Improving Parking Availability Information Using Deep Learning Techniques

Jamie Arjona a,*, MªPaz Linares a, Josep Casanovas-Garcia a,b, Juan José Vázquez a

aUniversitat Politècnica de Catalunya, Barcelona 08034, Spain
bBarcelona Supercomputing Center, Barcelona 08034, Spain

Abstract

Urban traffic currently affects the quality of life in cities and metropolitan areas as the problem becomes ever more aggravated by parking issues: congestion increases due to individuals looking for slots to park their vehicles. An Internet of Things approach allows drivers to know the state of the parking system in real time through wireless networks of sensor devices. This work focuses on studying the data generated by parking systems in order to develop predictive models that generate forecasted information. This can be useful in improving the management of parking areas, especially on-street parking, while having an important effect on urban traffic. This research begins by looking at the state of the art in predictive methods based on machine learning for time series. Similar studies and proposed solutions for parking prediction are described in terms of the technology and current state-of-the-art predictive models. This paper then introduces the recurrent neural network methods that were used in this research, namely Long Short-Term Memory and Gated Recurrent Unit, as well as the models developed according to real scenarios in different cities. In order to improve the quality of the models, exogenous variables like hourly weather and calendar effects are taken into account, and the baseline models are compared to the models that used this information. Finally, the preliminary encouraging results are described, followed by suggestions for corresponding future work.

© 2019 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 22nd EURO Working Group on Transportation Meeting.

Keywords: parking availability forecast; deep learning; smart cities.

1. Introduction

Nowadays, one of the most important problems in urban areas concerns traffic congestion. This, in turn, has an impact on the economy, nature, human health, cities architecture, and many other facets of life. Part of the vehicular traffic in cities is caused by parking space availability. The drivers of private vehicles usually want to leave their

* Corresponding author. Tel.: +34934011019
E-mail address: jamie.arjona@upc.edu

2352-1465 © 2019 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 22nd EURO Working Group on Transportation Meeting.
vehicle as close as possible to their destination. However, the parking slots are limited and may not be enough to sustain the demand, especially when the destination pertains to an attractive area. Thus, individuals looking for a place to park their vehicle contribute to increasing traffic flow density on roads where the parking demand cannot be satisfied.

This motivates the use of real-time parking availability in order to build a mathematical model that helps improve parking management. Some approaches to this can be found in Caicedo et al. (2012) and in Teodorovic and Lucic (2006), who use models to manage indoor parking reservations while taking into account future parking demand. However, the nature of the data remains a problem: It is not on-street parking data and also is not in real time.

An Internet of Things (IoT) approach allows us to know the state of the parking system (availability of the parking slots) in real time through wireless networks of sensor devices. An intelligent treatment of this data could generate forecasted information that may be useful in improving management of on-street parking, thus having a notable effect on urban traffic. Smart parking systems first appeared in 2015, with platforms in Santander, San Francisco and Melbourne as is explained in Lin et al. (2017), when those cities began to provide on-street real-time parking data for offering new services to their citizens.

One of the most interesting services is parking availability forecasting, for which the first works studied the temporal and spatial correlations of parking occupancy. For this purpose, a VARIMA model was proposed by Rajabioun and Ioannou (2015) for short-term forecasts (less than 30 minutes) without loss of accuracy.

Those short-term forecasts are not useful at all to the end user of this service; thus, prediction intervals on the order of hours are needed. This is similar to what happens in traffic forecasting, where the literature (Vlahogianni and Karlaftis (2013) has identified a potential class of models that can offer better predictions than ARIMA: Neural Networks (NN).

The purpose of this paper is to present models that accurately forecast parking availability in urban areas in order to create and improve services for parking management. These models are developed for real scenarios using data provided by sensors with certain inaccuracies due to their very nature. Furthermore, the developed models are based on Recurrent Neural Networks (RNN), as these are capable of learning the time-dependency component that occurs in parking availability data. Lastly, the forecast will be of a mid-term duration (an interval of time between 2 and 8 hours).

This paper is organized as follows. The next section presents a summarized literature review of parking availability forecasting methods and further indicates new findings in contrast to the existing literature. Then we present the proposed methods, their development and the validation tests performed. Finally, the last section discusses the empirical findings from the implemented scenarios, followed by some concluding remarks.

2. State of the Art

In recent years, the research on parking forecasting has grown more sophisticated due to the availability of real-time data on parking slots. This data has become available as a result of advances in the field of sensors and the so-called IoT paradigm. The usual methods for obtaining parking data are: sensors, VANET (vehicle ad-hoc network), and GPS for detecting the number of vehicles in a zone’s cruising traffic. The variables of interest are availability, reservation (Caicedo et al. (2012)) and occupancy (usually in the form of parking occupancy rate (OCCR) (Vlahogianni et al. (2014)).

Prediction for parking analytics has always been a driving force behind traffic predictive analytics, which has been an active area of research for many years. This in turn has motivated efforts to adapt promising traffic methods to parking prediction. Thus, this section considers the scope of both in the literature review.

One of the first works that used an NN model is Ishak et al. (2003), where upstream and downstream data from a target link is used for traffic flow prediction. In Vlahogianni et al. (2005), a Multilayer Perceptron approach with a Genetic Algorithm is used for meta-optimization (or hyper-parameter tuning) in the short-term forecasting of traffic flow. At the same time, Medeiros et al. (2006) presented a work that they called AR-NN, where an NN is modeled using a statistical approach like the ones used in time series.

Currently, the approaches are centered on Deep Learning (large NN models with more than one hidden layer). In Yanjie Duan et al. (2016), Long Short-Term Memory NN (LSTM) (Hochreiter and Schmidhuber (1997)) are used for training 66 models (one per link) in order to perform travel flow forecasting. Fu et al. (2017) presented a comparison
between ARIMA, LSTM and Gated Recurrent Unit (GRU) (Cho et al. (2014)) for travel flow. Another approach with a Deep NN is presented in Polson and Sokolov (2017), where the model is used for traffic flow forecasting.

As mentioned in the Introduction, the first works on parking prediction emerged in parallel with big IoT platforms that collected data in real time from parking spots. Vlahogianni et al. (2016) takes an NN approach to predict parking occupancy times. Blythe et al. (2015) propose transforming parking space availability time series in order to reconstruct them in phase space, and then use the transformed data to train what they call a Wavelet NN.

Naïve Bayes and a decision tree are used in Hsueh et al. (2018) for predicting parking occupancy as a categorical variable (level: low, mid, high) at arrival time. The relevance of the study is that it uses three sets of different types of features: temporal (day of the week, time of day (ToD) and occupancy rate); spatial (nearest parking sectors, similar parking sectors based on RMSE, similar parking sectors based on RMSE with different ToD); and others (sector id and hourly precipitation).

Another approach to the parking prediction problem is detecting traffic that results from cruising. In Jones et al. (2018), data from smartphones (GPS, accelerometer, magnetometer, and gyroscope) are used to detect cruising behavior. The detection of cruising can be useful for updating probabilistic parking availability in the desired zone. Using GPS data, map matching is carried out via a Hidden Markov Model to localize the zones where the vehicle is located. Cruising is detected as a significant local minimum in the GPS trace relative to the distance to the destination. Data from the other sensors are used for training the ML algorithms. The ML algorithms that ascertain whether or not the user is cruising are Decision Tree, K-NN, and SVM. One of the problems with this proposal is that accuracy depends on the number of users in the zone.

In this work, our objective is to take an RNN approach to forecast the occupancy of street parking sectors over a succeeding period of 6 hours. We used LSTM and GRU models, as they have been proven to be the best RNN architectures for solving a wide set of sequential data problems (Jozefowicz and Zaremba (2015)). As far as we know, this approach has been used recently with good results in traffic forecasting but not for parking. In addition, these kinds of models allow improving predictions via the use of exogenous data such as weather conditions and calendar effects (holidays, day of the week, month, etc.). The reason this is interesting is that sensor measurements are subject to environmental noise that can affect the quality of the data used for training the models. Thus, relying only on measurements cannot be enough.

3. Developed Models

RNNs are part of the NN framework and are specialized for working with sequential data. The computational units are connected to themselves in order to preserve past information that can be relevant in the present computation. An RNN is composed of a weight matrix \( W_e \) that is applied to the input, a hidden parametric state transition \( S \), a weight matrix \( W_y \) that is applied to the output and an output bias \( b_y \). The input consists of a sequence of \( T \) elements \( x = (x_1, ..., x_T) \). At each step \( t \), the RNN unit computes a hidden state \( s_t = S(s_{t-1}, W_e x_t) \) that is part of the hidden sequence \( s = (s_1, ..., s_T) \), and an output \( y_t = W_y s_t + b_y \) that is part of the output sequence \( y = (y_1, ..., y_T) \).

There exist different implementations of RNN and, as mentioned in the State of the Art section, two of them are well-known in the literature, namely LSTM and GRU. These architectures make use of auxiliary mechanisms called gates, which enhance the computational units in order to manage past information better.

In particular, LSTM stores memory in the form of an internal state similarly to RNN, but it also implements a gating system that modulates the flow of information within the unit in order to handle the interaction between the state and the current sample. A fundamental aspect in LSTMs is the cell state, where the information is added and removed carefully by a structure of gates that are composed of a sigmoid function and a pointwise multiplication operation. An LSTM has four of these gates, which control the hidden state of the LSTM neuron in the following manner. First, the forget gate is computed in order to ascertain the relationship between the current input and the hidden state. Then, the update gate is computed in order to decide if the input is relevant or not. With the results, the internal state gate is computed on the basis of how much relevant previous information and new information is retained. Finally, the output of the neuron is computed by taking into account the new internal state.

The difference between LSTM and GRU basically is that a single gate controls the forgetting and the decision to update the state, thus requiring less computation. It is interesting to note that GRU models have not been used before
in the literature for parking occupancy forecasting. In addition, these NN permit multiple output predictions that behave like different steps leading up to forecasts.

Intuitively, these models seem to be perfect candidates for learning temporal behaviors about time series data. Not only can those patterns be learnt using linear techniques (temporal correlations) and modeled in ARIMA, but also non-linear patterns (like trends) that may be present in the data and not be perceived because their granularity (hours) are hidden within larger scales (months or years). Another advantage of these models is the easy integration of other information sources (exogenous variables) that can improve accuracy.

4. Data

This work takes advantage of the availability of real-time parking data obtained from different pilot programs carried out as part of the Horizon 2020 project titled Fastprk2: Enhanced on-street parking management system (https://cordis.europa.eu/project/rcn/204264/factsheet/en). The objective of this project is to develop better parking detection systems and create intelligent transport services for cities and citizens. One of the participants, Worldsensing S. L., provided a new generation of sensors that combine both magnetic and infrared detection technologies that are nearly 99% accurate, thanks to better management of environmental noise that can affect magnetic sensors. This opens up the possibility of offering more reliable services.

Sensors are buried in the road within the area delimited by the parking spot and they are event-based. When they detect a car passing over them (change in sensor measurements), the data is then sent to a gateway (groups of sensors communicate to a single gateway), and from there to cloud servers.

4.1. Parking Data

The input data that Worldsensing provides can be handled at different scales: parking spot (sensor level) and sector (spatial aggregation of sensors pre-defined by the provider). The following specifies each of these measures:

- Sensor measurements are event-based and indicate when a parking spot is occupied and when it is free. They are indexed by timestamp and contain information on each sensor event at the time that it was generated. This data is collected from the gateways.
- Sector measurements contain information related to the sector status at each hour of the day. Sector data is computed by an extract-transform-load (ETL) process that aggregates the status of the sensors in each sector. This is the data that has been used for this study.

The developed forecasting models predict the occupancy/hour for a sector. This is computed by aggregating the occupancy/hour for all the sensors. Occupancy/hour is a measurement between 0 and 1, where 0 refers to an empty sector for an hour and 1 refers to full occupancy.

In order to perform the forecasts, the input data must be processed to feed the developed models. In particular, we take a time-window approach for RNN (Gers et al. (2001)) in which each model is fed with a look-back vector of 24 occupancy/hour values and it outputs a vector of 6 occupancy/hours (the forecasted values corresponding to the occupancy for the next 6 hours). The decision of using 24 hours of past data is based on empirical findings in which different intervals were used to train models and measure errors in the test set (see Fig. 1a).

4.2. Exogenous data

Because parking data may not be sufficient due to parking availability being affected by other sources of information, this study uses calendar effects and weather.

- Calendar effects: Parking behavior can be affected by so-called calendar effects. Occupancy on a weekend is different than on weekdays, but the pattern is maintained over weekends. In this study, the following calendar information is considered: day of the week, season, weekday and holiday (global and local). This data was collected from https://calendarific.com/api/v2/
4.3. Training Data

The proposed methodology begins by studying the amount of historical data needed for training models that make better predictions than random guesses. This quantification uses more than one year of historical data from different parking sensors deployed in Sandvika (Norway), Castelfranco Veneto and San Doná di Piave (Italy). These studies show that seven months’ worth of data is enough to enable RNN models to learn the system of patterns within parking sectors. As more data becomes available, the forecasts are improved (see Fig. 1b). Thus, the proposed work uses three months of training data for the different city parking sites in order to compare the two previously selected RNN methods: LSTM and GRU. An extra month of data is used to test the accuracy.

In each of the cities, there exists a different number of sectors. This study presents the results for the sectors where more sensors are involved.

5. Computational Experiments

5.1. Scenarios

The data used to train the models is collected from parking sectors in four different cities: Antwerp (Belgium), Barcelona (Spain), Wattens (Austria) and Los Angeles (USA). It also applies different characteristics like indoor/outdoor, public/private and environmental noise levels. In addition, each of the parking sites can include one or more sectors, which also have different characteristics between them. Table 1 characterizes the different parking zones in each of the cities.

- **Weather**: The objective of using this source of information is to find if the parking occupancy patterns are modified by the weather. The information considered related to weather is temperature, wind speed, precipitation intensity and weather summary for each hour. This data was gathered from https://api.darksky.net/forecast/.

It is worth mentioning that some of this data is categorical, so some preprocessing is needed. In order to learn how occupancy and these variables are related, an embedding model for each categorical value is considered. This is of special interest for the summary weather information (a short text about the climatology for a specific hour), as it allows modeling the meaning of the words into vectors. A similar idea was pursued by Guo and Berkahn (2016).
This heterogeneity makes it impossible to work with a single model, so a model was developed for each sector in each city. It is also important to note here that these parking sectors were deployed in the year 2018, therefore the historical data available covers April 1st to November 30th 2018. November data will be held as the test set of data in order to evaluate the models with unseen data.

5.2. Hyperparameter optimization

Neural networks are famously difficult to tune, due to the great number of hyperparameters that can be configured before finding a model that provides better forecasting. The hyperparameters that affect the architecture of the model are: number of layers, neurons per layer, activation function (linear, ReLu, SeLu, sigmoid, hyperbolic tangent), dropout (regularization), and model type (LSTM / GRU). Other hyperparameters were set to the default values indicated by the literature, namely: Adam, for optimization with a learning rate of 0.001 (Kingma and Lei Ba (2014)); Glorot, as a method for initializing the weights (Glorot and Bengio (2010)); and a batch size of 256, because values in multiples of 2 are common in the literature and its upper limit is based on the hardware used by some works on the effects of batch normalization (Masters and Luschi (2018)). In order to find the hyperparameter configuration for each model sector, we used a random search strategy (Bergstra and Bengio (2012)). An early stopping criterion proposed in Caruana and Lawrence (2000) is employed in the hyperparameter search in order to quickly discard models with poor initial results while training and to save computation time.

5.3. Results and discussion

Models are evaluated and compared using the Root Mean Square Error metric (RMSE), which is a commonly used metric for model goodness of fit and comparison. Other evaluation metrics of interest are model training time and complexity of the model (number of parameters, number of layers and neurons per layer).

Table 2 shows the comparison of the best LSTM and GRU endogenous models, which consider only the occupancy of the sector with the most sensors in each of the cities. For these models that use only endogenous variables, the GRU architecture achieves better results in nearly all cases compared to the LSTM proposal. It is interesting that the best LSTM models do use linear activation functions, thus indicating that the parking process could be modeled with a pure autoregressive model.

Table 3 shows the results of the exogenous models. Only the best architecture is shown. The exogenous variables used reduce the error for the Barcelona and Antwerp sectors, of which the latter shows a significant improvement. For the Wattens and Los Angeles sectors, the results are similar to those obtained with the endogenous model. Thus, the results shown that exogenous variables related to calendar and weather do improve the predictions of parking sectors with less number of sensors.

<table>
<thead>
<tr>
<th>City</th>
<th>No. of sensors</th>
<th>Indoor?</th>
<th>Private?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antwerp</td>
<td>9</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Barcelona</td>
<td>11</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Wattens</td>
<td>23</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>60</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 2. Results of endogenous models.

<table>
<thead>
<tr>
<th>City</th>
<th>Accuracy (RMSE)</th>
<th>Training Time (s)</th>
<th>No. of parameters</th>
<th>No. of layers</th>
<th>Neurons per layer</th>
<th>Activation function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antwerp</td>
<td>0.235 / 0.235</td>
<td>205 / 1185</td>
<td>1.126.782 / 9.123.420</td>
<td>2 / 5</td>
<td>353 / 503</td>
<td>ReLu / Linear</td>
</tr>
<tr>
<td>Barcelona</td>
<td>0.121 / 0.126</td>
<td>159 / 305</td>
<td>16.854.027 / 7.382.076</td>
<td>7 / 4</td>
<td>657 / 513</td>
<td>SeLu / Linear</td>
</tr>
<tr>
<td>Wattens</td>
<td>0.092 / 0.092</td>
<td>216 / 194</td>
<td>2.688.126 / 1.551.802</td>
<td>6 / 3</td>
<td>285 / 278</td>
<td>SeLu / Linear</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>0.114 / 0.116</td>
<td>569 / 1311</td>
<td>14.742.162 / 21.645.484</td>
<td>8 / 6</td>
<td>572 / 701</td>
<td>Linear / Linear</td>
</tr>
</tbody>
</table>
Table 3. Results of exogenous models.

<table>
<thead>
<tr>
<th>GRU / LSTM</th>
<th>Accuracy (RMSE)</th>
<th>Training Time (s)</th>
<th>No. of parameters</th>
<th>No. of layers</th>
<th>Neurons per layer</th>
<th>Activation function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antwerp</td>
<td>0.089 / 0.093</td>
<td>145 / 62</td>
<td>84.455.04 / 3.454.680</td>
<td>2 / 1</td>
<td>958 / 904</td>
<td>ReLu / SeLu</td>
</tr>
<tr>
<td>Barcelona</td>
<td>0.094 / 0.093</td>
<td>169 / 119</td>
<td>10.049.466 / 3.454.680</td>
<td>3 / 1</td>
<td>811 / 904</td>
<td>SeLu / ReLu</td>
</tr>
<tr>
<td>Wattens</td>
<td>0.089 / 0.092</td>
<td>254 / 48</td>
<td>14.950.594 / 3.454.690</td>
<td>3 / 1</td>
<td>992 / 904</td>
<td>SeLu / Linear</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>0.120 / 0.123</td>
<td>142 / 123</td>
<td>510.764 / 158.684</td>
<td>1 / 1</td>
<td>339 / 16</td>
<td>SeLu / SeLu</td>
</tr>
</tbody>
</table>

6. Conclusions and Future Research

In this paper, two RNN architectures (LSTM and GRU) are developed for accurately forecasting parking availability in urban areas. A further development has been carried out by including exogenous variables like hourly weather and calendar effects. Furthermore, the different proposals were intensively compared in order to analyze the suitability of whether or not to add these exogenous variables.

The results show that the GRU architecture achieves better results in nearly all cases compared to the LSTM version. While the proposed exogenous variables reduced the errors in the sectors with less number of sensors, the accuracy remains similar in bigger scenarios. Thus, the results show that weather and calendar effects improve the predictions of parking sectors with a small number of sensors (spots).

Finally, we have demonstrated that the amount of available historical data for training models directly affects forecasting capabilities. However, in new deployments, historical data is scarce. Thus, it may be valuable for future research to identify parking sectors with similar behaviors and transfer the knowledge of one sector model to a new sector that is anticipated to behave similarly. Exogenous variables like weather or calendar effects play a role in parking occupancy patterns, but it is necessary to be careful with the parking characteristics in order to choose the correct variables and improve the results. Other issues remain regarding the impact of environmental noise on sensor measurements. Lastly, further research could consider new exogenous variables that can affect a sector’s parking availability in the model. For example, it would be of interest to know how parking availability is affected by city events or the traffic flow in nearby parking sectors, among other factors.

Acknowledgements

Throughout this work, the authors have benefited from the support of the inLab FIB team at Universitat Politècnica de Catalunya and the company Worldsensing S.L. This research was funded by Secretaria d’Universitats i Recerca de la Generalitat de Catalunya (2017- SGR-1749) and under the Industrial Doctorate Program (2016-DI-79).

References


