2014 IEEE 17th International Conference on
Intelligent Transportation Systems (ITSC)
October 8-11, 2014. Qingdao, China

Using Exit Time Predictions to Optimize Self Automated Parking Lots

Rafael Nunes, Luis Moreira-Matias and Michel Ferreira

Abstract— **Private car commuting is heavily dependent on the subsidisation that exists in the form of available free parking. However, the public funding policy of such free parking has been changing over the last years, with a substantial increase of meter-charged parking areas in many cities. To help to increase the sustainability of car transportation, a novel concept of a self-automated parking lot has been recently proposed, which leverages on a collaborative mobility of parked cars to achieve the goal of parking twice as many cars in the same area, as compared to a conventional parking lot. This concept, known as self-automated parking lots, can be improved if a reasonable prediction of the exit time of each car that enters the parking lot is used to try to optimize its initial placement, in order to reduce the mobility necessary to extract blocked cars. In this paper we show that the exit time prediction can be done with a relatively small error, and that this prediction can be used to reduce the collaborative mobility in a self-automated parking lot.**

I. INTRODUCTION

Parking is a major problem of car transportation, with important implications in traffic congestion and urban landscape. It has been shown that parking represents 75% of the variable costs of automobile commuting [1], supported by a major public subsidisation of the space devoted to car parking, where the user does not pay in more than 95% of the occasions [2].

The sustainability of car transportation is nowadays facing several challenges. The number of cars in many cities has reached a level where the road infrastructure is unable to avoid systematic traffic congestions. In addition, the high cost of fossil fuels and pollutant emission levels are creating significant challenges for the sustainability of private car commuting in major cities. Tolls and prohibition of circulation in one or two week days for a given vehicle are already in place in some of our cities. Technology is trying to mitigate these challenges faced by car transportation. Zero-emissions electric propulsion and connected navigation are two examples of technologies that can help making car transportation more sustainable.

Technology has been focusing however in moving cars, disregarding the parked period of these cars, which represents 95% of the vehicle existence. Recently, a simple proposal that leverages on technology such as electric propulsion or wireless vehicular connectivity has addressed the issue of car parking, arguing that through a collaborative approach to the parking of cars, the area per car could be reduced to nearly half, when compared to the area per car in a conventional parking lot. This approach, known as *self-automated parking lots* [3], works as follows. An electric vehicle (EV) is left at the entrance of a parking lot by its driver. This EV is equipped with vehicular communications that establish a protocol with a Parking Lot Controller (PLC). The EV is also based on Drive-by-Wire (DbW) technology, where in-vehicle Electronic Control Units (ECUs) manage signals sent by the acceleration and braking pedal, and steering wheel. The Vehicle-to-Infrastructure (V2I) communication protocol allows the PLC to control the mobility of the EV in the parking lot. The PLC remotely drives the EV to its parking space, using in-vehicle positioning sensors (e.g. rotation per wheel), magnet-based positioning, or some other type of positioning system (e.g. camera-based). Alternatively to a fully-automated system, a scenario of human-based tele-operated driving could also be used [4]. In this concept of self-automated parking lots the cars are parked in a very compact way, without space devoted to access ways or even inter-vehicle space that allows opening doors. As a new vehicle enters the parking lot, the PLC sends wireless messages to move the vehicles in the parking lot to create space to accommodate the entering vehicle. If a blocked car wants to leave the parking lot, the PLC also sends messages to move the other vehicles, in order to create an exit path. In [3] it was shown that this concept could reduce the area per vehicle to nearly half, as well as reduce the overall mobility of cars in the parking lot, when compared to a conventional parking lot. However, in the original paper, a first-fit strategy was used to initially park each vehicle. Clearly, the initial placement can be improved if some knowledge about the expected exit time of each car is used. The basic idea is that a car should not be blocked by another car that will leave the parking lot later. If the cars in the parking lot are placed using an order that reflects their expected exit times, then the overall mobility in the parking lot to create exit paths can be reduced.

In this paper we use an entire year of entries and exits in a parking lot, where each vehicle uses a unique identifier, to be able to derive its expected exit time, using this information to improve the original placement of the car in order to reduce manoeuvring mobility. Our goal is not to obtain a precise exit time for each vehicle, but rather a time-interval that can be

Manuscript submitted June 22, 2014.
Manuscript reviewed August 19, 2014.

Rafael Nunes is with the Faculdade de Engenharia, U.Porto, 4200-465 Porto - Portugal (e-mail: rafael.nunes[at]dcc.fc.up.pt).
Luis Moreira-Matias is with the Instituto de Telecomunicações, 4200-465 Porto, Portugal and with Faculdade de Engenharia, U.Porto, 4200-465 Porto - Portugal (phone: 00351-91-4221647; e-mail: luis.matias[at]fe.up.pt).
Michel Ferreira is with the Instituto de Telecomunicações, U.Porto, 4169-007 Porto - Portugal (e-mail: michel[at]dcc.fc.up.pt).

This work was supported by the project I-CITY - "ICT for Future Mobility", as part of Grant NORTE-07-0124-FEDER-000064.

978-1-4799-6077-4/14/$31.00 ©2014 IEEE

used in conjunction with the parking lot layout (e.g. number of lanes) to reduce the probability of having to move parked vehicles to created exit paths for blocked vehicles.

The remainder of this paper is organised as follows: in the next section we present some considerations regarding parking lot design, and further describe our optimisation goal based on a typical layout for a self-automated parking lot. We then present our methodology to predict an exit time interval for each vehicle, and how this interval is used to select the original lane to park each vehicle. We then present our dataset set used as case study and present experimental results in the next section, including a discussion of these results. Finally, we end with some conclusions.

II. Parking Lot Design

The geometric design of the parking lot is an important issue in a self-automated parking lot. In conventional parking lots there are a number of considerations that have to be taken into account when designing them. For instance, width of parking spaces and access ways, one-way or two-way use of the access ways, entry angle in the parking bays ($90°, 60°, 45°$), pedestrian paths, visibility to find an available parking space, etc. In a self-automated parking lot, many of these considerations do not apply. Manoeuvring is done autonomously by the car following the instructions of the PLC, pedestrian access is not allowed, and the assigned parking space is determined by the PLC. The main design issue is defining a geometric layout that maximises parking space, leveraging on minimal buffer areas to make the necessary manoeuvres that allow the exit from any parking space under all occupancy configurations. This geometric design is ultimately determined by the shape of the space of the parking lot. The parking lot architecture also defines the trajectories and associated manoeuvres to enter and exit each parking space.

The parking lot has a V2I communication device which allows the communication between the vehicles and the PLC. In theory, this infrastructure equipment could be replaced by a vehicle in the parking lot, which could assume the function of PLC while parked there, handing over this function to another car upon exit, similarly to the envisioned functioning of a V2V Virtual Traffic Light protocol [5]. Note, however, that the existence of the actual infrastructure, which could be complemented with a video-camera offering an aerial perspective of the parking lot to improve the controller perception of the location and orientation of vehicles, could simplify the protocol and improve reliability.

Reducing and simplifying such trajectories and manoeuvres is also an important design issue, as they affect the reliability of the system and allow faster storage and retrieval of cars. Note also that the parking lot architecture can take advantage of the fact that the passenger does not enter the parking lot, and thus the inter-vehicle distances do not need to allow for space to open doors. To optimise and simplify manoeuvres, these self-automated parking lots will require specific minimum turning radius values for vehicles.

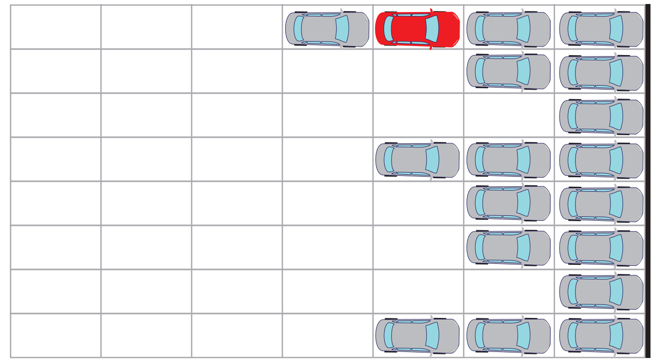

Fig. 1. An example layout for a self-automated parking lot. The parking lot can never be completely full, as buffer areas are necessary to be able to allow the exit of each vehicle under all possible configurations. In this example, a minimum of 6 empty spar are necessary.

Only vehicles that meet the turning radius specified by each parking lot will be allowed to enter it.

The geometric layout of the parking lot and its buffer areas can assume very different configurations for the self-automated functioning. One possibility is to have parallel lanes with minimal space between them, as illustrated in Fig. 1. In this type of layout, the PLC starts by assigning a lane to a vehicle. This initial decision is critical, as it should minimise the need to move a vehicle from one lane to another. Note that if the red vehicle in Fig. 1 needs to leave under the current configuration, then the vehicle behind it needs to be moved to another lane. If we could predict that the exit of the red vehicle would happen before the exit of the vehicle behind it, then this last vehicle would be better placed in a different lane. Our goal in this paper is exactly to be able to predict an exit-interval for each vehicle, and design a lane selection methodology that reduces the mobility needed to create exit paths.

Note that parking lots will not be able to be completely full, as buffer space needs to exist to allow the exit of each vehicle under all possible configurations. The minimum number of empty spaces, configuring buffer areas, depends on the parking lot layout. In the layout presented in Fig. 1, with a lane depth of 7, we need a buffer area with a minimum of 6 empty spaces.

III. Methodology

Our methodology consists on the following four steps: it starts by (A) dividing the original dataset in k smaller ones, containing users with similar parking habits; then, (B) data driven regression is performed over the newly created sub-datasets. Thirdly, a parking time interval is generated (C) based on such predictions and on their previous residuals (difference between a predicted value ($\hat{y}$) and its real one, y). Finally the selected lane (D) will be the one which minimizes the likelihood of performing *unnecessary* vehicle movements

[1]. This methodology is summarized in Fig. 2 and explained in detail throughout this section.

A. Profile Generation

Let $\mathbb{X} = \{X_1, X_2, ..., X_n\}$ be n timestamped data records on the parking lot entries describing the entry/exit behaviours of ρ distinct users. Let $U_i \subseteq \mathbb{X}$ denote the records of and individual user i (i.e. $U_{i=1}^{\rho} \equiv \mathbb{X}$) and Ψ_i describe the sample-based probability density function (*p.d.f.*) of its parking time habits. A clustering process is firstly made on $\mathbb{X}$ based on the extracted Ψ_i. The resulting k clusters can be defined as $\Pi = \{\pi_1, \pi_2, ..., \pi_k\}$. They will comprise sub-datasets containing data records on users having similar **profiles** (i.e. parking time-habits). Consequently, $\mathbb{X} \equiv \bigcup_{i=1}^{k} \pi_i$.

B. Parking Time Prediction

To perform the parking time prediction, we propose to use **data driven regression**. In regression, the goal is to determine a function $f(Z, \theta)$, given the input independent variables, Z, and the real values of the dependent variables, θ. The output of the model is not necessarily equal to the real value, due to noise in the data and/or limited number of entries. Consequently, a regression model commonly comprises an error e. The function f can be expressed as follows:

$$Y \approx f(Z, \theta) + e \tag{1}$$

Let $\mathbb{M} = \{M_{\pi_1}, M_{\pi_2}, ..., M_{\pi_k}\}$ be the set of k regression models and p_{j,π_i} denote the parking time prediction for a given timestamped user entrance with the profile π_i. $\mathbb{M}$ results of applying an induction method of interest to the datasets in Π. By doing so, the authors expect to approximate the real vehicles parking time given a set of describing variables (i.e.: Z).

C. Incremental Interval Generation

Given a prediction for the parking time of an user timestamped entrance (i.e. p_{j,π_i}), it is possible to estimate an **interval** for this value based on the residuals produced by its regression model. Hereby, we propose to do so by employing the residuals' *quantiles*. A quantile is a point taken from a cumulative distribution function of a variable. The first quantile represents the point that is greater than 25% of the data, while the third quantile the point that is greater than 75%. Let $e_{1,i}$ and $e_{3,i}$ denote the first and third quantiles of the regression *residuals* produced by a given model M_{π_i} on the previously tested data records in π_i. Our baseline interval I is given by the following equation:

$$I_{j,\pi_i} = [p_{j,\pi_i} - e_{1,\pi_i}, p_{j,\pi_i} + e_{3,\pi_i}] \tag{2}$$

Let a **hit** occur every time the real parking time is contained within the interval estimated. Otherwise, we consider the occurrence of a **miss**. Our goal is to produce intervals in order to maximize the number of hits and, at the same time, to minimize its **width**. To do so, we propose to extend the baseline described in eq. (2) by employing a *self-adaptive* strategy. Such strategy consists on multiplying the quantile-based interval width by a $0 \leq \beta \leq 2$ (starting on $\beta = 1$). This value is *incrementally* updated whenever an user of π_i leaves the parking lot (i.e. each time a newly real parking time is known on π_i). Let α_{π_i} denote the number of *consecutive* misses/hits of our interval prediction method in π_i. Whenever $\alpha_{\pi_i} > \alpha_{th}$, the value of β is incremented/decremented by τ. α_{th} and τ are two user-defined parameters setting how **reactive** the interval prediction model should be. Consequently, it is possible to re-write the eq. (2) into the following one:

$$I_{j,\pi_i} = [p_{j,\pi_i} - \Delta, p_{j,\pi_i} + \Delta], \Delta = (e_{3,\pi_i} - e_{1,\pi_i}) \times \beta \tag{3}$$

Everytime that a sequence miss/hit or hit/miss occurs, the respective α value is set to 0. The β ends up by controlling the interval width: the described algorithm aims to **adapt itself** to the current scenario by *narrowing* the intervals width whenever it is getting multiple hits or by *stretching* itself on the opposite scenario.

D. Parking Lane Selection

In this paper, the parking lot is assumed to follow a rectangular layout where the entrance and the exit are the same. It is possible to represent it as a $l \times r$ matrix, where l, r sets the number of **lanes** and the maximum number of vehicles in each lane, respectively. When a vehicle enters the parking lot,it is necessary to select a lane κ to park it in. Such selection should minimize the number of unnecessary vehicle movements (i.e. ϑ_κ). Consequently, each lane has an associated **score** W_κ. It can be faced as a likelihood of that selection force unnecessary movements given the i) current interval prediction for the newly arrived user (I_{j,π_i}) and ii) the vehicles already parked in κ. The lane with lowest score is predicted to be the one that minimizes ϑ_κ.

Empty lanes have a predefined score of $W = 1$ while a full one have $W = \infty$. Let h be the *last* vehicle in κ (i.e. the vehicle most recently parked), I^U_{j,π_i} be the upper limit and I^L_{h,π_b} be the lower limit of the estimated interval (note that the vehicle's j profile, π_i, may be (or not) the same of the vehicle h, π_b). If $I^U_{h,\pi_b} < I^L_{j,\pi_i}$, it is expected that the vehicle j of profile π_i exits the parking lot first than h (e.g.: Fig. 3-c). In this case, $W_\kappa = \infty$. If $I^U_{j,\pi_i} < I^L_{h,\pi_b}$, then it is expected that j and h can leave the parking lot provoking no unnecessary movements (i.e.: $\vartheta_\kappa = 0$; e.g.:

[1]Whenever a given vehicle c exits, all its lane's vehicles standing between c and the parking lot exit, have to be moved to a buffer zone. Such movements could be avoided by an exit-oriented sorting of each lane's vehicles.

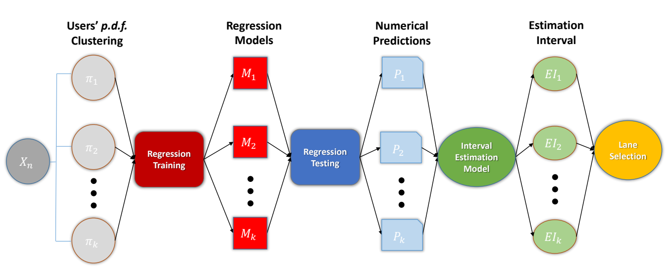

Fig. 2. An illustration on the different steps of the proposed methodology.

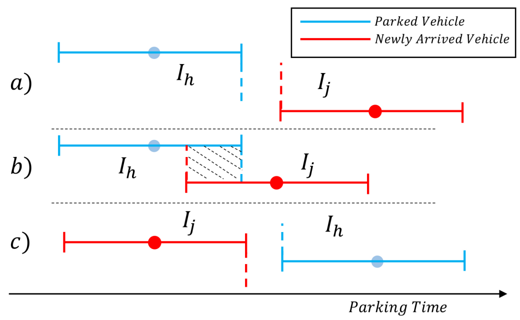

Fig. 3. In a), the upper limit of I_h is lower than the lower limit of I_j, so h is expected to leave the parking lot first than j. In b) there is an overlap between the two intervals. Its width is used to compute the lane's score. Finally, c) is the opposite scenario of a).

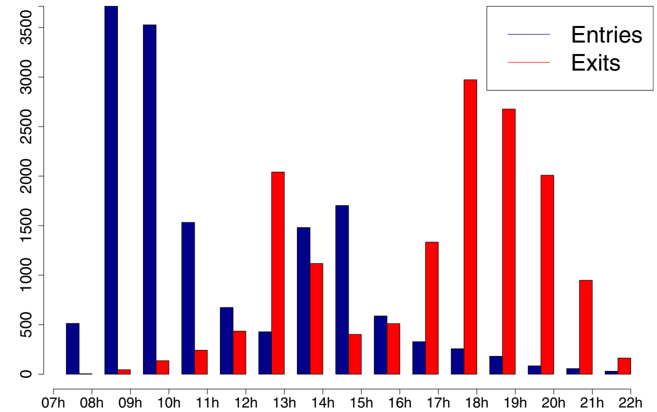

Fig. 4. Barplot chart representing histograms for the Entry/Exit times between 7am and 10pm.

Fig. 3-a). Consequently, the score is then $W_\kappa = 0$ on this case. Otherwise, W_κ can be computed as follows

$$W_\kappa = \frac{I^U_{j,\pi_i} - I^L_{h,\pi_b}}{I^U_{h,\pi_i} - I^L_{h,\pi_b}} + \frac{(N_\kappa - 1)^4}{r} \quad (4)$$

where N_κ stands for the number of vehicles currently in κ. This approach is inspired on the typical *p-value* statistical test considering a null hypothesis by setting the *extreme* data point as I^U_{j,π_i} and I_{h,π_i} as a *rough* approximation on the parking time distribution function for the parked vehicle h. The second term of eq. (4) is an exponential weight which aims to express the possible cost of having unnecessary vehicle movements caused by assigning the newly arrived vehicle j to the lane κ.

IV. Case Study

This case study consists on the parking lot of the Faculty of Science of University of Porto, Portugal. The data of 309 users during the year of 2013 was used to validate our methodology. This parking lot has the capacity to hold up to 100 vehicles. Since 96.4% of the data entries are in week days, only the workdays are considered in this study.

Each data record has the following features: (i) an user ID, (ii, iii) two timestamps for the parking entry/exit, (iv) type of day (e.g.: Monday), (v) holiday/not-holiday boolean and, finally, the (vi) department, (vii) sex and (viii) job role (e.g. Full Professor).

Ideally all data entries would have their entry and exit times properly labelled. However, it does not happen in this case because the parking entries/exits are not fully monitored. Consequently, there are entries without exits and vice-versa. To tackle such issue, a preprocessing task to pair the entries with the exits was performed. All the resulting data records with parking time smaller than 10 minutes or higher than 16 hours were removed. For the same reasons, we have also filtered the parking lot users by using the data records of the top-75%, regarding their number of parking entries.

In the resulting dataset, the average parking time is 5 hours and 25 minutes and with a standard deviation of 3 hours and 8 minutes. Fig. IV exhibits two histograms representing the hourly frequencies on the entry and exit times. It is possible to observe that the main entry times are between 8am and 10am and the main exit times between 5pm and 8pm. The vehicle's exits from the parking lot follows a bimodal distribution, with the modes at lunch time (between 12am and 2pm) and at late afternoon (between 5pm and 7pm).

V. Experimental Results

In this section, we start by describing the experimental setup used in our experiments and the evaluation metrics used to validate our methodology. Then, we present some experimental results and a brief discussion on their insights.

A. Experimental Setup

The initial dataset was divided in a training set (January to October) and a test set (November). All experiments were conducted using R Software [6]. The algorithms used were the k-Nearest Neighbours (kNN) [7], the Random Forests (RF) [8], the Projection Pursuit Regression (PPR) [9] and the Support Vector Machines (SVM) [10] from the R packages [`kknn`], [`randomForests`], [`stats`] and [`e1071`].

Regarding the feature selection, a well-known state-of-the-art technique was used: Principal Component Analysis (PCA) [11]. The tested features were type of day, holiday/not-holiday boolean variable and the user's department, sex and job role. For clustering we used the Expectation-Maximization algorithm with the R package [`MClust`]. This algorithm was chosen due to being able to determine the optimal number of clusters automatically based on Bayesian Information Criterion [12].

The last 2 weeks of the training set was used for model selection. In this stage, the following parameters were tested for each algorithm: for kNN, $distance = [1..5]$, $kMax = [2..15]$ and the kernels: rectangular, triangular, epanechnikov, gaussian, rank and optimal, for RF $mtry = \{3,4,5\}$ and $ntrees = \{500,750,1000\}$, for PPR $nterms = \{2,3,4\}$ and $max.terms = \{5,6,7,8\}$ and for SVM the kernels: linear, radial, polynomial and sigmoid. The best pair (algorithm,parameter setting) was selected to perform the numerical prediction in the test set.

Finally, the reactiveness parameters on the interval estimation model (τ, α_{th}) were set for the values 0.1 and 3, respectively.

To evaluate our method performance, we considered a baseline naive strategy. It consists on directing the newly arrived vehicle to the leftmost lane κ with an empty space. A series of simulations were conducted to compare the parking lot behavior using the aforementioned lane selection strategies (i.e. *naive* and *smart*). Multiple parking layouts were considered on this series of simulations. It aimed to demonstrate that the strategies behavior is **independent** on the parking layout. The averaged maximum number of parked cars on a daily basis on the considered dataset is 50. Consequently, every parked layouts with a capacity between 50 and 80 vehicles (i.e.: the 1^{st} quantile) containing, at least, 8 lanes, were considered on our experiences.

B. Evaluation

The root-mean-squared-error (RMSE) and the mean absolute error (MAE) were the metrics used to evaluate the predictions. They can be defined as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{g}(\hat{y}_t - y_t)^2}{g}}, MAE = \frac{\sum_{t=1}^{g}|\hat{y}_t - y_t|}{g} \tag{5}$$

where $\hat{y}$ is the predicted value, y the real one and g is the number of samples.

The parking time estimation interval is evaluated in two forms, a percentage of hits and a ratio between the hits and its width. If for a sample s there is a hit, then $hit_s = 1$, otherwise $hit_s = 0$. The ratio can be defined as:

$$ratio = \sum_{s=1}^{g} hit_s \times \frac{1}{\delta_I \times g} \tag{6}$$

where δ_I is the width of the estimation interval and g is the number of considered samples.

The evaluation criteria employed in the simulation was the total number of unnecessary vehicle movements forced by a given strategy (i.e., UM). Let us consider a exiting vehicle c, parked in a lane κ with g vehicles, in position i. The unnecessary number of movements UM caused for c to exit the parking lot can be computed as:

$$UM = \sum_{j=1}^{g-i} j \tag{7}$$

Let us consider a lane with $g = 5$ vehicles where the vehicle on the position $i = 2$ is requested to exit as an exemplification for the calculus of MU. In this case, $MU = 3 + 2 + 1 = 6$.

C. Results

The obtained results are three fold: (1) the PCA results have recommended to remove the user's sex and the holiday feature from the original set. (2) Table I exhibits the results of the numerical prediction using the remaining feature set for each profile π_i, by pointing the number of users contained in each group and the (RMSE,MAE) obtained in each one of them. (3) Table II shows the results from the parking simulation in every tested configurations, with the number of unnecessary vehicle movements, μ for both strategies. The intervals generated had **65%** hits and an average interval width of $\approx$ **11000** seconds. The *smart* strategy overcomes the *naive* one in all the considered configurations.

TABLE I

RESULTS FROM THE NUMERIC PREDICTION.

Group	# of Individuals	RMSE	MAE	Hit %	Interval
1	11	5124	3320	63	8942
2	9	4804	3255	66	3862
3	3	7047	5235	68	9584
4	6	4644	4047	78	9764
5	1	7716	5458	82	3482
6	1	376	340	72	3504
7	5	3968	3317	68	9196
8	7	7618	6101	58	11738
9	11	9106	7628	53	11900
10	6	8244	7403	55	12560
11	4	2609	2058	72	5255
12	10	7871	5436	67	9583
13	6	8901	5789	72	9558
14	7	8595	6883	54	11228
15	4	5981	4804	50	6258
16	10	6682	5356	70	10293
17	1	361	298	50	3158
	W.Average	**6601**	**5076**	**65**	**11188**

TABLE II

SIMULATION RESULTS WITH THE NUMBER OF UNNECESSARY VEHICLE MOVEMENTS FOR BOTH STRATEGIES.

Config.	Naive	Smart	Config.	Naive	Smart
10x05	1665	**1379**	*05x10*	7799	**7540**
11x05	1482	**1205**	*05x11*	7817	**7615**
12x05	1255	**1074**	*05x12*	7817	**7633**
13x05	1074	**914**	*05x13*	7817	**7633**
14x05	937	**813**	*05x14*	7817	**7633**
15x05	811	**771**	*05x15*	7817	**7633**
09x06	2234	**2032**	*06x09*	5596	**5423**
10x06	1819	**1583**	*06x10*	5596	**5444**
11x06	1510	**1282**	*06x11*	5596	**5453**
12x06	1255	**1139**	*06x12*	5596	**5453**
13x06	1074	**930**	*06x13*	5596	**5453**
08x07	2808	**2520**	*07x08*	3818	**3545**
09x07	2248	**2116**	*07x09*	3818	**3545**
10x07	1819	**1616**	*07x10*	3818	**3551**
11x07	1510	**1303**	*07x11*	3818	**3551**
07x08	3818	**3545**	*08x07*	2808	**2520**
09x08	2248	**2116**	*08x09*	2808	**2535**
10x08	1819	**1617**	*08x10*	2808	**2535**
08x08	2808	**2535**			

D. Discussion

Table I exhibits a large variation on RMSE/MAE produced by the models of the different groups. The groups size is also different from group to group. These groups can be faced as **profiles** which describe the *typical* parking behavior of the users within. It is possible to observe that some groups contain only one user (i.e. 5,6,17) which indicates that they have a completely different profile than the remaining ones. So far, such profiles are only based on each user's parking time (namely, by using the Euclidean Distance over their *p.d.f.*). However, some users can experience large variations on their parking time depending on some subsets of feature values (i.e. to enter the parking lot at morning or at afternoon). This fact can partially explain the above mentioned RMSE/MAE variability.

The averaged hits percentage (65%) and its large width uncover the stochasticity of the parking time variable given the current feature set. In fact, it is reasonable to admit that we may need other features to improve our prediction model such as weather or event-based ones (e.g. a sunny day or a special soccer match may reduce/increase the parking time). However, we cannot sustain these insights on the present results.

The *naive* strategy is clearly benefited by configurations with more lanes, where the UM can be naturally minimized by underusing the total lane's capacity by filling first the empty ones. In fact, this strategy is already focused on minimizing UM by maintaining the maximum number of vehicles parked on a lane as **low** as possible. Such behaviour can explain some of the lower gain margins presented by the *smart* strategy on some configurations (check Table II). Obviously, the UM could also be minimized by moving vehicles from one lane to another. However, the discussions about the optimal parking layout for each case study and on the parked vehicle's self-arrangements are out of this paper's scope.

Even considering the abovementioned drawbacks, the authors want to highlight that the **proposed methodology overcomes the *naive* strategy for all the presented parking layouts**. The aim with this work is to demonstrate that is possible to **mine** both the historical and the real-time data on the parking lot entrances/exits to improve the lane selection on a self-automated parking lot. This stepwise framework takes advantage of *off-the-shelf* Machine Learning algorithms to do it so. In our opinion, this proof of concept represents a consistent **breakthrough** on this relevant topic by opening promising research lines to be explored by other researchers.

VI. Final Remarks

Throughout this paper, a Machine Learning framework to predict the exit times on a self-automated parking lot is proposed. It consists on using historical data on the entries/exits on the parking lot to uncover user's profiles able to explain their parking habits. Our goal is to optimize the vehicle's initial placement by improving the lane selection using such predictions. The experiments demonstrated that our method can overcome a naive strategy by **reducing the collaborative mobility needs on roughly 10%**. By doing so, we hope to open new research lines on this topic.

As future work, we propose to explore the inter-lane vehicle movements to re-arrange their placements. Such movements aim to *react* to the parking current status by a) updating the exit time predictions while the vehicles are still parked or by b) moving the blocking vehicles to their neighbour lanes instead of using the buffer. The validity of such hypothesis comprise open research questions.

References

[1] D. C. Shoup, A. P. Association, *et al.*, *The high cost of free parking*. Planners Press Chicago, 2005, vol. 206.

[2] D. C. Shoup, "Cruising for parking," *Transport Policy*, vol. 13, no. 6, pp. 479–486, 2006.

[3] M. Ferreira et al., "Self-automated parking lots for autonomous vehicles based on vehicular ad hoc networking," in *Proc IEEE Intelligent Vehicles Symp. - IV*, Dearborn, Michigan , United States, June 2014.

[4] T. T. Sebastian Gnatzig, Frederic Chucholowski and M. Lienkamp, "A system design for teleoperated road vehicles," in *ICINCO 2013 - Proceedings of the 10th International Conference on Informatics in Control, Automation and Robotics*, 2014, pp. 231–238, 10th International Conference on Informatics in Control, Automation and Robotics, ICINCO 2013.

[5] M. Ferreira, R. Fernandes, H. Conceição, W. Viriyasitavat, and O. K. Tonguz, "Self-organized traffic control," in *Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking*. ACM, 2010, pp. 85–90.

[6] R. C. Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012. [Online]. Available: http://www.R-project.org

[7] N. S. Altman, "An introduction to kernel and nearest-neighbor non-parametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[8] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[9] J. H. Friedman and W. Stuetzle, "Projection pursuit regression," *Journal of the American statistical Association*, vol. 76, no. 376, pp. 817–823, 1981.

[10] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[11] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.

[12] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.