Journal of Intelligent Transportation Systems: Technology, Planning, and Operations
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gits20

A Real-Time Parking Prediction System for Smart Cities

Eleni I. Vlahogianni[a], Konstantinos Kepaptsoglou[a], Vassileios Tsetsos[a] & Matthew G. Karlaftis[a]
[a] Department of Transportation Planning and Engineering, National Technical University of Athens, Athens, Greece
Accepted author version posted online: 11 Apr 2015.Published online: 11 Apr 2015.

Journal of Intelligent Transportation Systems, 00:1–13, 2015

ISSN: 1547-2450 print / 1547-2442 online
DOI: 10.1080/15472450.2015.1037955

A Real-Time Parking Prediction System for Smart Cities

ELENI I. VLAHOGIANNI, KONSTANTINOS KEPAPTSOGLOU, VASSILEIOS TSETSOS, and MATTHEW G. KARLAFTIS†

Department of Transportation Planning and Engineering, National Technical University of Athens, Athens, Greece

A methodological framework for multiple steps ahead parking availability prediction is presented. Two different types of predictions are provided: the probability of a free space to continue being free in subsequent time intervals, and the short-term parking occupancy prediction in selected regions of an urban road network. The available data come from a wide network of on-street parking sensors in the "smart" city of Santander, Spain. The sensor network is segmented in four different regions, and then survival and neural network models are developed for each region separately. Findings show that the Weibull parametric models best describe the probability of a parking space to continue to be free in the forthcoming time intervals. Moreover, simple genetically optimized multilayer perceptrons accurately predict region parking occupancy rates up to 30 minutes in the future by exploiting 1-minute data. Finally, the real time, Web-based, implementation of the proposed parking prediction availability system is presented.

Keywords Duration Modeling; Internet of Things; Neural Networks; Parking Occupancy; Parking Sensors; Smart City

INTRODUCTION

Parking availability prediction is among the most important factors affecting both private car-based trip decisions and traffic conditions in urban areas. Drivers' decisions are temporally dependent, implying that they are influenced by past experience, as well as real-time (on-road) perceptions. Parking is such a case where prior knowledge on possible prevailing conditions (e.g., difficulty in finding a parking space, off-street parking costs, etc.) affects drivers' parking decisions. At the same time, vehicles in search of free parking spaces negatively impact traffic conditions and the environment. In this context, parking information provision is a research area of particular interest, since modern communication technologies offer alternative ways of delivering information to travelers in a timely and effective manner.

Various systems developed for providing parking information and guidance have been proposed by researchers in the past (Caicedo, Blazquez, & Miranda, 2012; Giuffrè, Siniscalchi, & Tesoriere, 2012; Liu, Lu, Zou, & Li, 2006; Oh, Lee, Kim, & Yang, 2002); parking information is usually disseminated by variable message signs or through the Internet, cellular phones, personal digital assistant (PDA), and geographic information system (GIS) technologies (Giuffrè et al., 2012; Liu et al. 2006). Teodorovic and Lucic (2006) argue that although parking guidance systems may not affect the occupancy rate or average parking duration, drivers tend to greatly appreciate the information provided by such systems. This is because such systems have certain positive effects on a city's traffic operations, as well as on the personal trips of road users. First, they significantly increase the probability of finding free parking spaces and mitigate frustration of those drivers/visitors unfamiliar with the city center. They are supposed to decrease queues in front of parking garages and decrease total vehicle-miles traveled (particularly in the city center). Additionally, they help road users to optimize their trips and thus improve vehicles' energy consumption and decrease emissions.

The usefulness of such predictive parking information is straightforwardly understood. If all drivers act without information and make "uninformed" choices, they will probably resort to similar optimal decisions leading to induced long waiting times, queues, and increased parking circling. On the other hand, dissemination of accurate and timely parking availability

Address correspondence to Eleni I. Vlahogianni, Department of Transportation Planning and Engineering School of Civil Engineering, National Technical University of Athens, 5 Iroon Polytechniou Str, Zografou Campus 15773, Athens, Greece. E-mail: elenivl@central.ntua.gr

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/gits.

information may lead to improved driver decisions and parking searching (Caicedo, Robuste, & Lopez-Pita, 2006). From a transport operator perspective, accurate parking availability predictions may lead to better management of the system, congestion mitigation due to queue formation avoidance, and so on.

The objective of this article is to exploit statistical and computational intelligence methods for developing a methodology that can be used for multiple steps ahead on-street parking prediction in "smart" urban areas. This work takes advantage of massive real-time parking data availability, obtained by an extended parking sensor network available in the "smart" city of Santander, Spain. Models are developed for predicting expected parking occupancy along with the probability of finding free parking spaces. Traditional survival analysis models, as well as neural network models, are developed. The methodology is evaluated and a real-time, Web-based system exploiting the proposed prediction models, for the city of Santander, Spain, is presented.

PARKING MODELING: OBJECTIVES AND METHODS

Parking modeling has been a topic of interest since the 1970s and has been studied from various angles, including parking patterns, impacts on traffic, off-street parking technologies, parking policy, choice and location models, and economic models describing parking conditions (Arnott & Rowse, 1991; Bonsall & Palmer, 2004; Hensher & King, 2001; Ibeas, Cordera, dell'Olio, & Moura, 2011; Van Der Goot, 1982; Wong, Tong, Lam, & Fung, 2000). An early review by Young, Thompson, and Taylor (1991) reported three categories of parking related models: driver behavior with respect to parking (parking choice), optimal positioning of parking lots (parking allocation), and interaction of parking operations with other transportation system elements and infrastructures (parking interaction).

Parking availability prediction, on the other hand, coupled with modern capabilities on data collection and processing, along with intelligent transportation systems (ITS)-based exploitation of such information, although indicated as a challenge since the early 1990s (Bonsall, 1991), has been introduced in recent years. David, Overkamp, and Scheuerer (2000) proposed a model for event-oriented forecasting of parking occupancy based on standardized daily distribution occupancy rate curves and online data obtained from parking lots equipped with detectors. On-street occupancy estimation in the absence of detectors was investigated by David and Keller (2001). An event-driven model was developed for that purpose using historical socioeconomic and parking specific data; the model was successfully validated for the city of Munich, Germany. Teodorovic and Lucic (2006) developed a system based on fuzzy logic, simulation, and optimization models, which decides whether to accept or reject new parking requests in real time, according to estimated availability for parking lots. Martens and Benenson (2008) exploited agent-based modeling for representing parking behavior by integrating the effect of considering real-time and expected parking availability, prices, and parking enforcement. Caicedo (2009) developed a discrete choice model for combining online and historical data for real-time, off-street parking availability prediction; this model was later used by Caicedo et al. (2012). Fabusuyi, Hampshire, Hill, and Sasanuma (2011) developed ParkPGH, a system for predicting parking availability in eight Pittsburg parking facilities using historical and real-time data. A fuzzy logic model for estimating the uncertainty of peak period parking availability in park-and-ride facilities was proposed by Chen, Xia, and Irawan (2013).

Modern technological advances have revolutionized the ways of monitoring and recording transportation operations and data especially in the case of parking (Thornton, Redmill, & Coifman, 2014). These have, however, been exploited for parking prediction to a lesser extent, and have mainly focused on off-street parking facilities.

Smart Cities and Parking Prediction Challenges

Cities are characterized as "smart" when their transportation and communication (information and communications technology, ICT) infrastructures along with their human and social capital investments cooperate and actively support sustainable growth and high quality of life, through participatory action and engagement, while preserving natural resources (Caragliu, Del Bo, & Nijkamp, 2009). Indeed, as noted by Komninos (2009), innovation and use of ICT for improving capacity of infrastructures are key elements of "smart" cities. A novel type of "smart" city infrastructure, applicable to the transportation sector, is the so-called Internet of Things (IoT). IoT consists of a variety of devices or objects—such as radiofrequency identification (RFID) tags, sensors, actuators, mobile phones, and so on—which, through unique addressing schemes, are able to interact with each other and cooperate with their neighbors to reach common goals (Atzori, Iera, & Morabito, 2010). By continuously collecting, analyzing, and redistributing transportation information, IoT networks can offer valuable, real-time information to both travelers and operators, and thus support and improve the operations of ITS, traffic, and public transportation systems. Although one can trace several reasons that may prevent IoT being fully developed in urban environments, this unique technological paradigm is expected to substantially support sustainable development of future smart cities (Vlacheas et al., 2013).

The SmartSantander project is such a case of an IoT architecture deployed in the city of Santander, to achieve a massive deployment of sensors and network communications in order to provide efficient and equitable transportation and other services to citizens (http://www.smartsantander.eu). The IoT network of the SmartSantander project consists of (Gutiérrez et al., 2013):

- IoT nodes: These are responsible for sensing and collecting information from the natural and socioeconomic environment

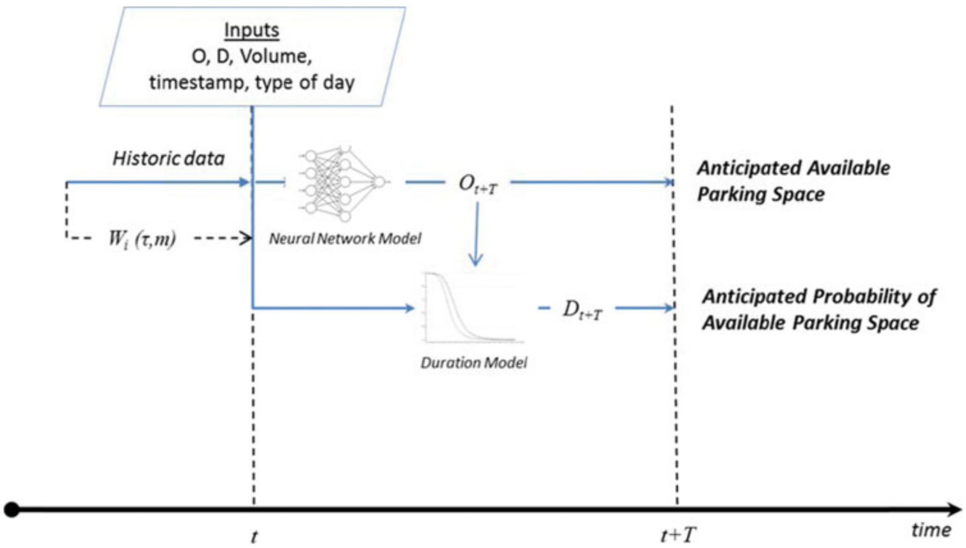

Figure 1 Functional architecture of the proposed parking occupancy prediction scheme.

and activities (temperature, CO, noise, light, car presence, etc.).
- Repeaters: These nodes are placed high above ground in street lights, semaphores, information panels, and so on, and behave as forwarding nodes of IoT information.
- Gateways: Both IoT nodes and repeaters are configured to send all the information to gateways. Once information is received by gateways it can be either stored locally or sent to central processing units through different interfaces (e.g., WiFi, GPRS/UMTS or Ethernet).

SmartSantander and similar test beds fuse high-resolution data sets stemming from both static and mobile sensors. Such data refer to macroscopic traffic flow in road sections of interest, detailed parking information in real time in urban areas, and high-resolution transit information. In the case examined in this article, having parking sensors collecting real-time information on their occupancy provides accurate, high-resolution parking information. New technologies can replace common parking metrics of average duration, turnover rate, and occupancy in extended time windows of 1 to 3 hours (which are manually collected in small-scale regions) with accurate, both aggregate and disaggregate, information on occupancy and parking duration. In this framework there are several questions that may arise; this article focuses on the following:

- Is it possible to predict parking occupancy using time-series modeling approaches based on data collected from an IoT network of sensors?
- How accurate are parking occupancy predictions produced by such models in relation to the predictive horizon?
- What are the statistical properties of parking space duration and how can we predict the probability of having free parking spaces in the area of interest?

MODELING PARKING AVAILABILITY BASED ON SENSOR DATA

Methodology

The architecture of the proposed parking occupancy prediction system is presented in Figure 1. In each sliding time window T, parking efficiency is defined by the following metrics:

- Duration ($\bar{D}_t$) of free parking space: the average time duration that a slot is free, over a certain time period.
- Occupancy (O_t): the percentage of parking slots occupied during a predefined time period.

The free space duration is an indication of how often a parking space becomes available, whereas occupancy relates the parking accumulation, meaning the number of parked vehicles in the study area at any specific time period, to the parking capacity.

The proposed methodology has two modules. The first module is a real-time time-series occupancy prediction scheme based on recurrent artificial neural networks. Simple yet flexible memory mechanisms will be applied in order to replicate the temporal dynamics of parking occupancy O_t. The model is presented with past information of occupancy ($O_{t-\tau}$, ..., $O_{t-(m-1)\tau}$) to predict occupancy one step ahead, O_t. The characteristics of the model's memory (the time delay τ and the dimension m) will be evaluated through a nonlinear analysis of the dynamics of parking occupancy. The use of exogenous variables, such as the type of day or day of week, will be evaluated relative to the improvement on the accuracy of predictions they impose.

When considering short-term prediction systems that operate in real time and in an "intelligent" technology-based environment, the effectiveness depends, mostly, on predicting traffic information in a timely manner (Smith & Oswald, 2003, (Vlahogianni, Karlaftis, Golias, & Kourbelis, 2006). Real-time system effectiveness depends both on the results and on the time in which these are produced (Shin & Ramanathan, 1994). The computational time for making a prediction mainly depends on the functional form of the prediction system; empirical results show that data-driven prediction systems that include recursive data-search algorithms exhibit "best" prediction accuracy but need extensive computational time for convergence at acceptable results (Smith & Oswald, 2003; Vlahogianni, Karlaftis, & Golias, 2014). For this, the final structure of the prediction model is kept as simple and flexible as possible. The trade-off between simplicity and efficiency is studied in a preliminary stage of analysis using genetic algorithms to optimize the structural and learning parameters of the different models.

The second is a static approach for estimating the probability of finding available parking space with relation to a series of factors such as the type of day (weekday, weekend) and the time period (peak, off-peak, morning, evening). The modeling will be based on survival analysis. Parametric hazard based modeling may be developed under a variety of functional forms. The model will be also presented with the predicted value of occupancy as provided by the first module. The output of the model (the probability of finding an available parking space and the anticipated number of available space) can be visually depicted using various graph methods (such as heat maps), to provide useful information to users.

Neural Network Models for Time Series Prediction

Time-series modeling is a popular approach for making predictions in transportation problems (Karlaftis & Vlahogianni, 2011). This approach is suitable for analyzing parking occupancy due to the temporal structure of the performance measures of parking systems. A common prediction strategy implemented in transportation problems is based on the autoregressive moving average family of models. These models are relatively straightforward mathematically and easy to produce; however, they are severely constrained by stationarity and linearity, characteristics that most frequently violated in real transportation time series. Treating nonstationarity and nonlinearity may lead to a tedious process without achieving the desired levels of accuracy in predictions and modeling reliability (Washington, Karlaftis, & Mannering, 2010). Neural networks (NNs) for time series provide a good alternative, as they relax many of these constraints and also appear to provide short-term forecasting models that are more adaptable to sudden shifts in the data (Vlahogianni, 2009). A recent study on traffic time-series prediction has shown the structural equivalencies between nonlinear univariate and multivariate ARIMA models with exogenous variables and dynamic forms of multilayer perceptrons (MLPs) (Vlahogianni & Karlaftis, 2013b). The simplest of all is the NAR(p) structure of order p. In general, The MLP presented with p lagged values of parking occupancy O_k ($k = 1, \ldots, n$) may act as a predictor of the form (Mandic & Chambers, 2001):

$$\hat{O}_k=\hat{h}(O_{k-1}, ..., O_{k-p})=\sum_{i=1}^{I} W_i f\left(\sum_{j=1}^{p} w_{ij} O_{k-j}+\theta_i\right) \quad (1)$$

where $f(\cdot)$ is a smooth monotonic function, W_i and w_{ij} are the weights of the connections (synapses)—coefficients estimated through learning (training), thereby obtaining an estimate of the nonlinear approximation $\hat{h}$. The network converges to an estimate of $\hat{h}$ by minimizing the residuals $\sum_k \left(O_k - \hat{O}_k\right)^2$ using a learning algorithm (usually back-propagation).

For a time-series consideration in an NN framework, the MLP should be modified to account for the time sequence of events under study. This is usually accomplished by adding memory structures in the MLP that retain the effect of past information to the system and use it during learning. The memory is accomplished using local—at an intraneuron level—and global—between neurons of different layers—recurrent connections in a neural network. Memory mechanisms may be of a simple tap delay form, realizing structures:

$$\mathbf{O}(t) = \{O(t-\tau), ..., O(t-(m-1)\tau)\} \quad (2)$$

where τ is the delay and m the dimension of the horizon of past information introduced to the model or of other more complex mathematical forms (Gamma memory etc.). The networks with memory usually require cumbersome and slow learning procedures that may not be always stable. To avoid this, static MLPs can be externally modified to represent the temporal characteristics of transportation time series (e.g., parking occupancy) in a manner resembling the common statistical prediction techniques. The introduction of such data inputs in MLPs that are unchanged in their internal structural logic may conceptually approximate very complex and multivariate statistical structures with efficiency equal to that of classical MLPs (Vlahogianni, Karlaftis, & Golias, 2005).

Hazard-Based Free Parking Duration Modeling

Hazard-based duration modeling deals with the statistical representation of time-to-event data, a very frequent form of data in transportation problems; typical examples of such transportation data are the time to clear an incident (Chung, 2010; Vlahogianni & Karlaftis, 2013a), the time until the end of congested phenomena (Stathopoulos & Karlaftis, 2002; Vlahogianni, Karlaftis, & Kepaptsoglou, 2011), the time until the end of transit vehicle repair (Karlaftis, 2011), the time to an activity (Habib, 2012; Zhong & Hunt, 2010), the time to household evacuation under the emergence of a physical disaster (Hasan, Mesa-Arango, & Ukkusuri, 2013), the time for a pedestrian to cross a signalized intersection approach (Tiwari, Bangdiwala, Saraswat, & Gaurav, 2007), the time to vehicle transaction (Rashidi, Mohammadian, & Koppelman, 2011), the time to complete an overtake (Vlahogianni, 2013), and so on. Extensive

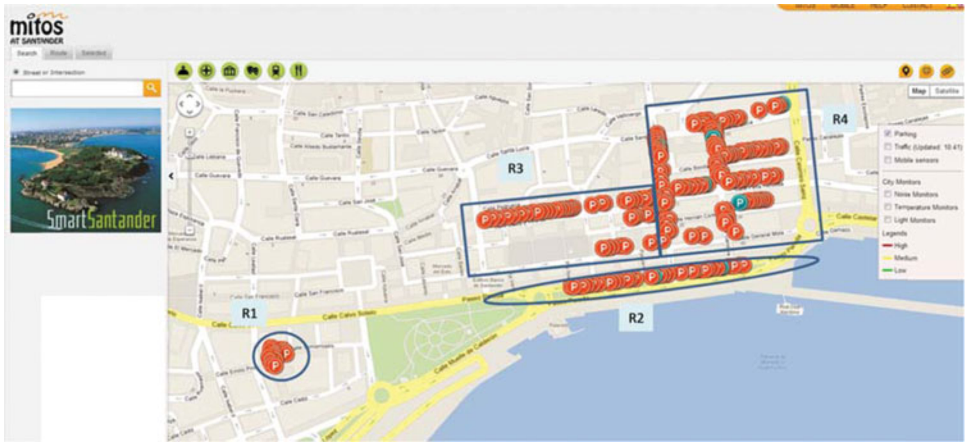

Figure 2 Parking sensors location in the area of Santander (map accessed at: http://mitos.ploigos.gr).

review on the hazard based transportation modeling applications may be found in Hensher and Mannering (1994) and Bhat (2000). Moreover, methodological, computational, and estimation issues in duration modeling with focus on transportation problems may be found in Washington et al. (2010).

Parametric hazard-based modeling is based on two concepts: the survival function and the hazard function. Let T be a nonnegative random variable representing the time a vehicle occupies a space; the survival function is defined as the probability that T is of length at least t (i.e., a parking space is occupied at least t minutes) and is given by:

$$S(t) = P(T > t) = 1 - F(t), 0 < t < \propto, \tag{3}$$

where $F(t)$ is the cumulative probability. The survival function $S(t)$ can have a variety of shapes following certain restrictions; it is bounded by 0 and 1 and, as T cannot be negative, $S(0) = 1$. Moreover, with larger t, S never increases (and usually decreases).

For continuous survival data, the hazard function specifies the instantaneous failure rate at $T = t$ conditional upon survival to time t and defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \tag{4}$$

where $f(t)$ is the probability density function:

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt} \tag{5}$$

By combining Eqs. 2 and 3 we get:

$$h(t) = -\frac{d}{dt}\log S(t) \tag{6}$$

The hazard function *h(t)* is always nonnegative and, unlike survival functions, has no upper bound. Parametric hazard based models may have a range of different functional forms aiming at modeling different distributional characteristics. The hazard rate $h(t|\mathbf{X})$ with covariates is given by:

$$h(t|\mathbf{X}) = h_0(t)\exp(\beta\mathbf{X}) \tag{7}$$

where $h_0(t)$ is the baseline hazard rate, β is a vector of estimated parameters, and $\mathbf{X}$ is a vector of covariates.

IMPLEMENTATION AND RESULTS

Data Set

Parking information comes from a network of 400 sensors located in the area of Santander (Figure 2). Parking sensors based on ferromagnetic technology, buried under the asphalt, have been installed at the main free parking areas of the city center, in order to detect parking sites availability in these zones. The parking sensors are deployed in a regulated zone (working days from 10:00 to 14:00 and from 16:00 to 20:00); there is a 2-hour time limit in the parking duration with a fare of 0.12 €/min. The preceding restrictions do not apply for citizens living in the study area.

The information from the parking sensors is wirelessly transmitted to the IoT corresponding repeaters that are installed in high-rise places (e.g., street lights, semaphores, information panels, etc). The communication between IoT sensors and repeaters is done through 802.15.4 protocol. All information is directed (through 802.15.4 protocol) and stored to the getaways (more details on the infrastructure description may be found at http://www.smartsantander.eu/index.php/testbeds/item/132-santander-summary).

The available data set consists of time series of the state (free/occupied) of each sensor every 1 minute for the months April to September 2013. With this information, it is possible to straightforwardly calculate parking metrics such as parking accumulation, occupancy, duration, and so on. In this study we focus on (a) parking occupancy (%) of a specific area, which is the percentage of parking spaces that are occupied by vehicles within a time interval; (b) the turnover rate (vehicles/space/t), which is equal to the number of vehicles per parking space in a time interval; and (c) the duration that a parking space is not occupied by vehicles.

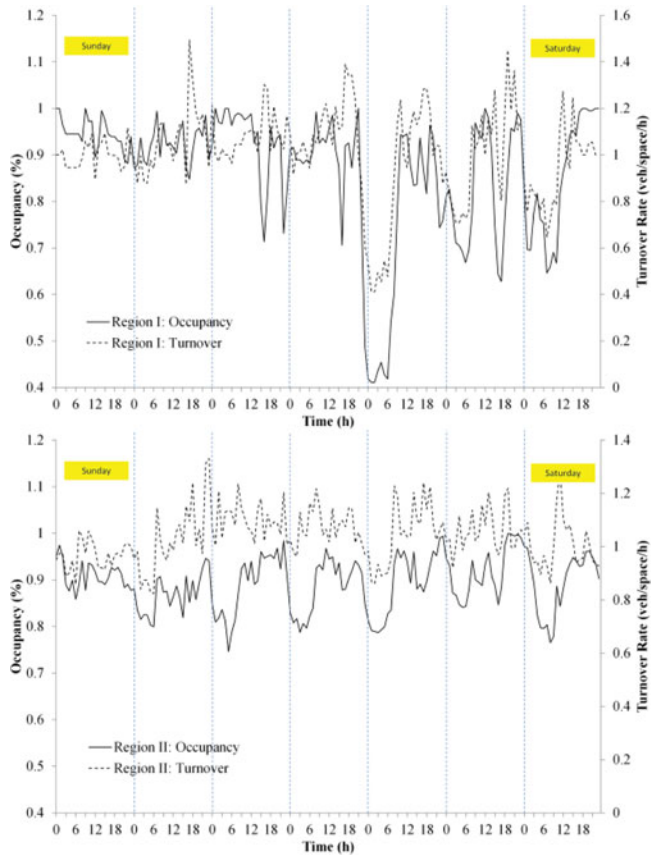

Figure 3 Hourly evolution of parking occupancy (%) and turnover rate (vehicles/space/h) for Regions I and II for a typical week.

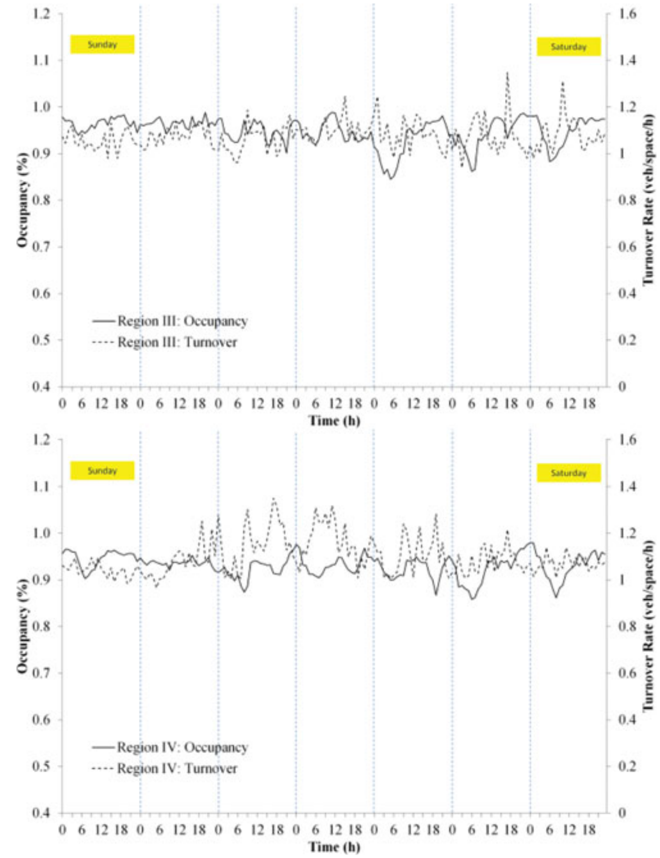

Figure 4 Hourly evolution of parking occupancy (%) and turnover rate (vehicles/space/h) for Regions III and IV for a typical week.

For efficiently analyzing the parking characteristics, the entire area covered by the parking sensors is further segmented into four regions; this segmentation is empirically done based on drivers' feasible cruising paths, while searching for available parking spaces. The specific regions are depicted in Figure 2. The four regions are of different shape. Region 1 controls an area of approximately 1,800 m^2, and Region 2 an area of 12,000 m^2. Regions 3 and 4 are 30,000 m^2 and 76,000 m^2, respectively. The proposed segmentation is coarse enough to reflect the process of search for free parking space in the study area and is dependent on the existing road network. Evidently, a more exhaustive segmentation may be proposed (e.g., more than four regions), which may result in more detailed information on the predictive parking occupancy patterns in the study area.

Figures 3 and 4 show the time series of parking occupancy (%) and the turnover rate (vehicles/space/h) per region for a typical week. Further statistical testing shows that there exist differences in the mean of occupancy and turnover rate between weekdays and weekends for all regions. Moreover, high variability in average parking occupancy for Region I in the morning and afternoon is observed; setting 85% as the critical occupancy, we may distinguish between high-occupancy and low-occupancy periods within a day. High-/low-occupancy periods for the other parking regions have an occupancy threshold of 90%.

The duration of occupied parking spaces for all regions is best described by a Weibull distribution; the survival curve has the form $S(t) = \exp(-\lambda t^p)$, with parameters λ and p to be estimated. Figures 5 and 6 show the survival probabilities for the duration of occupied parking spaces per region. Higher probabilities after 60 minutes on average are expected in Regions I and II when compared to those of Regions III and IV, where the probability of a space being occupied after 60 minutes drops below 0.5.

Parking Occupancy Prediction

A unique MLP for each parking region is developed in order to predict the overall occupancy (%) of parking spaces using past information on its evolution. The input space is introduced with information on occupancy with a look-back time window ranging from 5 to 10 minutes. For the prediction system to be useful, multiple steps ahead predictions are required. Here, the prediction horizon of the developed models is extended from 1 step to 30 steps ahead (i.e., from 1 to 30 minutes ahead).

Following previous research on short term prediction using neural networks in transportation problems (a detailed description is given in Vlahogianni et al., 2005), the neural networks are

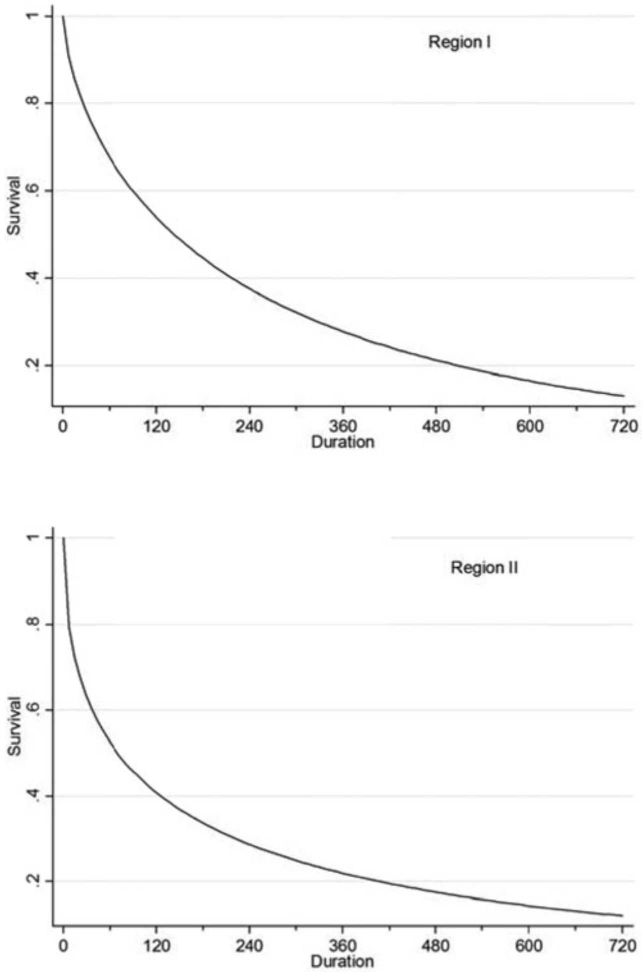

Figure 5 Survival probabilities for the duration of occupied parking spaces for Regions I and II.

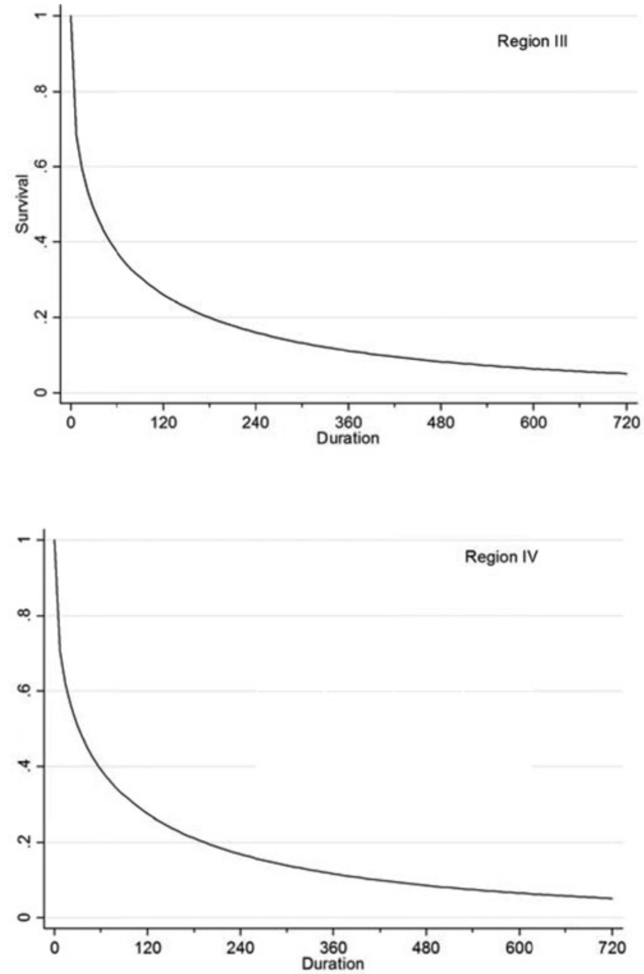

Figure 6 Survival probabilities for the duration of occupied parking spaces for Regions III and IV.

optimized with respect to their structural and learning properties using genetic algorithms. Training is conducted using a simple back-propagation algorithm with genetically optimized learning rate and momentum. The structure of the hidden layer (number of hidden units) is also optimized using genetic algorithms. The model's look-back time window is also genetically optimized. Based on the preceding description, a population of 100 generations is evaluated. Each generation involved the development and testing of 50 chromosomes that include the number of hidden units h (structure), the learning rate η, and the momentum μ (learning), as well as the time dependent inputs (parameters τ and m of Eq. 2). A roulette selection mechanism is employed. The crossover probability is set to 0.9 and the mutation probability to 0.02. The fitness function is the mean square error in the cross-validation set.

The available data are divided into the training (60%), cross-validation (20%), and test set (20%), used for training the models, evaluating the training of the models, and testing the generalization power of the models, respectively. The genetic optimization showed that an MLP of 8 hidden layers and a look-back time window of 5 minutes in the past may be efficiently used to predict parking occupancy (%) up to 30 steps in the future with high accuracy.

Prediction results (test set) are seen in Table 1 with respect to the following forecasting metrics:

- Mean absolute error (MAE): $\frac{\sum_i |\hat{y}_{i+\tau} - y_{i+\tau}|}{N}$
- Root mean squared error (RMSE): $\sqrt{\frac{\sum_i (\hat{y}_{i+\tau} - y_{i+\tau})^2}{N}}$
- Mean absolute percentage error (MAPE): $\frac{1}{N} \sum_i \frac{|\hat{y}_{i+\tau} - y_{i+\tau}|}{y_{i+\tau}}$
- Root relative squared error (RRSE): $\sqrt{\frac{\sum_i (\hat{y}_{i+\tau} - y_{i+\tau})^2}{N}} \Big/ \sqrt{\frac{\sum_i (\hat{y}_{i+\tau} - y_{i-\tau})^2}{N}}$

where $y_{i+\tau}$ and $\hat{y}_{i+\tau}$ are the actual and predicted value with $i = 1, .., N$, τ is the time step, N is the number of samples, and $y_{i-\tau}$ is the last known value relative to the prediction step. Results seem encouraging particularly for longer prediction horizons. For example, the models are able to predict each region's parking occupancy 15 minutes into the future with less than 3.6% MAPE.

Table 1 Results for different prediction horizons (prediction step τ equals to 1 minute) for each parking region (test set).

	Prediction horizon			
	1 min	5 min	15 min	30 min
Region 1				
MAE	0.007	0.023	0.030	0.032
RMSE	0.015	0.032	0.043	0.048
MAPE	0.869	2.713	3.564	3.913
RRSE	85.065	88.954	71.682	59.250
Region 2				
MAE	0.006	0.019	0.024	0.028
RMSE	0.011	0.024	0.031	0.036
MAPE	0.749	2.192	2.824	3.262
RRSE	87.223	92.559	80.614	74.776
Region 3				
MAE	0.004	0.012	0.014	0.015
RMSE	0.007	0.015	0.017	0.019
MAPE	0.412	1.213	1.415	1.576
RRSE	86.008	88.895	78.422	75.549
Region 4				
MAE	0.004	0.010	0.012	0.014
RMSE	0.006	0.012	0.015	0.017
MAPE	0.436	1.049	1.302	1.480
RRSE	86.245	89.879	79.914	77.151

Parking occupancy predictability is different between the four regions. Note that MLPs manage to produce predictions up to 30 minutes ahead with acceptable levels of accuracy. Although this may hint at a rather smooth temporal evolution of 1-minute parking occupancy data, the efficiency of these models may prove to be critical in delivering useful information services to travelers in urban areas.

Interestingly, all models perform well and produce predictions that are more accurate than a naive prediction (e.g., taking the last known value relative to the step of prediction), as seen from the RRSE. The larger the predictive horizon, the more accurate the MLP becomes relative to a naive prediction, an indication of the robustness of the NNs in dealing with problems of ranging levels of complexity. This is a strong indication that the proposed models may provide solid predictions that may enable potential predictive information dissemination in real time using Web-based tools, cell phone, or other media.

Survival Analysis of Free Parking Spaces

Hazard-based free parking space duration models are developed for each parking region. Based on preliminary analysis, a Weibull survival function is used. Several forms of Weibull regression models are tested, with independent variables being the TypeDay (weekend/weekday), the Period of day (high /low occupancy period) and the Weekday (1 to 7, from Monday to Sunday). Results are seen in Table 2.

As can be observed, different independent variables are significant in each parking region when modeling free parking space duration. In Region I, free space durations are dependent

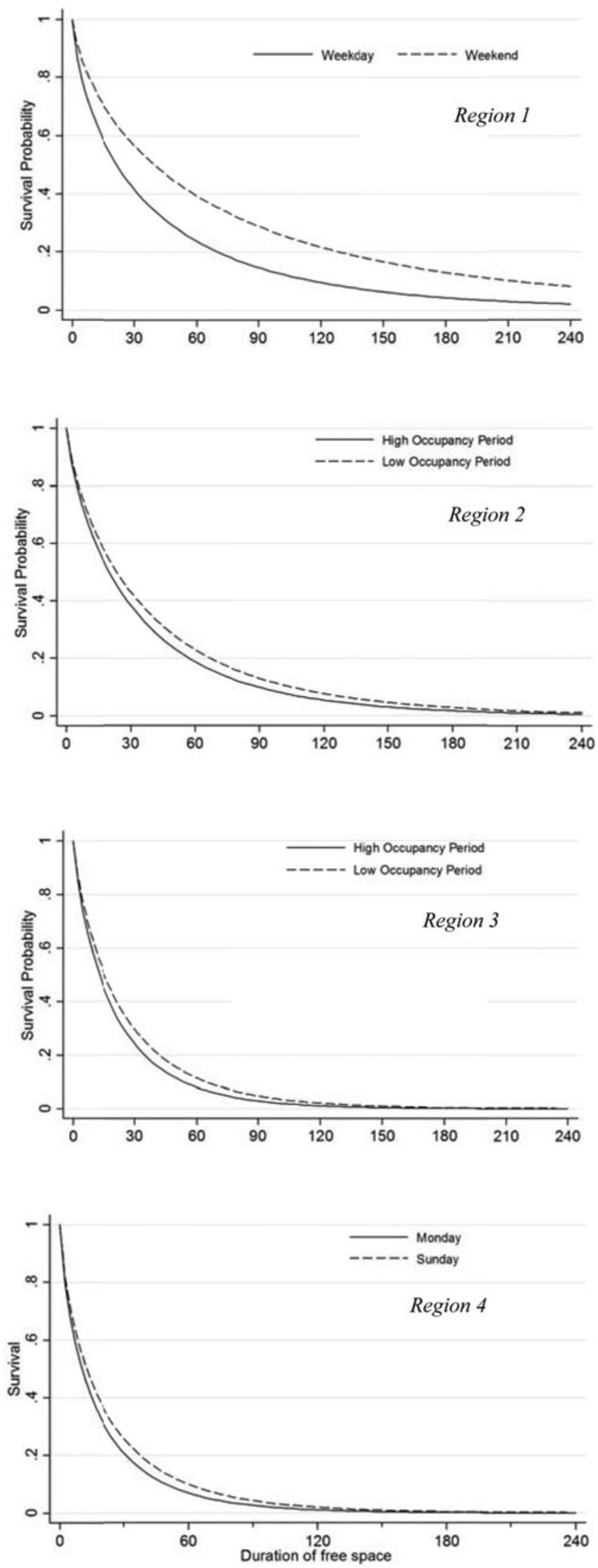

Figure 7 Survival curves based on the models developed for free space durations in the four regions under study.

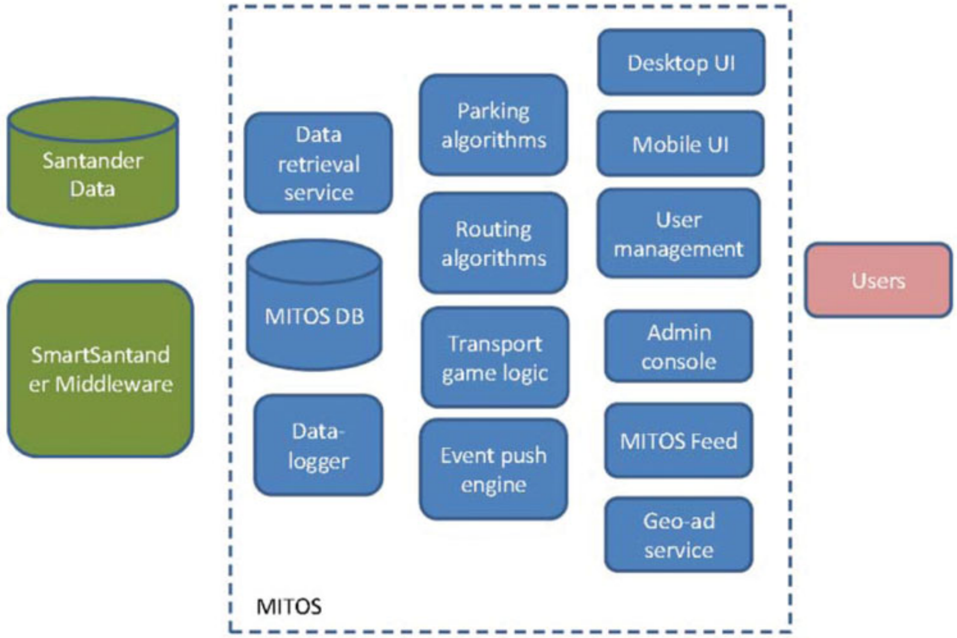

Figure 8 Architecture of MITOS platform.

on the type of day (weekend or weekdays), whereas in Region II and III, free space durations depend on the period of the day. Finally, in Region IV, free space durations change with regard to the daily evolution of parking demand. Interesting results arise from observing the survival probability curves (Figure 7) that show whether the phenomenon will continue to be observed in the forthcoming intervals, given that it has lasted up to time t. For example, in Region I there is a probability of a parking space continuing to be free after 5 minutes of 0.76 for weekdays; this increases to 0.83 for weekends. The same probability for Region II spaces is 0.77 for high-occupancy periods and increases to 0.79 in low-occupancy periods. In Region III the probability of a space to be free after 5 minutes is 0.69 and 0.73 for high- and low-occupancy periods, respectively. Finally, in Region IV the differences of the same probability are small but significant; the probability ranges from 0.63 in Mondays and 0.69 in Sundays.

The preceding survival curves may be used as a predictive tool for parking availability. Interestingly, the estimations provided by the survival model may act complementary to the parking occupancy predictions that require relatively extensive and continuous data streams. This may significantly improve the real-time operability of an intelligent parking prediction framework.

Table 2 Survival analysis results.

	β_i			
Variable	Region I	Region II	Region III	Region IV
Constant	−2.98	−2.90	−2.65	−2.06
TypeDay	0.43	—	—	—
Period	—	0.13	0.15	—
Weekday	—	—	—	−0.04
Log likelihood	770.86	−1764.64	−1932.55	−4367.82
LR chi2(1)	10.15	9.78	8.28	14.36
Prob > chi2	0.000	0.002	0.002	0.0002

REAL-WORLD APPLICATION

The models just described have been implemented in the MITOS system (multi-input transport planning system). MITOS is a system developed within the SmartSantander project (http://mitos.smartsantander.eu), in the city of Santander, Spain. The goald of MITOS is to deliver novel intelligent transportation services (ITS) to smart city citizens and capture perceived user experience. The fact that an IoT (Internet of Things) infrastructure is available provides new ways of quantifying the effect that ITS may have on the daily transportation of the commuters and the environment. The MITOS platform provides an integrated multi-modal transportation guide, which allows citizens and visitors to optimally choose their trips and get accurate information and guidance before and during their trip. The MITOS platform includes the following services:

- Stop/Line/POI Survey and Search for public transport.
- Route guide.
- Parking monitoring and short-term parking occupancy prediction.

The architecture of the MITOS platform is seen in Figure 8. The services and applications deployed and demonstrated in the context of MITOS include a Web portal for transportation information, which is used as a city guide for citizens and tourists, and

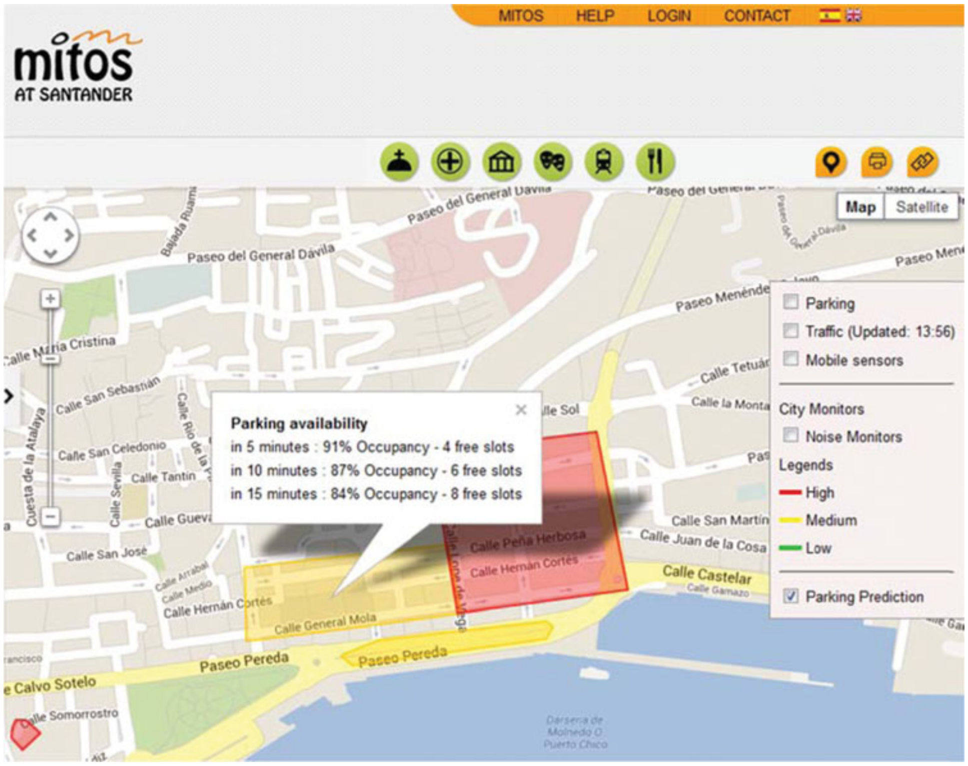

Figure 9 Screen shot of Web parking prediction implementation.

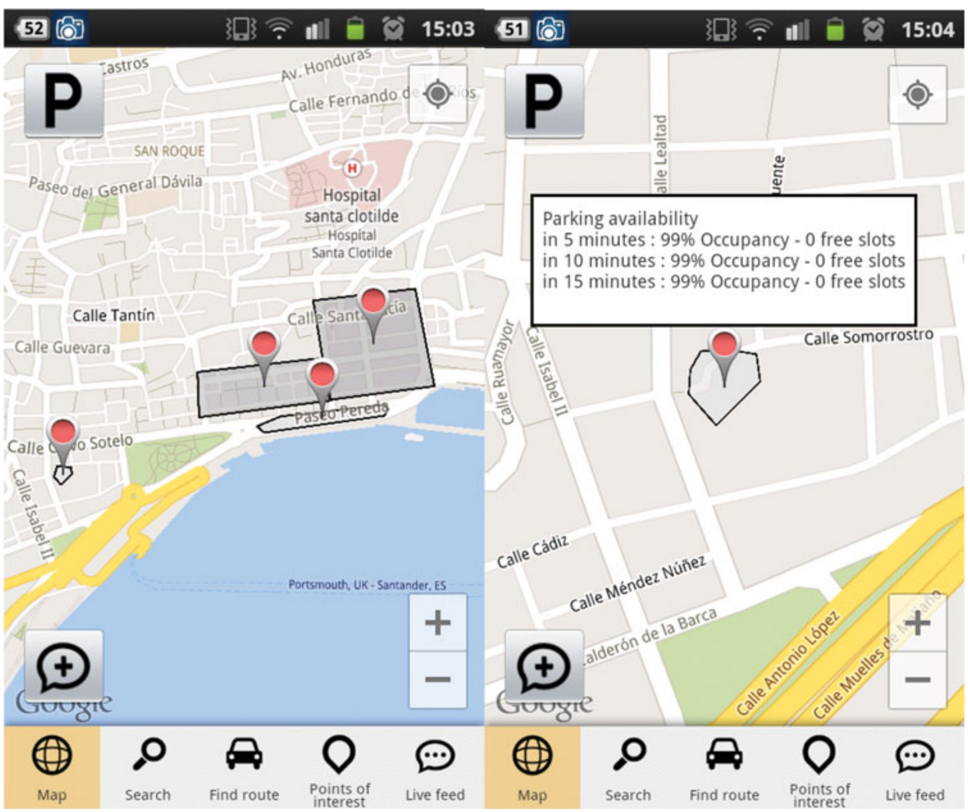

Figure 10 Screen shot of mobile parking prediction implementation.

a mobile application for advanced transportation services. The last provides all type of information required by commuters and enables intuitive ways for searching routes and geo-referenced information based on the personal preferences and abilities of each user.

An added value of the MITOS platform comes from the exploitation of various diverse sensor sources such as:

- Participatory sensors: Users will act as sensors that provide relevant traffic and travel information in the form of free text or predefined messages (e.g., "heavy traffic," "too much noise," etc.) and/or image.
- Environmental data sensors, measuring noise, temperature, and CO/CO_2 emissions.
- On-street parking space occupancy sensors.
- Traffic occupancy sensors.
- On-vehicle devices (GPS) installed in buses.

The implementation of a parking prediction service involves retrieving parking sensor data every minute (through the SmartSantander framework APIs) and storing them in a relational database (MITOS DB). Next, the algorithms described earlier are executed, and a continuous stream of predictions is generated. Those predicted values are stored back in the database, so that related user requests on parking availability can be served as rapidly as possible. Following user requests for parking occupancy, data are retrieved and visualized on a digital Web map (Figure 9) and on a mobile phone (Figure 10). The visualization involves data easily perceived by the end users, such as estimated occupancy and estimated number of free parking slots in the short-term future (5, 10 and 15 minutes ahead).

CONCLUSIONS

In this article we exploited high-resolution parking occupancy data and developed and tested a system for short-term and longer term parking availability prediction in urban areas. The system encompasses two modules; the first introduces neural networks for the prediction of the time series of parking occupancy in different regions of an urban network, while the second applies survival analysis to predict the probability that a parking space will be free in the following time intervals. Findings show that the duration of free parking space follows a Weibull distribution. Moreover, the neural networks adequately captured the temporal evolution of parking occupancy and may accurately predict occupancy up to half an hour ahead. The proposed model, as it stands, is flexible and of reduced complexity aiming to be implemented online. This is shown by its incorporation within an innovative routing service (MITOS) for Santander, Spain, using real-time data from the city's "smart" infrastructure.

From a methodological perspective, the proposed approach is tested on limited data that may not claim to be representative of the monthly variations in parking demand. The evolution of parking occupancy may significantly differ between summer and winter months and may have other cyclicity features that were not addressed. Thus, a more comprehensive data set of historical parking occupancy and duration information may improve forecasting. Toward the treatment of seasonality features, retraining strategies based on the daily or monthly seasonality should be also considered. The treatment of seasonal features may be done by introducing specific dummy variables to NNs, or by using more complex NN structures, such as the temporal NNs that include memory mechanisms. Nevertheless, these networks should be treated with caution due to the difficulties in training and convergence.

Moreover, a critical limitation of the present approach is the lack of traffic data that would have provided a more consistent formulation of the parking prediction problem to the evolution of traffic demand. Correlations with traffic's spatiotemporal demand are expected to improve the predictability of parking occupancy. Nevertheless, these recommended directions for improving forecasts may result to more complex and difficult to train networks. Thus, a thorough consideration should be given to issues related to computational efficiency during model development. Further, traffic spatiotemporal patterns should be also jointly considered with the extent and coverage of the free parking space prediction models. Evidently, area-wide and streetwide prediction models may serve different purposes in relation to the type and characteristics of the ITS application.

From a conceptual perspective, the proposed approach can provide the anticipated free parking slots, as well as the probability of having free parking spaces in the specific area. In the absence of real-time data, the second information set that is based on historical data may be found extremely useful. The proposed approach may be extended to include parking fares (where they apply), as well as underground parking stations near each zone to enhance the explanatory power of the model and improve predictions.

FUNDING

This work is part of the MITOS experiment, incorporated in the SmartSantander project and funded by the European Union Seventh Framework Programme (FP7).

REFERENCES

Arnott, R., & Rowse, J. (1991). Modeling parking. *Journal of Urban Economics*, **45**(1), 97–124. http://dx.doi.org/10.1006/juec.1998.2084

Atzori, L., Iera, A., & Morabito, F. (2010). The Internet of Things: A survey. *Computer Networks*, **54**(15), 2787–2805. http://dx.doi.org/10.1016/j.comnet.2010.05.010

Bhat, C. R. (2000). Duration modeling. In D. A. Hensher & K. J. Button (Eds.), *Handbook of transport modeling* (pp. 91–111). London, UK: Elsevier Science.

Bonsall, P. W. (1991). The changing face of parking-related data collection and analysis: The role of new technologies. *Transportation* **18**(1), 83–106. http://dx.doi.org/10.1007/BF00150560

Bonsall, P. W., & Palmer, I. (2004). Modelling drivers' car parking behaviour using data from a travel choice simulator, *Transportation Research Part C: Emerging Technologies*, **12**(5), 321–347. http://dx.doi.org/10.1016/j.trc.2004.07.013

Caicedo, F. (2009). The use of space availability information in "PARC" systems to reduce search times in parking facilities. *Transportation Research Part C: Emerging Technologies*, **17**(1), 6–68. http://dx.doi.org/10.1016/j.trc.2008.07.001

Caicedo, F., Blazquez, C., & Miranda, P. (2012). Prediction of parking space availability in real time. *Expert Systems with Applications*, **39**(8), 7281–7290. http://dx.doi.org/10.1016/j.eswa.2012.01.091

Caicedo, F., Robuste, F., & Lopez-Pita, A. (2006). Parking management and modeling of car park patron behavior in underground facilities. *Transportation Research Record*, **1956**, 60–67. http://dx.doi.org/10.3141/1956-08

Caragliu, A., Del Bo, C., & Nijkamp, P. (2009). *Smart cities in Europe. Series Research Memoranda 0048*. Amsterdam, The Netherlands: VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics.

Chen, Z., Xia, J., & Irawan, B. (2013). Development of fuzzy logic forecast models for location-based parking finding services. *Mathematical Problems in Engineering*, Article ID 473471.

Chung, Y. (2010). Development of an accident duration prediction model on the Korean freeway systems. *Accident Analysis and Prevention*, **42**(2), 282–289. http://dx.doi.org/10.1016/j.aap.2009.08.005

David, A., & Keller, H. (2001, September). Event-driven modelling of on-street parking probability. *Proceedings of the 8th World Congress on Intelligent Transport Systems,* Sydney, New South Wales, Australia.

David, A., Overkamp, K., & Scheuerer, W. (2000, November). Event-oriented forecast of the occupancy rate of parking spaces as part of a parking information service. *Proceedings of the 7th World Congress on Intelligent Systems,* Turin, Italy.

Fabusuyi, T., Hampshire, R. C., Hill, V., & Sasanuma, K. A. (2011, October). Predictive model and evaluation framework for smart parking: The case of ParkPGH. *Proceedings of the 18th ITS World Congress,* Orlando, FL.

Giuffrè, T., Siniscalchi, M. S., & Tesoriere, G. (2012). A novel architecture of parking management for smart cities, *Proceedia—Social and Behavioral Sciences*, **53**(3), 16–28. http://dx.doi.org/10.1016/j.sbspro.2012.09.856

Gutiérrez, V., Galache, J. A., Sánchez, L., Muñoz, L., Hernández, J. M., Fernandes, J., & Presser, M. (2013). SmartSantander: Internet of Things research and innovation through citizen participation, Future Internet Assembly—Dublin (Ireland). *Lecture Notes in Computer Science*, **7858**, 173–186. http://dx.doi.org/10.1007/978-3-642-38082-2_15

Habib, K. M. N. (2012). Modeling commuting mode choice jointly with work start time and work duration. *Transportation Research Part A: Policy and Practice*, **46**(1), 33–47.

Hasan, S., Mesa-Arango, R., & Ukkusuri, S. (2013). A random-parameter hazard-based model to understand household evacuation timing behavior. *Transportation Research Part C: Emerging Technologies*, **27**(1), 108–116. http://dx.doi.org/10.1016/j.trc.2011.06.005

Hensher, D. A., & King, J. (2001). Parking demand and responsiveness to supply, pricing and location in the Sydney central business district. *Transportation Research Part A: Policy and Practice*, **35**, 177–196.

Hensher, D. A., & Mannering, F. L. (1994). Hazard-based duration models and their application to transport analysis. *Transport Reviews*, **14**(1), 63–82. http://dx.doi.org/10.1080/01441649408716866

Ibeas, A., Cordera, R., dell'Olio, L., & Moura, J. L. (2011). Modelling demand in restricted parking zones. *Transportation Research Part A—Policy and Practice*, **45**(6), 485–498. http://dx.doi.org/10.1016/j.tra.2011.03.004

Karlaftis, M. G. (2011). Modeling transit vehicle repair duration and active service time. *Transportation Planning and Technology*, **34**(5), 433–442. http://dx.doi.org/10.1080/03081060.2011.586113

Karlaftis, M. G., & Vlahogianni, E. I. (2011). Statistics versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, **19**(3), 387–399. http://dx.doi.org/10.1016/j.trc.2010.10.004

Komninos, N. (2009). Intelligent cities: Towards interactive and global innovation environments. *International Journal of Innovation and Regional Development*, **1**(4), 337–355. http://dx.doi.org/10.1504/IJIRD.2009.022726

Liu, Q., Lu, H., Zou, B., & Li, Q. (2006, June). Design and development of parking guidance information systems based on Web and GIS technology. *Proceedings of the 6th International Conference on ITS Telecommunications,* pp. 1263–1266, Chengdu, China.

Mandic, D. P., & Chambers, J. A. (2001). *Recurrent neural networks for prediction: Learning algorithms, architectures, and stability. Wiley Series in Adaptive and Learning Systems for Signal Processing, Communications, and Control*. Hoboken, NJ: John Wiley & Sons.

Martens, K., & Benenson, I. (2008). Evaluating urban parking policies with agent-based model of driver parking behavior. *Transportation Research Record*, **2046**, 37–44. http://dx.doi.org/10.3141/2046-05

Oh, B., Lee, S., Kim, M., & Yang, Y. (2002, June). Spatial applications using 4S technology for mobile environment. *Proceedings of the Geoscience and Remote Sensing Symposium,* pp. 2444–2446, Toronto, Canada.

Rashidi, T. H., Mohammadian, A., & Koppelman, F. S. (2011). Modeling interdependencies between vehicle transaction, residential relocation and job change. *Transportation*, **38**(6), 909–932. http://dx.doi.org/10.1007/s11116-011-9359-4

Shin, K., & Ramanathan, P. (1994). Real-time computing: A new discipline of computer science and engineering. *Proceedings of the IEEE*, **82**, 6–24. http://dx.doi.org/10.1109/5.259423

Smith, B. L., & Oswald, R. K. (2003). Meeting real-time requirements with imprecise computations: A case study in traffic flow forecasting. *Computer Aided Civil and Infrastructure Engineering*, **18**(3), 201–213. http://dx.doi.org/10.1111/1467-8667.00310

Stathopoulos, A., & Karlaftis, M. G. (2002). Modeling duration of urban traffic congestion. *Journal Transportation Engineering*, **128**(6), 587–590. http://dx.doi.org/10.1061/(ASCE)0733-947X(2002)128:6(587)

Teodorovic, D., & Lucic, P. (2006). Intelligent parking systems. *European Journal of Operational Research*, **175**(3), 1666–1681. http://dx.doi.org/10.1016/j.ejor.2005.02.033

Thornton, D. A., Redmill, K, & Coifman, B. (2014). Automated parking surveys from a LIDAR equipped vehicle. *Trans-*

portation Research Part C: Emerging Technologies, **39**, 23–35. http://dx.doi.org/10.1016/j.trc.2013.11.014

Tiwari, G., Bangdiwala, S., Saraswat, A., & Gaurav, S. (2007). Survival analysis: Pedestrian risk exposure at signalized intersections. *Transportation Research Part F*, **10**(1), 77–89. http://dx.doi.org/10.1016/j.trf.2006.06.002

Van Der Goot, D. (1982). A model to describe the choice of parking places. *Transportation Research Part A: General*, **16**, 109–115. http://dx.doi.org/10.1016/0191-2607(82)90003-6

Vlacheas, P., Giaffreda, R., Stavroulaki, V., Kelaidonis, D., Foteinos, V., Poulios, G., ... Moessner, K. (2013). Enabling smart cities through a cognitive management framework for the internet of things. *IEEE Communications Magazine, IEEE*, **51**(6), 102–111. http://dx.doi.org/10.1109/MCOM.2013.6525602

Vlahogianni, E. I. (2009). Enhancing predictions in signalized arterials with information on short-term traffic flow dynamics. *Journal of Intelligent Transportation Systems*, **13**(2), 73–84. http://dx.doi.org/10.1080/15472450902858384

Vlahogianni, E. I. (2013). Modeling of the overtaking duration in two lane highways. *Transportation Research Part F: Traffic Psychology and Behavior*, **20**, 135–146. http://dx.doi.org/10.1016/j.trf.2013.07.003

Vlahogianni, E. I., & Karlaftis, M. G. (2013a). Fuzzy-entropy neural network freeway incident duration modeling with single and competing uncertainties. *Computer Aided Civil and Infrastructure Engineering*, **28**(6), 420–433. http://dx.doi.org/10.1111/mice.12010

Vlahogianni, E. I., & Karlaftis, M. G. (2013b). Testing and comparing neural network and statistical approaches for predicting transportation time series. *Transportation Research Record: Journal of the Transportation Research Board*, **2399**(1), 9–22. http://dx.doi.org/10.3141/2399-02

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach. *Transportation Research Part C: Emerging Technologies*, **13**(3), 211–234. http://dx.doi.org/10.1016/j.trc.2005.04.007

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, **43**, 3–19. http://dx.doi.org/10.1016/j.trc.2014.01.005

Vlahogianni, E. I., Karlaftis, M. G., Golias, J. C., & Kourbelis, N. D. (2006, September). Pattern-based short-term urban traffic predictor. *Proceedings of the 9th International IEEE Conference on Intelligent Transportation Systems,* pp. 389–393, Toronto, Canada.

Vlahogianni, E. I., Karlaftis, M. G., & Kepaptsoglou, K. (2011). Nonlinear autoregressive conditional duration models for traffic congestion estimation. *Journal of Probability and Statistics*, Article ID 798953.

Washington, S., Karlaftis, M. G., & Mannering, F. L. (2010). *Statistical and econometric methods for transportation data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC Press.

Wong, S. C., Tong, C. O., Lam, W. C. H., & Fung, R. Y. C. (2000). Development of parking demand models in Hong Kong. *Journal of Urban Planning and Development*, **126**, 55–74. http://dx.doi.org/10.1061/(ASCE)0733-9488(2000)126:2(55)

Young, W., Thompson, R. G., & Taylor, M. A. P. (1991). A review of urban parking models. *Transport Reviews*, **11**(1), 63–84. http://dx.doi.org/10.1080/01441649108716773

Zhong, M., & Hunt, J. D. (2010). Exploring best-fit hazard functions and lifetime regression models for urban weekend activities: Case study. *Journal of Transportation Engineering*, **136**(3), 255–266. http://dx.doi.org/10.1061/(ASCE)0733-947X(2010)136:3(255)