Computers and Electrical Engineering 50 (2016) 79–90

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A battery-friendly data acquisition model for vehicular speed estimation☆

Sevgi Kaya*, Necati Kilic, Taskin Kocak, Cagri Gungor

[a] Department of Computer Science, ETH Zurich, Switzerland
[b] Department of Computer Engineering, Bahcesehir University, Turkey
[c] Department of Computer Engineering, Abdullah Gul University, Turkey

ARTICLE INFO

Article history:
Received 12 May 2014
Revised 16 January 2016
Accepted 18 January 2016
Available online 12 February 2016

Keywords:
Vehicular speed estimation
Cellular positioning
GPS positioning
Microscopic traffic simulation
Mobile tracking
Traffic theory

ABSTRACT

Modeling traffic flow and gathering accurate traffic congestion information are two challenging problems in smart transportation systems. Most of the traffic flow models and velocity estimation methodologies that have been proposed so far gather the data from GPS-equipped smart phones and extract the flow model based on GPS sampling. However, these approaches tend to fail in real life scenarios due to the insufficient vehicle data and unpredictable dynamics of the flow. Furthermore, utilization of GPS sensor leads to a battery drainage and hence reduces the overall system performance. In this paper, we propose a new battery-friendly data acquisition model to obtain the raw data. We then evaluate our model under various traffic conditions to determine its feasibility in vehicle speed estimation. The proposed model results in 88% location accuracy whereas it reduces the battery consumption by half.

1. Introduction

For many years, transportation systems relied on infrastructure based equipment such as loop detectors and cameras for data gathering. Although the accuracy rate of such systems is high, they can not be deployed to the every segment of the roads due to their high costs. Hence the trends in the smart transportation industry changed from infrastructure based systems to on-board and mobile data gathering devices in order to adapt to the dynamic environment of traffic and also the spatio-temporal features of roads while reducing the cost.

New ways of traffic data acquisition have been investigated and classified as on-board and mobile sources where former one refers to an embedded hardware deployed on vehicles and the latter one is based on the utilization of smart phones for data gathering. Under the caption of on-board data sources, Radio Frequency Identification (RFID) transponders, License Plate Recognition(LPR) systems and GPS devices have been investigated and compared according to their benefits and drawbacks. Although, usage of on-board devices for traffic data gathering leads highly accurate results in the concept of traffic status estimation, their high deployment costs and low penetration rates on roads put away the on-board devices from ideal traffic data gathering technology. Therefore, new methods that have low cost and extensive usage features, but yields accurate results as RFID [1,2], LPR [3,4] and GPS methods [5–7], have been investigated. The technologies based on mobile devices

☆ Reviews processed and approved for publication by the Editor-in-Chief.

* Corresponding author. Tel.: +90 5546236456.
E-mail address: sevgikaya2010@gmail.com (S. Kaya).

http://dx.doi.org/10.1016/j.compeleceng.2016.01.017

have been proposed and classified as cellular positioning, GPS positioning and hybrid systems where the former one stands for gathering of location data from triangulation of cellular network signals, the middle one refers to usage of GPS sensors deployed on smart phones as a data source and finally the latter one refers to the combination of these technologies to provide more accurate positioning data while avoiding the battery drainage arose from GPS based positioning. As a result of the experiments that have been conducted to compare the ways have been proposed so far for data gathering, mobile devices got an edge over others with their high accuracy and low cost features.

Before the rapid emergence and penetration of GPS technology within the mobile devices, cellular phone based monitoring mostly relied on identification of the cellular devices via the base station signal information. In this approach, triangulation of cell tower signals, trilateration, handovers between the base stations or a hybrid of these techniques have been utilized to estimate the location of vehicle probes. Most of the mobile data gathering work uses base station signals [8] to find out the approximate location of probes. Although these methods satisfy the low cost and high penetration requirements of an ideal data gathering technology, accurate vehicle location estimation remains as a challenge. Consecutively, when the precision of the location information is low, considerable difficulties arise in the process of computing the vehicle speed.

Together with the fact that mobile communications have been developed immensely in the last decade and become readily available all over the world, low cost, increasing penetration and high accuracy of GPS sensors attracted the attention of the transportation authorities for using GPS-equipped smart phones on traffic data gathering. [9–11] This approach yields more reliable position information, and hence provides more accurate velocity, travel time, acceleration/deceleration, direction estimation. As it is emphasized in [9], accurate traffic status estimation highly depends on the reliability of vehicle's spatio-temporal data. Although, GPS-sensor equipped smart phones have been proposed to improve state-of-the-art cellular phone based monitoring which rely on tracking base station signals for location estimation with their high accuracy, spatio-temporal extensive coverage and low cost features, there are some other concerns related to privacy of the users, energy consumption of the phone and overload on the communication network. Since spatio-temporal data need to be transmitted to a central server between specific time intervals, in most cases, it leads to network overload issues. Besides the overload of communication network, frequent data updates cause to battery drainage of smart phones that involved in data acquisition process. Furthermore, obtaining almost exact position and speed data of vehicle probes may lead to infringing privacy [12]. Consequently, albeit GPS-enabled smart phones can be considered as a strong candidate for data gathering with its low cost and high accuracy features, many concerns of the technology in terms of privacy, maintenance and battery consumption still remain challenging.

For many years, lots of studies have been conducted to address these challenges by proposing hybrid technologies to gather traffic data from both GPS sensors and cellular network signals [12–19]. While the idea of using both GPS sensors and cellular network signals to overcome the issues evolved from these technologies is clearly a good one, most of the state-of-the-art work has utilized GPS sensors as a supplementary source to cellular network signals and disregarded the scheduling of data sources in order to avoid battery drainage of smart phones. In this paper, we propose a data acquisition model to address the battery drainage issue of smart phones by electing data sources with respect to the battery information.

In our approach, GPS sensors are utilized for gathering location data when battery is sufficient, otherwise cellular positioning techniques have been applied to estimate location. This paper also introduces an experiment that have been conducted to estimate the battery status of each vehicle probe by measuring battery rates on different modes where phone is in idle mode, involved in a GPS based and cellular network based location look-ups respectively. Gradient descent algorithm has been applied to fit the data sets to a function to find out the battery rates on each mode. Results of the experiments have been illustrated in Section 2. The techniques that have been used for cellular network and GPS sensor based positioning have been introduced in detail. Furthermore, this paper proposes an approach to come up with a solution for the issue that arises from the determination of update frequencies of each location look up. In GPS sensor based positioning, it is determined with respect to battery status, conversely, update frequencies of cellular positioning have been determined according to handovers between two cells where phone signals are transferred from one cell to another. Furthermore, optimization and calibration methodologies for data gathering process have been studied to achieve effective results.

Our research primarily focuses on investigation of data gathering methodologies to develop an accurate and effective dynamic route guidance system which plays an essential role in the prevention of traffic congestion. To illustrate the feasibility of our model, we performed experiments to estimate vehicle velocity. We conducted real-life experiments with an Android application deployed in a smart phone on a particular road segment. Furthermore, the results of proposed data gathering method are illustrated in our simulator by considering two cases to present the performance of our methods over real life limitations where former one requires a fixed number of vehicles and continuous GPS data transmission occurs while the latter one uses GPS sampling to extract the velocity data dynamically throughout the simulation. Velocity measurements of our model are compared with the ground truth data. We considered velocity values extracted from GPS sensors as ground truth to illustrate for analysis purposes. Given the output of a hybrid data acquisition model prototype, we propose a system that obtains highly accurate location data via several gathering methodologies while reducing the energy consumption. The main contributions of this paper are as follows:

- Evaluate the effectiveness of using cellular positioning techniques along with GPS-sensor based data collection via performing experiments on energy consumption and location accuracy for various data acquisition modes where only GPS positioning, only cellular positioning and both of the techniques are executed.

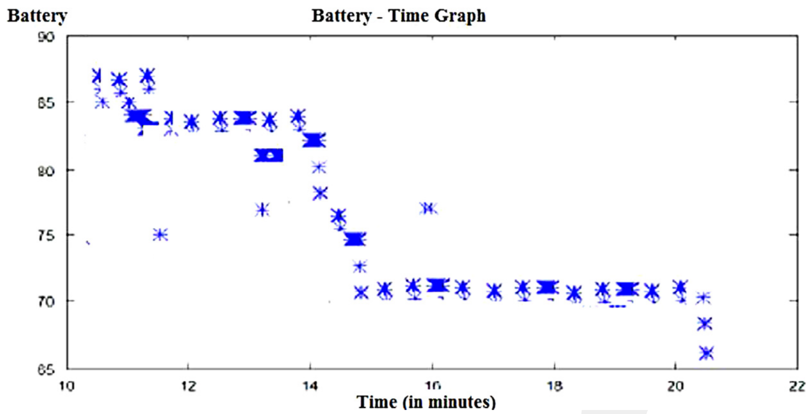

Fig. 1. Battery/Time data set in idle mode.

- Our proposed system is designed to be used for average velocity calculation which is an essential parameter in traffic monitoring systems. Therefore, we perform experiments to compute the average velocity by extracting location data with the hybrid data acquisition model. We assume average speed measurements obtained from GPS positioning as the ground truth data to assess the performance of our model.
- Assess the feasibility of cellular positioning in cases where the GPS sensors are not available for location estimation. In this paper, we propose a hybrid acquisition model which applies cellular positioning techniques to obtain the raw location data where GPS sensor is not available or battery of a particular smart phone is too low for location look up.

The rest of the paper is organized as follows. Section 2 describes the data gathering model by presenting experiments on battery consumption and explains system components and main contributions of our model. Section 3 illustrates the evaluation of extracted data by conducting real-life and simulation experiments. Finally, Section 4 concludes the paper by evaluating the outcomes of data gathering, extraction methodologies and simulation results.

2. Battery friendly spatio-temporal data gathering

With the recent advancements in sensor-equipped smart phones, many applications that provide location based mobile services have been developed. These kind of applications generally process location information obtained from satellites via their GPS sensors. However, fundamental problem of location based mobile services is providing inaccurate position samples. Although, this problem can be solved by using GPS sensors that give accurate results for all position monitoring tasks, GPS sensors have some limitations [20]. GPS sensors can not provide accurate results in places such as bridges and tunnels. Moreover, GPS chip sets lead to battery drainage by consuming excessive amounts of power. Besides the limitations, when smart phones without GPS sensors are taking into account, application developers rely on other positioning techniques such as cellular GSM base station fingerprinting [21,22]. Since cell phones always communicate with base stations, using base station fingerprints to obtain location information has no additional cost over battery. In this paper, we propose a new data acquisition model based on GPS sensors and cellular base stations. A noteworthy fact about our approach is that, appropriate data source for positioning is determined by battery status. When battery is sufficient, it utilizes GPS sensor data, otherwise it gathers cellular base station fingerprints for positioning.

2.1. Energy measurements

With dynamic election of positioning source, our method consumes less energy than other acquisition methods which utilize only GPS sensors for positioning. This situation can be explained by the effective radiated power(ERP) difference of GPS satellites and GSM signals. GSM satellites consume $2x10^{-11}$ mW/m^2 ERP in average, while GSM signals consumes at most 10 mW/m^{2}ERP [23,24]. The difference (117dB) causes GPS signals to be more sensitive than GSM signals under places like tunnels and some weather conditions [25,26].

In this experiment, our goal is to measure battery consumption rates on different modes of smart phones. We logged battery status and time information when phone is in idle mode, only GPS sensor is used to get location information and only phone involved in an active call. In idle mode, we charged phone's battery to 100%, and turn off all of the services that run in background except the idle mode logger. We also recorded location information for every minute to be sure that location methods are working correctly. Battery consumption measurements on idle mode is illustrated in Fig. 1.

Consecutively, we tracked battery status, location and time-stamp data of the phone on modes where only one of the GPS sensor and GSM signal positioning techniques has been applied for spatio-temporal data gathering. We, then applied Gradient Descent Linear Regression algorithm to our data set to find out the cost function that fits to data. Finally, we

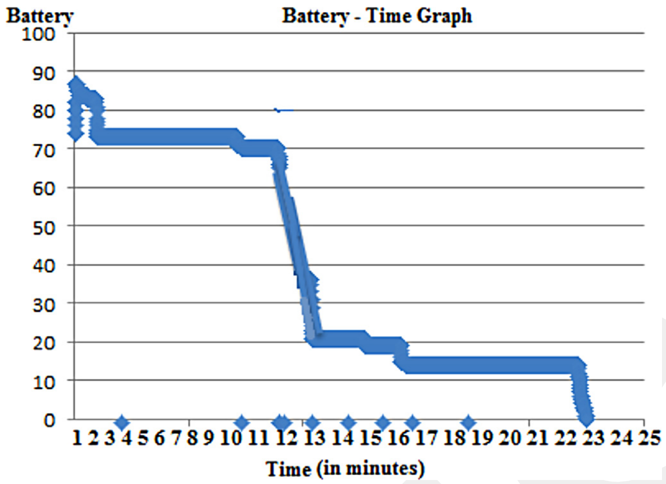

Fig. 2. Cost function that measures battery consumption rate in idle mode.

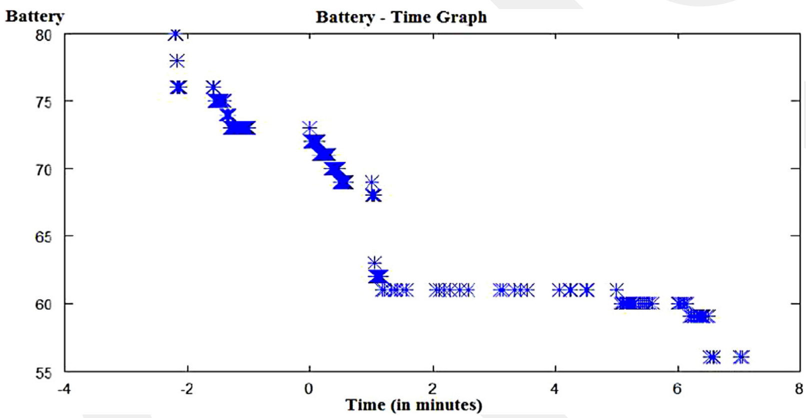

Fig. 3. Battery/Time data set when only GPS sensor is used for data gathering.

obtained battery consumption rates for all modes by using slope of linear cost function(h). When phone is in idle mode, gradient descent algorithm generates a cost function that has 0.6 as a slope. Consequently, battery reduces by 0.6 units per hour in the idle mode. The data set for idle mode and its cost function are demonstrated in Figs. 1 and 2.

When GPS is used to obtain location information for every minute, Gradient Descent algorithm generates a cost function that has 1.2 units as a slope. In other words, battery reduces by 1.2 units per hour when GPS positioning method is applied. Battery consumption rate is doubled according to the idle mode when location information is obtained from GPS sensors per minute. On the other hand, since, obtaining cell based parameters from cellular base stations has no additional cost over battery, estimating location via GSM signal fingerprints causes 0.6 battery consumption rate as in the idle mode. As a conclusion, battery consumption rate can be reduced by 50% by using GSM signal fingerprinting instead of GPS positioning. Associated data sets and cost functions that fit the data are illustrated in Figs. 3 and 4.

2.2. Estimating location from fingerprinting signal strengths

In the previous section, we measured battery consumption rate when GPS positioning is applied and we concluded that battery power consumption is doubled when GPS sensor is used to obtain location information instead of GSM signaling from cellular base stations which has no additional cost over battery. In our approach, when smart phone battery is insufficient, location data is gathered by using the signaling information that is already generated by the standard operation of mobile phone network. Since mobile phones are always connected with surrounding networks or base stations, the information about surrounding base stations can be obtained without any additional cost.

In order to estimate location of a phone with cell phone signal fingerprinting, we need to obtain base station coordinates and their associated identifier parameters. Although, many public cell ID databases are available for this purpose, most of them have poor coverage [27]. In our experiments, we use the database and system settings proposed in [28] due to its high coverage with 3.900.000 cell records. Location of a mobile phone is estimated via matching surrounding cell ID values and

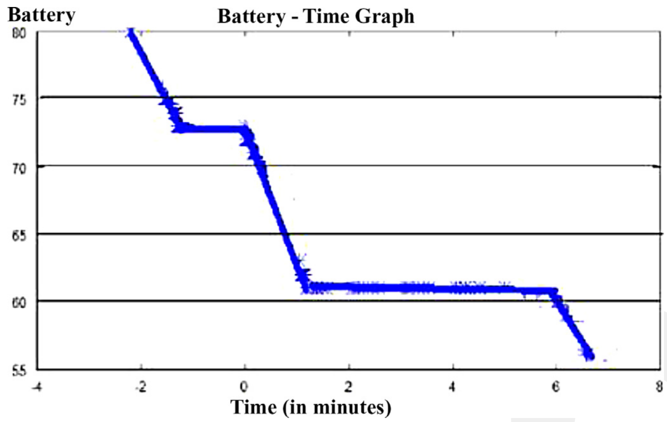

Fig. 4. Cost function that measures battery consumption rate when only GPS sensor is used for data gathering.

their associated coordinates with the values in the database [28]. By tracking cell phone signals to obtain spatio-temporal data in case of insufficient battery, we overcome the challenges emerge from utilization of only GPS sensors for mobile data gathering.

2.2.1. Handover notification

Location data which has been estimated by cellular positioning, can also be utilized for vehicular speed estimation. However, the essential issue in vehicular speed estimation is the adjustment of location update frequencies to obtain accurate results. A variety of sampling methodologies have been proposed so far to determine the data gathering frequency from smart phones. Most of the state-of-the-art work rely on temporal or spatial sampling where the former one refers to a continuous data transmission to a central server at predetermined time intervals regardless of its spatial status whereas the latter one denotes a data sampling when the vehicles cross a specific fixed point. In temporal sampling, a traffic data which consist of longitude, latitude, velocity and time-stamp parameters are transmitted to the central server within the specific time interval. Although, temporal sampling rate depends on the application type, in most of the cases, once-per-minute sampling frequency is sufficient [29]. However, in spatial sampling, the vehicles transmit the traffic data as they pass by each fixed location on the roadway independent of the time. This sampling strategy uses a similar method as the inductive loop detectors, LPR readers or RFID transponders do.

The main advantage of spatial sampling is that it forces the mobile phone to send the vehicle's traffic data at spatially important locations to ensure not to miss a critical spatio-temporal information. Although spatial sampling provides a rich amount of available traffic data, and hence increase the accuracy of traffic monitoring and estimation systems, the existence of large data will cause complications in terms of battery life of the phone, communication load on the cellular network, as well as well as privacy protection of the users who participate in the system. Therefore, effective optimization algorithms are required to overcome such challenges.

Although lots of methods have been proposed so far, most commonly used method for update frequency adjustment is based on handovers [27–29]. Overall objective of this method is to obtain cell information when phone signals are transferred from one cell to another. An anonymous string which consists of signal timing information and cell information is stored for each handover. When next handover occurs, location information is updated via performing cellular positioning method on new cellular information. Consequently, approximate distance traveled between handovers can be easily calculated. Since time information is also gathered with each update, average velocity of a phone can be estimated. To illustrate the usage of handover notifications for providing such services, we developed a service to collect location and time data. Our implementation collects GSM and sensor data on the phones where former one includes a list of GSM towers within a specific region and their distances from a smart phone where the data are collected, whereas the latter one consists of GPS coordinates along with a time-stamp. We update this data batch when the handover occurs to determine the vehicle velocities. The service also notifies users when handover occurs via a listener that listens for changes in cell ID values. Fig. 5 presents our experiments in detail.

2.3. System overview

In our system design, the speed information is extracted after the processing of raw data in several steps. First, a smart phone application is executed to perform the location measurements by using GPS sensors or GSM data. The data source is selected with respect to the battery status. For each randomly chosen vehicle, a battery look-up is executed to detect the current battery status for data source selection. The application collects a list of GSM towers, their associated identification numbers and GPS coordinates along with a time-stamp. For GPS positioning, the location data are directly transmitted via any available wireless network (3G, WiFi), whereas for cellular positioning additional processing is performed for location

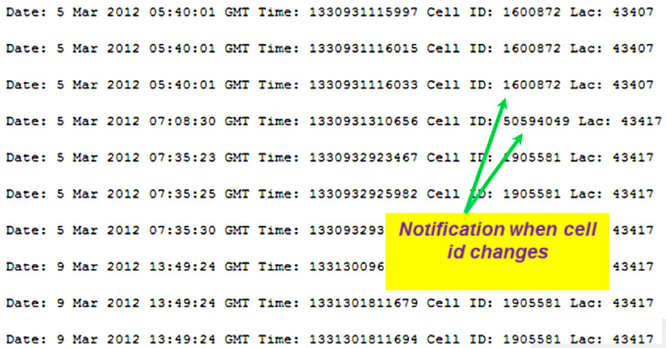

Fig. 5. Handover notification for vehicle speed estimation.

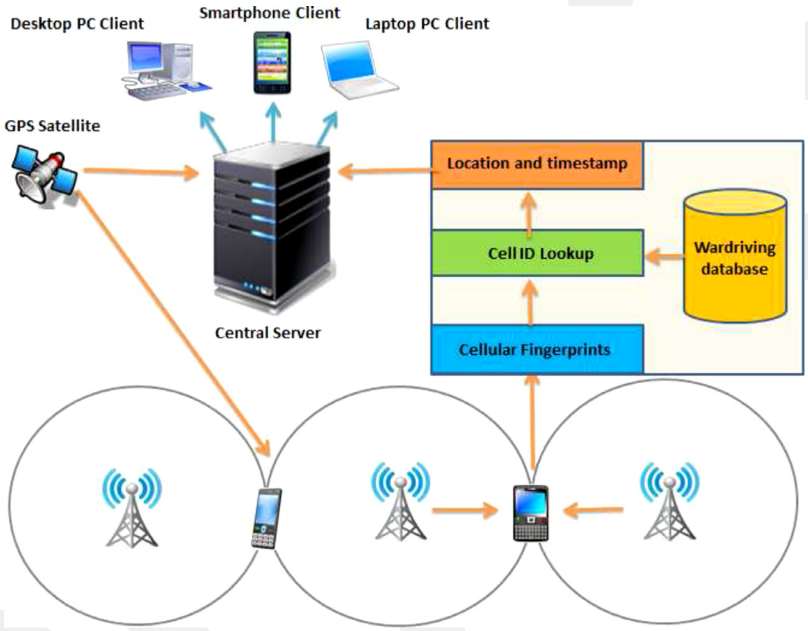

Fig. 6. Overview of the system architecture.

estimation which includes base station location look-up and distance calculations. After the extraction phase, the location information is transmitted to a central server for calculating speed information for further evaluation and analyses. Furthermore, an average speed information of a particular road segment where the vehicle resides can be sent directly back to the smart phone. Section 4 describes the average speed information extraction process by performing real-life and simulation experiments. Fig. 6 illustrates the system architecture in detail.

3. Experimental results

This section analyses the performance of our proposed model by presenting real-life and simulation experiments. The purpose of real-life experiments is to determine the location and speed accuracy of cellular positioning with respect to GPS-sensor based estimation. The experiment is performed with a vehicle on a particular 12 km road segment between Maltepe to Suadiye districts of Istanbul. The location and speed data are extracted with an Android application deployed in a smart phone. We update the measurements on each handover where a cell information is transferred from one cell to another. Fig. 7 illustrates the road segment where the real-life experiment is performed along with the location measurements obtained via cellular and GPS positioning techniques.

So far, we introduced a battery-friendly data gathering model for obtaining spatio-temporal data without suffering from the limitations of GPS sensors in terms of energy consumption and communication network overload. To evaluate the effectiveness of our model, we developed a Java-based microscopic traffic simulator with several battery consumption modules associated with it. Our proposed data gathering and extraction methods are fully simulated via the features provided by our traffic simulation tool.

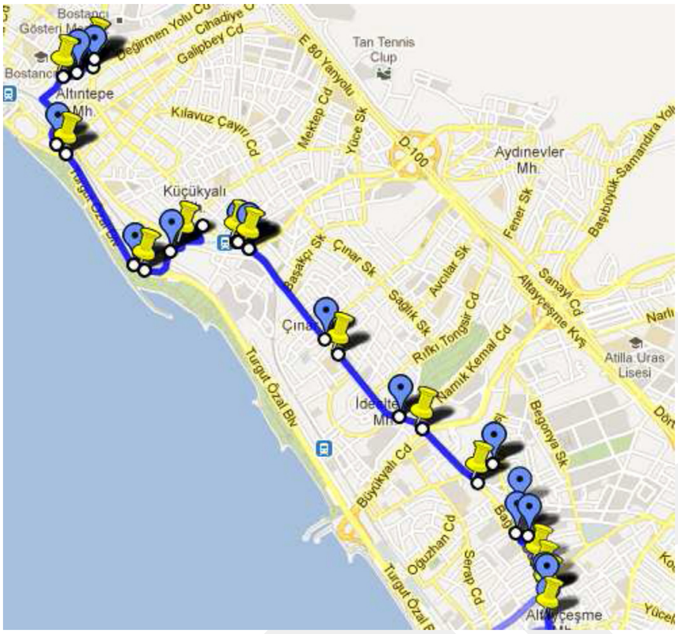

Fig. 7. Coverage map of the data set for the experiment.

In the data gathering phase, dynamic data source selection model, which is proposed in Section 2, is operated based on instant tracking of smart phone battery status. For this purpose, we apply the following assumptions for the new battery module that has been implemented to all vehicles simulated in the framework. According to the battery model, initial battery value is assigned as 100 for every vehicle. Next, every vehicle is associated with the modes that are explained in Section 2. Then, in every location update, batteries of vehicles are reduced with a rate specific to the mode that the vehicle is associated. For instance, in the case of GPS-only mode, where only GPS sensor is used for spatio temporal data gathering, battery value is decreased 0.8% in each update. Furthermore, in order to evaluate the performance of the system by considering real time limitations, we set the battery power dynamically in the cases where the phone battery has been charged or data gathering mode has been altered.

In the data extraction phase, the simulation framework allows us to create road sections with specified number of lanes to find out the average velocity and density for each section on the roadway to be simulated. This feature enables us to simulate the results of our data extraction methods by estimating the traffic status on each section depicted on the simulator. We also illustrate the results by coloring the segments to red, orange, yellow and green to denote congested, less congested, medium and smooth traffic flow respectively.

Analysis of the proposed methods have been conducted in two setups where the former one requires a fixed number of vehicles and continuous GPS data transmission while the latter ones samples available GPS sensor data dynamically throughout the simulation. Furthermore, we also present a real life scenario on a particular road segment to demonstrate the feasibility of our model for dynamic traffic environments.

3.1. Prototype implementation

We performed real-life experiments to evaluate our data acquisition model and tested its operation while the phone was carried out by a mobile user. For this experiment, we used the built-in GPS sensor on Android smart phone for gathering GPS location data, whereas for cellular positioning we used Ericsson's Cell API [28] to find out an estimated location based on the war-driving data. The Android OS Phone Library has been configured to record the ground truth GPS location data and cell tower fingerprints for every minute. Although, the data were gathered in every minute, for data analysis we only consider the changes in location data where a handover occurs. We performed real-time velocity estimation to assess the speed measurement accuracy of cellular positioning. Our data set covers 12 km of road segments driven. In these experiments, our goal is to demonstrate the location accuracy of cellular positioning with respect to GPS data which we accept as a ground truth for this case and presenting the benefits of cellular positioning as a proposed solution to the battery drainage problem in GPS sensors.

The data logged in the internal storage of smart phone and transmitted to the central server between specified time intervals for further evaluation. The trajectory data were processed to assess the feasibility of cellular positioning for location and velocity estimation. We use greatest circle distance [29] in our evaluation of accuracy which refers to shortest spherical distance between two points on the surface of Earth. Since, spherical geometry differs from the Euclidean geometry where

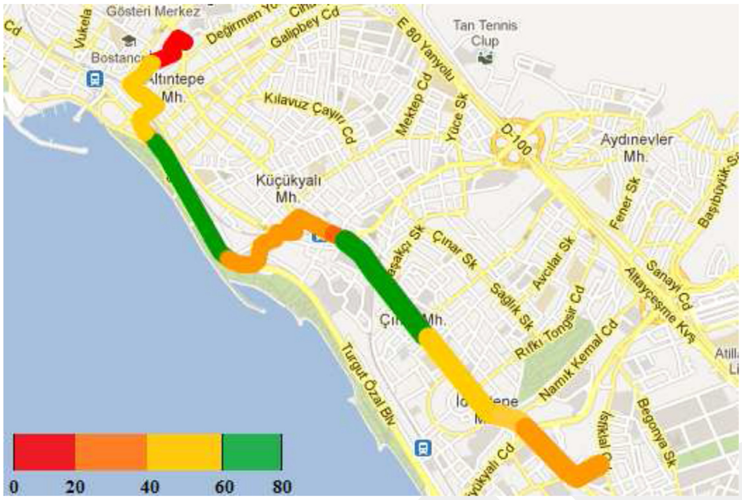

Fig. 8. Average velocity map for cellular data.

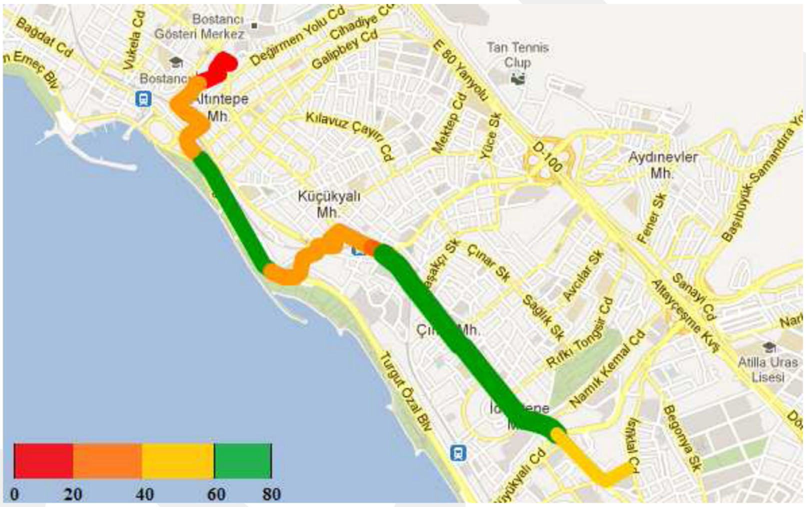

Fig. 9. Average velocity map for GPS data.

the straight line distance between data points have been used to find out the location difference, we define the accuracy as it is formulated in [30]. We tracked the ground truth data by sampling GPS feeds for every minute and presented the difference from cellular positioning data to find out the accuracy rate of our proposed data acquisition method. Fig. 7 demonstrates a map where yellow markers refer to GPS location data and purple markers stand for cellular positioning data. Our real time experiments result in 95% precision in both the mean and median, and a median geographic error of 5 m. The geographic error range varies from 5 to 12 m excluding the outliers. As it can be observed from the map, utilization of cellular data along with GPS sensor data for location estimation gives highly accurate results while reducing the energy consumption by half.

Moreover, hybrid data acquisition systems produce reasonable average velocity estimates for traffic monitoring purposes. To illustrate this, we conducted a real-life experiment in which a particular route is divided into different segments that have varying lengths due to the handover locations. The trajectory data projected through the GPS and cellular data along with the timestamp were used to extract the average speed in each section denoted by a road segment between two handovers. We classified the speed measurements into four parts according to the scale of the ground truth velocity data range from 0–80 km/h: highly congested conditions ($0 < v(d) < 20$), congested conditions ($20 < v(d) < 40$), medium flow conditions ($40 < v(d) < 60$), free flow conditions ($60 < v(d) < 80$) and illustrated the velocity classes with red, orange, yellow and green colors respectively. Figs. 8 and 9 show the map in which the route is divided with respect to handover locations on the road. Each segment colored to present the velocity class it belongs to. The mean and standard deviations are calculated in both of the cases. Our speed experiments result in 92% accuracy with a median speed error of 7 km/h. The speed error varies from 3 to 12 km/h. The average velocity measurements extracted via GPS and cellular data are presented in Fig. 10.

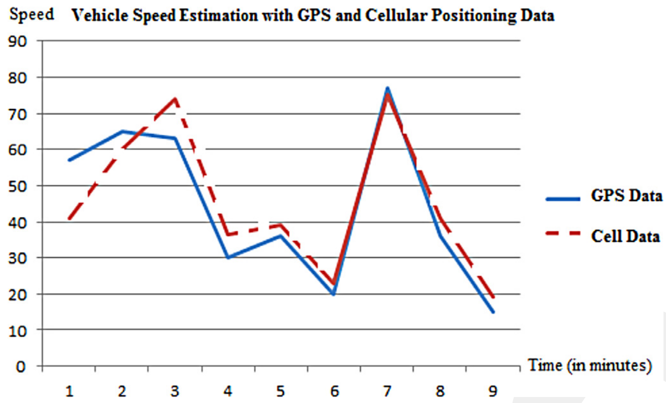

Fig. 10. Comparison of GPS based and hybrid data acquisition models without GPS sampling.

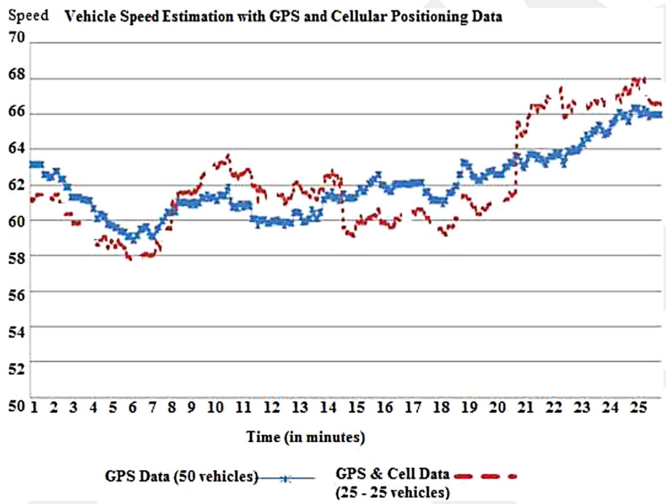

Fig. 11. Vehicle speed estimation using GPS and cell data.

3.2. Simulation

In addition to the real world experiments, we also performed a simulation to verify the efficiency of the hybrid data acquisition model for vehicle speed estimation. In these experiment, our goal is to measure the average accuracy and battery level of our hybrid acquisition model where the number of vehicles is set to a specific number and the number as well as the already chosen vehicles remain as the same during simulation. For the first simulation setup, following configurations have been set: (1) First, a route is divided into different sections; (2) Next, constant number of vehicles has been created; (3) Then for each vehicle, random speed values vary between 0–120 km/h range has been set. Overall simulation setup is illustrated in Fig. 13. Furthermore, in every minute we logged GPS location and cellular positioning data of 20 vehicles. We initiated our experiment by gathering all the location data from GPS sensors (m = 20, n = 0) where the m and n refer to number of vehicles which provide GPS location and cellular positioning data to the simulation, respectively. In each simulation cycle, we decreased the m parameter by 1 and increased the n parameter by using a one new log which have been gathered by cellular positioning. At the end of the simulation, we obtained 5.5 m median location error.

Moreover, we expanded our simulation tool by adding multiple battery modules with respect to the data gathering method as it is discussed in Section 2. In real life experiments, we found out that the battery reduces by 0.6 unit per hour in idle and cellular positioning mode whereas the overall battery level is reduced by 1.2 unit per hour in GPS-based data acquisition. In the simulation, we applied these values to demonstrate the battery consumption over time according to the data gathering method of each vehicle. Since GPS positioning doubles the battery consumption, utilization of cellular location data increases the average battery level of vehicles which involve in the data gathering process. The results are illustrated in Fig. 12. Our proposed hybrid data acquisition model gives 95 % per cent location accuracy. As it can be observed from the graph, our proposed system provides highly accurate results while reducing the battery consumption.

Fig. 11 demonstrates the average speed calculation with 50 vehicles in two cases: blue line represents 50 vehicles where only the GPS data were used for data gathering whereas in the red line half of the randomly selected vehicles collect the

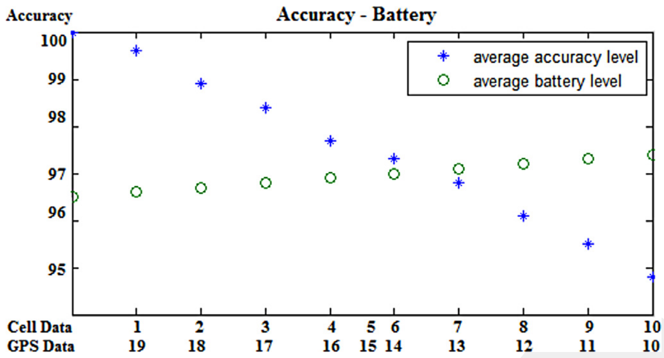

Fig. 12. Accuracy battery trade-off without GPS sampling.

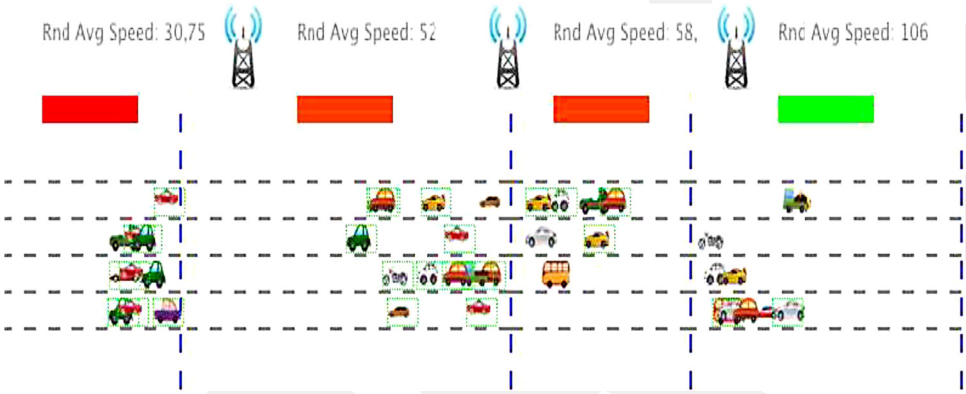

Fig. 13. The simulation environment for our proposed acquisition model.

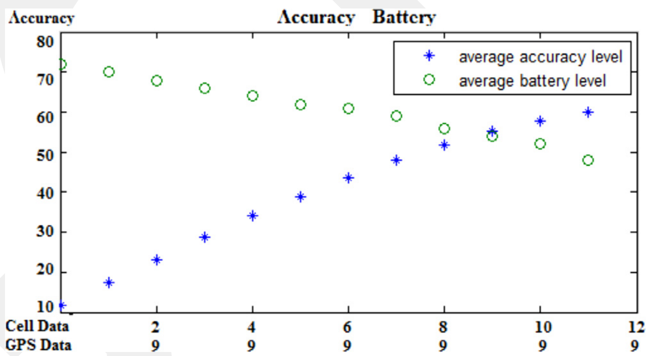

Fig. 14. Accuracy battery trade-off with GPS sampling.

location data via GPS and other half use cellular data for positioning. The randomly chosen vehicles are static in this analysis. The speed of the vehicles created in this setup are random values between 55 km/h and 70 km/h. The speed differences between two lines representing the setup are lower than 8 km/h. The median variation is around 2 km/h. Although the experiment enables us to simulate our proposed data gathering and extraction methods, such configurations of vehicles do not reflect a real case. Because in the real-life, vehicles do not continuously send GPS data to the server. The mobile application may be closed or run arbitrarily which leads to randomized effects in terms of simulating such real-life conditions. Hence the number of vehicles as well as the individual vehicles sending GPS information to the server keep varying in real-time (Fig. 14).

Although, simulation parameters are set based on the real life experiments, this experiment assumes that the location data is always available. However, in real life, there may be cases where the location data can not be obtained either via GPS or cellular positioning. In this analysis, our goal is to present a complete simulation of our proposed model by changing the number of vehicles which provide GPS data in every iteration. We performed a simulation experiment to illustrate this real-life scenario by applying GPS data sampling to random 20 vehicles in the simulation. The random selections of individual

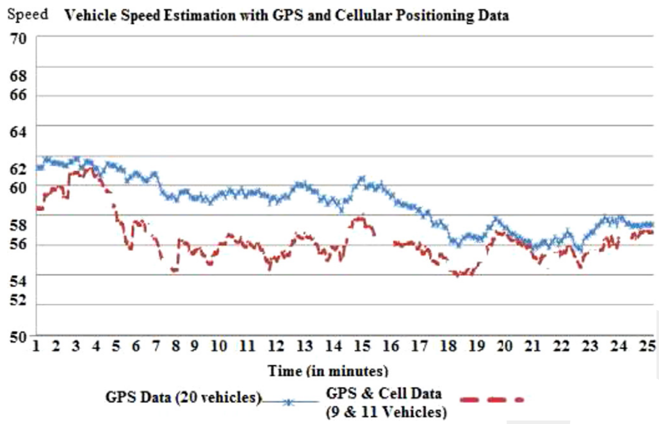

Fig. 15. Comparison of GPS based and hybrid data acquisition models with GPS sampling.

vehicles as well as determination of the random number of vehicles that are assigned to execute the task of transmitting GPS and cellular data (location, speed and time) to the server are repeated in every specified time interval (Fig. 15).

For the setup configuration we generated 20 vehicles and set the parameter m to 9 which refers to the number of vehicles which use GPS sensors for finding location, p to 11 which stands for the number of vehicles where GPS data is not available. In the initialization step, we set the parameter n to 0 to illustrate the case where cellular positioning is not applied to gather location data. Next, for every iteration, we increased the number of vehicles which use cellular positioning for data gathering to understand how utilization of cellular positioning method affects the overall accuracy and battery consumption. Moreover, we implemented the same battery modes discussed in Section 2 to measure the battery consumption in the sampling case. We then evaluate our results based on the ground truth data to obtain the accuracy of our traffic status estimation system. The results are illustrated in Fig. 12. Our proposed hybrid data acquisition model achieves 40% higher location accuracy in the case where only the vehicles (9 vehicles) which provide GPS data are used for calculating the average location. The results verify the benefits of the usage of cellular data in the cases where the GPS data are not available and the overall energy consumption is high.

The average velocity calculated with 20 vehicles indicated with the blue and red lines where only GPS data were collected for speed measurements and both GPS and cellular data were collected with a ratio of 1:4 (GPS data = 4, Cellular Data = 16), respectively. The randomly chosen vehicles are dynamic in this analysis. The speed of the vehicles created in this setup are random values between 55 km/h and 70 km/h. The speed differences between two lines representing the setup are lower than 10 km/h. The median speed variation is around 5 km/h.

4. Conclusions

Data gathering and extraction phases are important parts of traffic monitoring systems. Utilization of cost effective and highly accurate traffic monitoring devices and processing the raw data that have been gathered with the efficient data extraction methodologies are required in order to come up with accurate traffic information. However, data extraction phase has solid challenges such as the determination process of locations that are contained by successive segments. Furthermore, data gathering phase has some issues in terms of accuracy and cost.

In this paper we propose a hybrid data acquisition model which uses GPS sensors and cell phone fingerprints to reduce the energy consumption while preserves the required accuracy level as a solution to GPS battery drainage problem. The model has significant benefits in the cases where the phone battery is depleted or the user does not want to send the data and/or closes the application. Furthermore, we evaluated the effectiveness of applying cellular positioning techniques along with GPS-sensor based data collection via performing experiments on energy consumption and location accuracy for various data acquisition modes where only GPS positioning, only cellular positioning and both of the techniques are executed. Our real-life experiments reveal the fact that the battery consumption rate can be reduced by 50% by using GSM signal fingerprinting instead of GPS positioning.

Moreover, we performed experiments to compute the average velocity by extracting location data with our hybrid data acquisition model. A real time microscopic traffic simulator has been developed to evaluate the method. Optimization and calibration methodologies for data gathering process have been studied to achieve effective results. The model is improved by considering real life limitations that mostly occur when data cannot be obtained from some of the sensors. System architecture and battery consumption model of the traffic simulator have been introduced in detail. Real life and simulation experiments were conducted to illustrate the cases in which fixed number of vehicles are required for continuous data transmission while the latter ones randomize vehicle selection for data transmission dynamically throughout the simulation. Our proposed model results in 88% location accuracy whereas it reduces the battery consumption by half.

References

[1] Wright J, Dahlgren J. Using vehicles equipped with toll tags as probes for providing travel times. In: California PATH Working Paper UCB-ITSPWP-2001-13, Institute of Transportation Studies, UC Berkeley, CA; 2001.
[2] Herrera JC, Work D, Ban J, Herring R, Jacobson Q, Bayen A. Evaluation of traffic data obtained via GPS-enabled mobile phones: the mobile century experiment. Transp Res Board Part C 2010;18(4):568–83.
[3] Bertini R, Lasky M, Monsere C. Validating predicted rural corridor travel times from an automated license plate recognition system: Oregon's frontier project. In: Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems; 2005. p. 706–11.
[4] Department of Transport. Intelligent Transportation Systems in United Kingdom: Initial Report, Technical Report, UK, August 2011.
[5] Sanwal KK, Walrand J. Vehicles as probes. California PATHWorking Paper UCB-ITS-PWP-95-11. Institute of Transportation Studies, University of California, Berkeley, CA; 1995.
[6] Zito R, Este GD, Taylor M. Global positioning systems in the time domain: how useful a tool for intelligent vehicle highway systems. Transp Res Board, Part C 1995;3(4):418–25.
[7] Kwon J, McCullough B, Petty K, Varaiya P. Evaluation of peMS to improve the congestion monitoring program. In: California PATH Research Report UCB-ITS-PRR-2007-6, Institute of Transportation Studies, University of California, Berkeley, CA; 2007.
[8] Ygnace J-L, Drane C, Yim YB, de Lacvivier R. Travel time estimation on the san francisco bay area network using cellular phones as probes. In: California PATH Working Paper UCB-ITS-PWP-2000-18, Institute of Transportation Studies, University of California, Berkeley, CA; 2000.
[9] Yim YB, Cayford R. Investigation of vehicles as probes using global positioning system and cellular phone tracking: field operational test. In: California PATH Working Paper UCB-ITS-PWP-2001-9, Institute of Transportation Studies, University of California, Berkeley, CA; 2001.
[10] Yim YB. The state of cellular probes. In: California PATH Working Paper UCB-ITS-PRR-2003-25, Institute of Transportation Studies, University of California, Berkeley, CA; 2003.
[11] Menard T, Miller J, Nowak M, Norris D. Comparing the GPS capabilities of the samsung galaxy s, motorola droid x, and the apple iphone for vehicle tracking using freesim_mobile. In: Proceedings of International IEEE Conference on Intelligent Transportation Systems (ITSC); 2011. p. 985–90. 5–7.
[12] Myr D. Traffic information gathering via cellular phone networks for intelligent transportation systems. In: United States Patent; 2003. p. 10.
[13] Valerio D. Road traffic monitoring from cellular network signaling, Technical Report, FTW-TR-2009-003, March 2009.
[14] Final evaluation report for the CAPITAL-ITS operational test and demonstration program. Technical report. University of Maryland; College Park; 1997. University of Maryland Transportation Studies Center.
[15] Remy J. Computing travel time-estimates from GSM signaling messages: the STRIP project. In: Proceedings of IEEE Intelligent Transportation Systems Conference, Oackland, CA, USA; 2001.
[16] Fehmi Ben Abdesslem, Andrew Phillips, Tristan Henderson. Less is more: energy-efficient mobile sensing with senseless. In: Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhold. Barcelona, Spain; 2009. p. 17–19.
[17] Valerio D, D'Alconzo A, Ricciato F, Wiedermann W. Exploiting cellular networks for road traffic estimation: a survey and a research road map. In: Proceedings of the IEEE 69th Vehicular Technology Conference (IEEE VTC '09-Spring). Barcelona, Spain; 2009. p. 26–9.
[18] Ramm K, Schwieger V. Mobile positioning for traffic state acquisition. Journal of Location Based Services, Universitat Stuttgart, Institute for Applications of Geodesy to Engineering 2007;1:133–44.
[19] Birle C. Via GSM network data to traffic information in real time. In: Proceedings of International Conference on Intelligent Transport Systems, Munich; 2007.
[20] Cellint. Source: http://www.cellint.com, Dec. 2012.
[21] Avni O. Performance and limitations of cellular-based traffic monitoring technologies: based on case study of bench marking cellint's trafficsense with road sensors. In: Proceedings of ITE 2007 Technical Conference, San Diego; 2007.
[22] Holdings I. GSM based traffic information services. In: Proceedings of International Conference on Intelligent Systems, University of Sydney; 2005.
[23] Thiagarajan A, Ravindranath L, LaCurts K, Toledo S, Ericsson's J. Vtrack: accurate, energy-aware road traffic delay estimation using mobile phones. In: Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys); 2009.
[24] Tom Tom Technical Report, Travel Time Measurements using GSM and GPS Probe Data, Source: http://www.tomtom.com/lib/doc/download/HDT_White_Paper.pdf, October, 2006.
[25] Singh R, Bhargava P, Kain S. Smart phones and interactive reports leave traffic in the rear view mirror. IEEE Trans Intell Transp Syst 2008;27:33–8.
[26] Ericsson Mobile API, Source: https://labs.ericsson.com/apis/mobile-location/, June 2012.
[27] Gundlegard D, Karlsson JM. Handover location accuracy for travel time estimation in GSM and UMTS. In: Proceedings of IEEE Conference on Intelligent Transportation Systems; 2009. p. 87–94.
[28] Hoh B, Gruteser M, Herring R, Ban J, Work D, Herrera JC, et al. Virtual trip lines for distributed privacy-preserving traffic monitoring. In: Proceedings of International Conference on Mobile Systems, Applications, and Services. ACM; 2008. p. 15–28.
[29] Lin K, Kansal A, Lymberopoulos D, Zhao F. Energy-accuracy trade-off for continuous mobile device location. In: Proceedings of International Conference on Mobile Systems, Applications, and Services; 2010. p. 285–98.
[30] Thiagarajan A, Ravindranath LS, Balakrishnan H, Madden S, Girod L. Accurate low energy trajectory mapping for mobile devices. In: Proceedings of Symposium on Networked Systems Design and Implementation; 2011.

Sevgi Kaya received her B.S. degree in Electrical and Computer Engineering from Bahcesehir University, Istanbul, Turkey in 2013. She is currently doing her M.S. in computer science at ETH Zurich, Switzerland. She worked as a research assistant at Stanford University, Microsoft Research and CERN. Her research interests are in applied machine learning, data mining, smart sensor based systems and optimization.

Necati Kilic received his B.S. and M.S degree in Electrical and Computer Engineering from Bahcesehir University, Istanbul, Turkey. His research interests are in wireless communications, smart grid, smart transportation and machine to machine communication.

Taskin Kocak received his Ph.D. degree in Electrical and Computer Engineering from Duke University, USA in 2001. He is currently a Full Professor and Dean of Engineering at Bahcesehir University, Turkey. Previously, he was on the faculties of Bristol University, UK and the University of Central Florida, USA. His research interests are in computer communications with over 100 peer-reviewed publications.

Cagri Gungor received his B.S. and M.S. in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey. He received his Ph.D. in Computer Engineering from Georgia Institute of Technology, Atlanta, USA, in 2007. Currently, he is an Associate Professor and Chair at Abdullah Gul University, Kayseri, Turkey. His research interests are in smart grid, next-generation wireless networks.