telecom Italia, Atac, and Samarcanda to provide locational data, both using secure file transfer protocol (SFTP) transfer and UDP datagram transfer. A database was designed and ran on MySQL in the Senseable City Laboratory server, both with some Java applications, which collected the data from the external servers and preprocessed them, providing the results to an internal SFTP server. The database was composed of several tables whose structures are shown in Fig. 3. Finally, six computers at the Venice Biennale exhibit continuously accessed the Senseable City Laboratory file transfer protocol server and ran Java software (developed using Processing and

---

[14]We considered as belonging to pedestrian people those MEs that are evaluated as "not moving" in all the evaluation windows that form their trajectories; this is because when a car is blocked in the traffic, it is not always still, but it sometimes moves, and consequently, the average speed in at least one of its evaluation windows can make it to be considered "moving." This category also contains information about people in indoor environments.

[15]The first three digits of the IMSI are the MCC and identify the nationality of the mobile network operator.

[16]Report Telecom Italia, first semester 2007 (http://www.telecomitalia.it/TIPortale/docs/investor/Monografia-1sem2007.pdf).
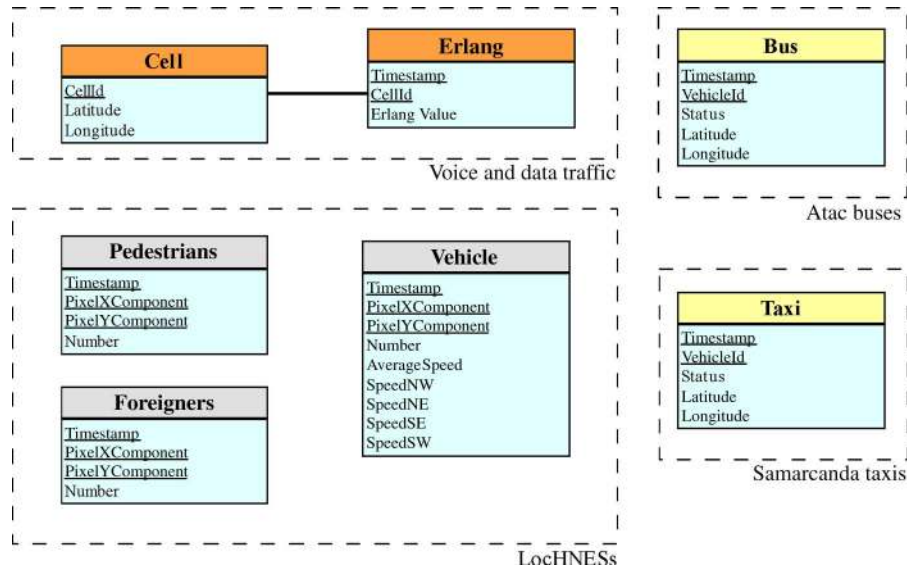
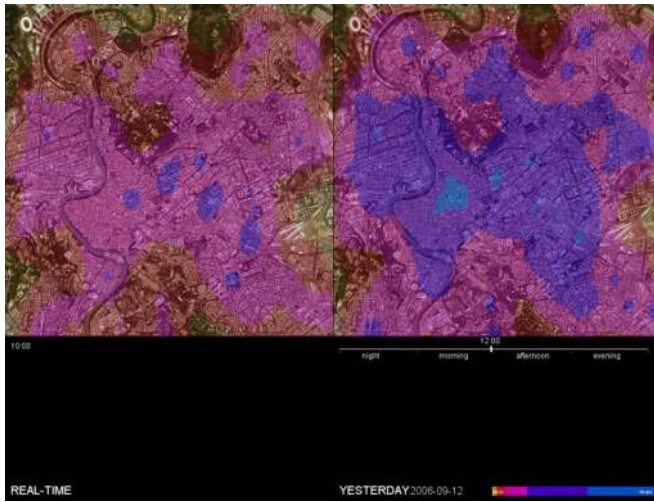Fig. 3.   Schematic of the database tables.

Fig. 4.   Pulse: What are the patterns of use in Rome?
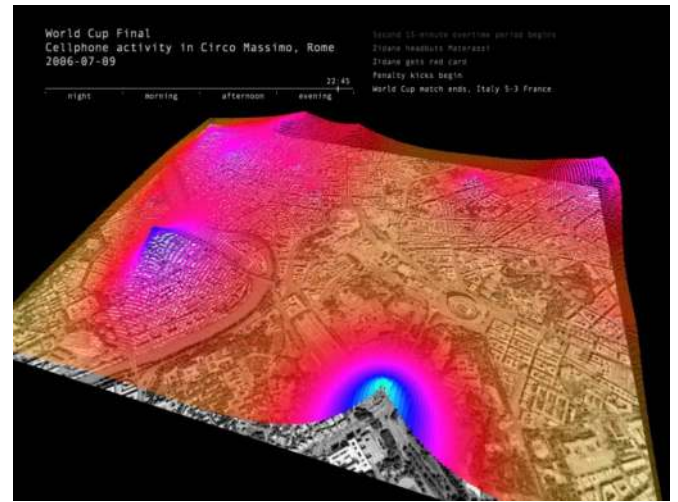

Fig. 5.   Gatherings: What does Rome look like during special events?

OpenGL) to visualize the different dynamic maps of the city in real time, using Google maps as background. Furthermore, three computers that are connected to three audio-streaming sources (coming from the three audio sensors installed in Rome) played locational traffic noise in time synchronization with the visual software.

V. PROCESSING AND VISUAL SOFTWARE

In the following, the detailed description of the six processing and visual software is given, starting from the questions that they address.

Pulse: Where in Rome Are People Converging Over the Course of a Day?: This software visualized the intensity of mobile phone calls in Rome at the present moment and compared it to the previous day's data (see Fig. 4). It is based on a geographic interpolation of the served traffic intensity provided by each monitored antenna. To this end, the Rome urban area

was divided in 40×40 m pixels, and the traffic intensity was assigned at each pixel considering the distance between the pixel and the surrounding antennas, the cell type, the location of the pixel in the city, and an exponential distribution function.

Gatherings: How Do People Occupy and Move Through Certain Areas of the City During Special Events?: This software showed the prerecorded intensity of mobile phone usage during important events: viewing the World Cup final match between Italy and France on July 9, 2006 and Madonna's concert in Rome on August 6, 2006 (see Fig. 5). The software was realized by concatenating 3-D visualizations of the served traffic in specific areas of Rome at different times of the day.

Icons: Which Landmarks in Rome Attract More People?: This software showed the density of people using mobile phones at different historic attractions in Rome (see Fig. 6). To this end, the served traffic intensity at each site was calculated as the average of the traffic intensities of the pixels located in

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

CALABRESE *et al.*: REAL-TIME URBAN MONITORING USING CELL PHONES

9

Fig. 6. Icons: Which landmarks in Rome attract more people?

Fig. 8. Connectivity: Is public transportation where the people are?

Fig. 7. Visitors: Where are tourists congregating?

Fig. 9. Flow: Where is traffic moving?

correspondence with the site. A bar on the top of each attraction showed the relative traffic intensity, whereas at the bottom of the screen, a weeklong data comparison between the most popular site and the least popular site was shown.

Visitors: Where Are the Concentrations of Foreigners in Rome?: This 3-D software highlighted a 24-h loop of the locations around the Stazione Termini neighborhood of Rome, where foreigners were speaking on mobile phones (see Fig. 7). An algorithm created a 48-length data queue, where, with a sampling rate of 1/30 min, the newest foreigner locational data were added (deleting the oldest one). For 3-D visualization purposes, an algorithm was used to spatially interpolate the foreigners' locational data to create a 10×10 m pixel matrix (based on a 2-D smoothing technique).

Connectivity: Is Public Transportation Where the People Are? How Do the Movement Patterns of Buses, Taxis, and Pedestrians Overlap in the Stazione Termini Neighborhood of Rome?: This software showed the changing positions of buses and taxis indicated by yellow points and the relative densities of mobile phone users, which are represented by the red areas (see Fig. 8). An algorithm was used to acquire and update the

bus and taxi location in real time (based on a hash table). It also estimated bus and taxi paths based on the previous three locations, drawing a yellow tail on the map. The algorithm acquired the pedestrian locational data every 5 min, showing a red layer on the top of the map (areas colored by a deeper red had a higher density of pedestrians). Public transportation authorities can use this type of a dynamic map to understand how the bus connectivity changes in space and time and how that is correlated to a real-time estimation of the demand.

Flow: Where Is Traffic Moving?: This software visualized the locational data of mobile phone callers traveling in vehicles (see Fig. 9). It focused on the area around the Stazione Termini and the Grande Raccordo Anulare (Rome's ring road). The software created a layer on the top of the map, showing 250×250 m pixels whose colors were related to vehicle speeds. Red indicated areas where traffic was moving slowly; green showed areas where vehicles were moving quickly. If the average speed that is associated with the pixel was higher than 40 km/h, the software also showed an arrow in the center of the pixel whose direction was the dominant direction of travel and magnitude was proportional to the related speed.

VI. CONCLUSION

In this paper, we have presented a new real-time urban monitoring platform and its application to the City of Rome during the Tenth International Architecture Exhibition in Venice. The project used the Telecom Italia LocHNESs platform for the real-time evaluation of statistical indexes based on the anonymous monitoring of the ME movements; in particular, this platform was used to obtain a set of maps describing the vehicular traffic status and the movements of pedestrians and foreigners. Moreover, the system developed for the project acquired the instantaneous location of buses and taxis and the voice and data traffic served by all the BTSs of the network. All these data were opportunely combined and updated in real time to realize several information layers on top of a map of Rome.

The visualizations proposed in this paper are exploratory and give a qualitative understanding of how the use of mobile phone and vehicle real-time location data can be used to provide valuable services to citizens and authorities. Different analysis is being done on the acquired data (see, for instance, [15] and [16]) to quantitatively evaluate the usability of the data for urban-planning purposes and tourist management, with the principal aim being to show how modern digital technologies and their applications could give city dwellers deeper knowledge of urban dynamics and more control over their environment by allowing them to make decisions that are more informed about their surroundings, reducing the inefficiencies of present-day urban systems.

REFERENCES

[1] J. L. Ygnace, J. G. Remy, J. L. Bosseboeuf, and V. Da Fonseca, *Travel Time Estimates on Rhone Corridor Network Using Cellular Phones as Probes: Phase 1 Technology Assessment and Preliminary Results.* Arceuil, France: French Dept. Transp., 2000.
[2] J. White, J. Quick, and P. Philippou, "The use of mobile phone location data for traffic information," in *Proc. 12th IEE RTIC*, 2004, pp. 321–325.
[3] Y. B. Y. Yim and R. Cayford, "Investigation of vehicles as probes using global positioning system and cellular phone tracking: Field operational test," California PATH Program, Univ. Calif., Berkeley, CA, Rep. UCB-ITS-PWP-2001-9, 2001.
[4] R. Herring, A. Hofleitner, S. Amin, T. Nasr, A. Khalek, P. Abbeel, and A. Bayen, "Using mobile phones to forecast arterial traffic through statistical learning," in *Proc. 89th Transp. Res. Board Annu. Meeting*, Washington, DC, Jan. 10–14, 2010.
[5] B. L. Smith, M. L. Pack, D. J. Lovell, and M. W. Sermons, "Transportation management applications of anonymous mobile call sampling," in *Proc. 80th Annu. Meeting*, Washington, DC, 2001.
[6] R. Bolla and F. Davoli, "Road traffic estimation from location tracking data in the mobile cellular network," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2000, vol. 3, pp. 1107–1112.
[7] D. J. Lovell, "Accuracy of speed measurements from cellular phone vehicle location systems," *J. Intell. Transp. Syst.*, vol. 6, no. 4, pp. 303–325, Oct. 2001.
[8] B. Smith, M. Pack, D. Lovell, and M. Sermons, "Transportation management applications of anonymous mobile call sampling," in *Proc. 11th Annu. Meeting ITS Amer.*, Miami, FL, 2001.
[9] R. Cayford and T. Johnson, "Operational parameters affecting use of anonymous cell phone tracking for generating traffic information," in *Proc. Inst. Transp. Studies 82nd TRB Annu. Meeting*, 2003.
[10] K. Sohn and K. Hwang, "Space-based passing time estimation on a freeway using cell phones as traffic probes," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 559–568, Sep. 2008.
[11] K. Sohn and D. Kim, "Dynamic origin—Destination flow estimation using cellular communication system," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 2703–2713, Sep. 2008.
[12] *3rd Generation Partnership Project; Technical Specification Group GSM EDGE Radio Access Network; Base Station Controller—Base Transceiver Station (BSC-BTS) Interface; Layer 3 Specification (Rel. 1999)*, Third Generation Partnership Program, TS 08.58, 1999.
[13] A. Kupper, *Location-Based Services—Fundamentals and Operation.* New York: Wiley, 2005.
[14] E. Damosso, *Digital Mobile Radio Towards Future Generation Systems*, 1998, COST 231 Final Rep.
[15] J. Reades, F. Calabrese, A. Sevtsuk, and C. Ratti, "Cellular census: Explorations in urban data collection," *Pervasive Comput.*, vol. 6, no. 3, pp. 30–38, Jul.–Sep. 2007.
[16] F. Girardin, F. Calabrese, F. Dal Fiore, C. Ratti, and J. Blat, "Digital footprinting: Uncovering tourists with user-generated content," *Pervasive Comput.*, vol. 7, no. 4, pp. 36–43, Oct.–Dec. 2008.
[17] Holosonics Research Labs, *AudioSpotlight*. [Online]. Available: http://www.holosonics.com/

Francesco Calabrese (M'04) received the M.S. and Ph.D. degrees in computer and systems engineering from the University of Naples Federico II, Naples, Italy, in 2004 and 2007.

From 2007 to 2009, he was a Postdoctoral Associate with the Massachusetts Institute of Technology, Cambridge. He is currently a Research Scientist with the Senseable City Laboratory, Massachusetts Institute of Technology, where he coordinates the Network and Society Group. His research interests include ubiquitous computing, the analysis of urban dynamics through sensor networks, and the analysis and design of distributed and hybrid control systems.

Massimo Colonna received the M.S. degree in electronic engineering from Politecnico di Torino, Turin, Italy, in 1997.

He is currently a System Engineer with the Service Layer and Messaging Group of TILab, Innovation and Engineering Division, Telecom Italia, Turin. His interests are related to localization techniques and service platforms in wireless systems (Wi-Fi and ultra-wideband in particular) and to the use of georeferenced mobile phone data in applications of urban analysis and transportation system planning.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

CALABRESE *et al.*: REAL-TIME URBAN MONITORING USING CELL PHONES   11

Piero Lovisolo received the M.S. degree in electronic engineering from the Politecnico di Torino, Turin, Italy, in 1989.

In 1989, he joined CSELT (now TILab), the R&D Center of Telecom Italia, Turin, in the Satellite Systems and Services Division, where he was involved in the Italsat, Olympus, Orion, and Iridium projects. Since 2002, he has been acting as the Project Leader of TILab activities for Location-Based Services and Technologies Evolution. He led research activities on location technologies based on mobile networks [Global System for Mobile Communications (GSM) and Universal Mobile Telecommunication System (UMTS)] and Global Navigation Satellite System [Global Positioning System (GPS), Assisted GPS, European Geostationary Navigation Overlay Service, Galileo]; his group developed the Telecom Italia Mobile Hybrid/Assisted GPS platform and participated on many edge research project on location based services and Galileo.

Dario Parata received the M.S. degree in electronic engineering from the Politecnico di Torino, Turin, Italy, in 1996.

Since 2001, he has been with TILab, which is the innovation and engineering division of Telecom Italia, Turin, where he was involved in projects related to positioning in mobile radio networks with particular emphasis on Third-Generation Partnership Project standardization and localization technology selection. As part of his research activity, he studied the application of statistical tracking to network quality-of-service monitoring and planning. Since 2005, his main interest is in the use of mobile phone data for road traffic estimation and urban dynamics analysis.

Carlo Ratti received the Ph.D. degree from the University of Cambridge, Cambridge, U.K.

He is currently the Director of the Senseable City Laboratory, Massachusetts Institute of Technology, Cambridge, and an Adjunct Professor with Queensland University of Technology, Brisbane, Qld., Australia. He is also a founding partner (with W. Nicolino) and the Director of the architectural firm Carloratti.

Dr. Ratti is a member of the Ordine degli Ingegneri di Torino and the Association des Anciens Elèves de l'École Nationale des Ponts et Chaussées.