

Characterising Demand and Usage Patterns in a Large Station-based Car Sharing System

Chiara Boldrini

IIT-CNR

Pisa, Italy

Email: chiara.boldrini@iit.cnr.it

Raffaele Bruno

IIT-CNR

Pisa, Italy

Email: raffaele.bruno@iit.cnr.it

Marco Conti

IIT-CNR

Pisa, Italy

Email: marco.conti@iit.cnr.it

Abstract—Car sharing is a new mode of transportation that is gaining increasing popularity with its promise to reduce traffic congestion, parking demands and pollution in our cities. Despite this potential, the properties of car sharing systems, e.g., in terms of spatiotemporal characterisation of how customers use the service, remain largely unexplored in the research literature. In order to fill this gap, in this work we analyse one month of online car-sharing map data from a large station-based car-sharing operator in France, which has 960 stations and more than 2700 electric cars. First, we study the spatial and temporal patterns of station utilisation, uncovering a dichotomy in station usage (stations that attract cars mostly in the morning vs. stations attracting cars mostly in the evening). We also find that this dichotomy is linked to the destination (residential or business) of the zone in which the station is located. In addition, we statistically model the users' demand in terms of drop-off and pickup rates, and the parking times of vehicles. Finally, we propose a classifier that exploits simple average statistics (average pickup rate and car availability of a station) in order to understand whether the station is profitable or not for the operator.

I. INTRODUCTION

In the last years, car sharing systems have emerged as a credible alternative to car ownership [1], [2]. Existing car sharing solutions have shown great potential in reducing parking needs and in promoting both public transport and non-motorised transportation modes (biking and walking), thus lowering households' transportation costs and traffic congestion in our cities [3]. Traditionally, car sharing systems can be distinguished into two-way, one-way, and free floating. Both in two-way and one-way car sharing, cars can be picked up only at designated locations (called *stations*). In two-way car sharing (e.g., Zipcar, Modo), customers are required to drop-off the vehicle at the same station where they have initially picked it up. This constraint is dropped into one-way car sharing. Examples of one-way car sharing are Autolib, Ha:Mo ride, CITIZ. A further step towards maximum flexibility for the users is represented by free floating car sharing (such as Car2go, DriveNow, Enjoy), in which the concept of stations disappear and cars can be picked up and dropped off in any public parking within the area (geofence) in which the car sharing is operating.

The increasing flexibility provided by one-way and free-floating car sharing has a cost in terms of managing the complexity of the car sharing operations. In fact, some lo-

cations end up being more popular than others (hot spots vs. cold spots), so typically there is a shortage of cars in hot spots and idle cars in cold spots. When this happens, the operator can resort to redistributing vehicles, i.e., moving vehicles from where they are not needed (taking into account the expected demand in the near future) to hot spots [4]. Clearly, this has a cost for the operator, thus redistribution should be performed only when economically viable. In the case of station-based systems (which will be the focus of this work), a major cost entry is also the deployment of stations. Additional investments are needed when the fleet is composed of electric vehicles (EV). In fact, EVs need recharging, which implies that i) setting up a station becomes more expensive due to the need for the recharging infrastructure ii) some EVs cannot be picked up until they have recharged enough for a customer to complete a trip. Thus, it is crucial to understand how many stations should be deployed, where, what should be their capacity, and how many charging points are needed [5].

To address the above problems, optimisation frameworks and operational decision tools for car sharing systems have been studied in the literature, but the proposed solutions have been evaluated either on simulated scenarios [6], [7] or using as input the demand (in terms of origin/destination matrix) obtained from surveys [8], [9]. On the contrary, the availability of a statistical characterisation of the general properties of real car-sharing systems, as well as a precise understanding of their emerging trends, is essential to both researchers and operators in order to design more effective decision support tools, and for the calibration and validation of simulations of car sharing systems. However, to the best of our knowledge the properties of car sharing systems, e.g., in terms of spatiotemporal characterisations of how customers use the service, remain largely unexplored in the research literature.

In order to fill this gap, in this work we exploit the availability of public data about a large one-way station-based car-sharing system operating in Paris (France) and its suburban areas, which has 960 stations and more than 2700 electric cars. A detailed description of this dataset is provided in Section III. Using this dataset as input, we carry out an analysis whose main contributions are the following.

- We characterise the car sharing system in terms of station capacity and station utilisation. As expected, the evolution of usage during the day is marked by the typical daily

milestones such as travelling to work in the morning, lunch, dinner, going out in the evening. In the center of Paris we also observe marked differences between weekdays and weekends, which are not present in the suburban area.

- We apply clustering techniques to group stations based on how they are used by the car sharing customers. Only two temporal patterns emerge, highlighting a clear dichotomy in how stations are used. Specifically, there are stations that attract cars mostly in the morning and stations attracting cars mostly in the evening. The first type of stations are typically located in business districts, while the others are in residential areas.
- We model statistically the demand at each station in terms of drop-off and pickup rates. Given its importance for recharging, we also statistically characterise the distribution of parking times. We show that well-known distributions (such as Normal, Lognormal and Gamma) are suitable to represent the distribution of these important properties of the car sharing system, and we provide the best fitting values of each distribution.
- We design a heuristic method to classify stations based on the profit they generate, using only two simple average statistics (average pickup rate and car availability of a station).

II. RELATED WORK

Knowledge about car sharing systems has been so far (to the best of our knowledge) acquired through surveys [3], [10], in which car sharing operators and members are interviewed. The understandings and advancements brought about by these works are invaluable, but the collection of survey data is expensive, time consuming, and does not scale. For these reasons, in this work we depart from this approach and we exploit public, web-based, digital records, whose geotagged and time-stamped variety of data can be analysed employing data mining techniques.

From the methodology standpoint, this work is close to [11], [12], in which bike-sharing, rather than car-sharing, systems have been analysed. Due to the different nature of the two systems, people use them differently, hence the results obtained for bike sharing systems cannot be applied directly to car sharing. However, similar methodologies can be exploited, e.g., to group stations based on how they are used by the customers.

This work is also orthogonal to the research efforts in the area of car pooling/ride sharing [13], [14]. The idea of car pooling/ride sharing is that people may share a vehicle (be it a private or public vehicle, e.g., a taxi cab) to perform their trips. Works in the area of car pooling typically focus on the amount of rides that can be shared, based on the historical or real-time trajectories of users, hence their focus is very different from that of this work.

III. THE DATASET

The dataset that we use in this study is composed of all the pickup and drop-off times at 960 car sharing stations in Paris and the surrounding area for the whole month of

April 2015. Data are scraped from the car sharing online map. We collect observations every 2 minutes, for a total of 1,881,727 observations. Each observation reports the state of the station in terms of: number of available cars, number of available parking lots, status of the station (operational or in maintenance). We also know the address and the GPS coordinates of each station, hence we can derive whether a station is within Paris or in the suburban areas outside Paris.

The only cleaning that we perform on the dataset is based on the operational status of stations. Specifically, we filter out all stations that are operational for less than 99% of the time, in order to get rid of anomalous behaviours due to stations in maintenance. This leaves us with 900 stations (535 inside Paris, 365 in the suburban area) and 1,664,751 observations. Similarly to [12], we define the state of station $s_i \in S$ at time $t \in T$ as $s_i(t) : c_i(t), p_i(t)$, where $c_i(t)$ and $p_i(t)$ denote the number of available cars and the number of available parking lots¹, respectively. The interplay between $c_i(t)$ and $p_i(t)$ is explained in detail in the following section.

IV. THE ANALYSIS

A. Station capacity and car availability

We start our analysis with the station capacity ρ_i , defined as the maximum number of available cars at a station i . Since not all parking lots may be filled, we compute the station capacity as the sum $c_i(t) + p_i(t)$ of the number of cars parked at the station and the number of unoccupied parking lots. The car sharing system under study allows users to reserve both cars and parking lots in advance. When a car is reserved, $c_i(t)$ is decreased, but $p_i(t)$ is increased only when the car is actually picked up. Similarly, when a parking lot is reserved, $p_i(t)$ is decreased but $c_i(t)$ is not. For this reason, the instantaneous value of $c_i(t) + p_i(t)$ may not tell the whole story about the capacity of the station. Owing to this fact, we compute the capacity as the maximum observed value of $c_i(t) + p_i(t)$ during the observation period. One month is a large enough window to assume that there is at least one t at which no reservations are active at a station.

In Figure 1 we plot the histogram of the capacity of all stations in the systems, distinguishing between stations in Paris and in the suburbs. We observe that most stations have a capacity greater than 5, and that station size tends to be larger in the suburban area. However, if, for each station, we compute the average number of stations that are within a 300m radius², we obtain 1.15 for the Paris area (with the closest station being on average at 264m) and 0.2 for the suburban area (with the closest station being on average at 502m). So, even if suburban stations tends to be larger, they are also sparser.

¹Please note that the available parking lots are an asset for the car sharing system under study for two main reasons. First, customers can only end their trip at a station. Thus, if there is no parking space at their destination station, they have to drive to the closest station with available parking lots, and this has a cost in terms of customer satisfaction. Second, each parking lot is equipped with a charging pole, which is a costly infrastructure for the operator.

²300m is the rule of thumb used in London bikesharing for station placement (<http://oobrien.com/category/bike-share/>). The rationale is that people are not willing to walk more than 300m to reach a station or to move from a station to another one in case the previous one is empty.

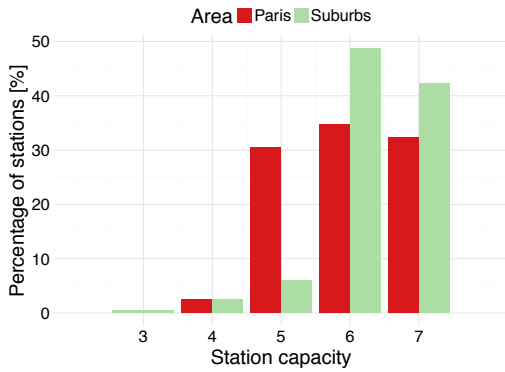


Fig. 1. Histogram of station capacity.

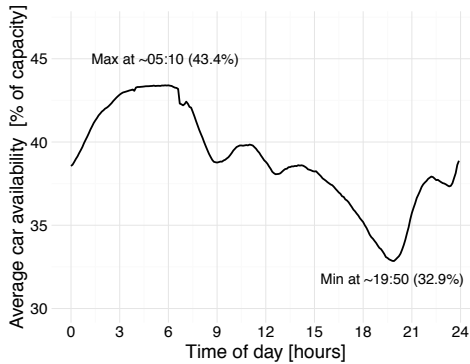


Fig. 2. Time series of car availability in the system.

The capacity ρ_i is typically a static metric associated with stations, and its changes are sporadic (e.g., if a station is expanded). What changes frequently, instead, is the car availability at stations, which is an indicator of how stations are used. We compute the (percentage) car availability as $\frac{c_i(t)}{\rho_i} * 100$. In order to get a compact view of the system, we divide each 24 hours day into bins of 6 minutes. For each station, we take the average availability within each bin. Then we average across all stations. The resulting time series is plotted in Figure 2. The average availability never goes above 50%. The maximum availability is observed at night (between 3am and 6am), the minimum in the evening (roughly at 8pm). The local minima corresponds to typical milestones in the day, such as people starting to work (~9am), people going for lunch (~1pm), people ending their working day (~7pm), people going out in the evening (~11pm).

In order to investigate more in detail the temporal and geographical patterns of usage, in Figure 3 we plot the time series of availability in the two main areas (Paris and suburbs) in which the service is operating, distinguishing between availability during weekdays and during weekends. We observe that people start their day later during weekends (car availability drops later in the day), both in Paris and in the suburban area. Apart from that, in the suburban area there are no big differences between weekdays and weekends (in fact, the correlation coefficient between the two time series is 0.92). In particular, in the suburban area there is a roughly continuous negative slope from the morning until

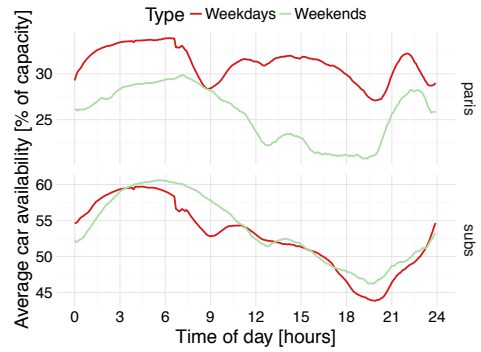


Fig. 3. Time series of car availability in Paris and the suburban area.

the evening, implying that the rate at which cars are dropped off is smaller than the rate at which cars are picked up. This might be due to cars being picked up in the suburban area then dropped off in Paris or to cars being used for the whole day (e.g., in the weekend for short trips). In Paris we observe two opposite behaviours during the working hours in a weekday and the same time of the day in weekends. In weekdays, the availability roughly plateaus towards the higher end during the working hours in a weekday. This might be due to customers using less the car sharing service or to the injection of vehicles from the suburban areas. Vice versa, during weekends there is a plateau of low availability in the afternoon, meaning that many are in use. It is also interesting to notice the striking match between the evening and night behaviour in weekends/weekdays, both within Paris and within the suburban area. If we look at the absolute numbers of the availability, we observe that in Paris car availability is always much lower than in the suburban area. In addition, car availability in Paris is lower during the weekend, even at night (which can be considered a baseline availability, since the usage should be at its lowest). Speculations can be made about the fact that either maintenance is performed during the weekend (hence there are fewer cars available to the customers) or people simply use car sharing more (this comes in two flavours: the same basin of users keep a car for a longer time or more users are active in the car sharing system at weekends). This effect is not detected in the suburbs though, where the availability during weekends can even be higher than during weekdays.

B. Many stations, two usage patterns

So far, we have performed an aggregate analysis of the car sharing system, in which stations have been aggregated in order to capture some general spatiotemporal trends in the system. The next step of our analysis aims at characterising stations individually. Specifically, considering again car availability, we investigate whether stations can be classified based on their usage patterns. For this classification, we use three standard techniques: Dynamic Time Warping (DTW) [15] for measuring how “close” two time series are, Partitioning Around Medoids (PAM) clustering to create k groups of similar stations based on the DTW distance, and the Silhouette

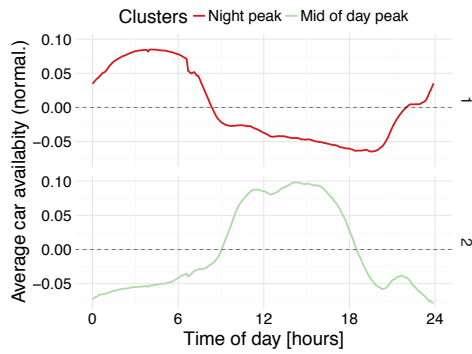


Fig. 4. Time series of car availability in the two clusters, normalised.

method for selecting the best k . The result of this analysis is that the most informative clustering is obtained when $k = 2$ (mean silhouette width 0.48), thus bringing out a dichotomic behaviour in the car sharing system.

In order to better characterise the features of these clusters, in Figure 4 we plot the car availability averaged within the two clusters (normalising by its time average to highlight zones in which the availability is above and below the average value). The behaviour in the two clusters is intuitive and informative: in the first group, stations have high availability at night and below-average availability during the day, vice versa in the other group there is a high availability during the day and below average availability at night. We call these two groups Night Peak and Mid of Day Peak, respectively. A night peak is most likely observed at stations in residential areas, where people leave in the morning and come back in the evening. A mid of day peak is typical of commercial and industrial areas, where people arrive in the morning and leave in the afternoon. This is confirmed if we plot the stations on a map and we color them based on the cluster they belong to (Figure 5). There is a strong geographical correlation between stations belonging to the same cluster, especially inside Paris. Stations in downtown Paris all show a mid of day peak (cars arrive in the morning and leave in the afternoon), vice versa for the other stations. In the suburbs Night Peak stations dominate: people leave in the morning and come back in the afternoon. There are some Mid of Day Peak stations in the suburbs as well, which in some cases are easy to relate to notable business districts in the zone (e.g., the group of green spots north-west of Paris corresponds to La Défense).

This classification based on the way stations are used can be very useful for the day-to-day operations of car sharing systems, for example for vehicle redistribution. It is a well known problem [4] that sometimes there are empty stations in an area of high demand and at the same time stations with several cars in areas of low demand³. The solution to this unbalance in the system is to move cars from one area to the other one, but this redistribution operation is costly for the car sharing operator (personnel costs plus the costs of

³In this dataset, for example, there can be up to 75% of Paris stations that are empty (the suburban area seems to suffer much less from this problem, as this percentage is rarely above 30%).

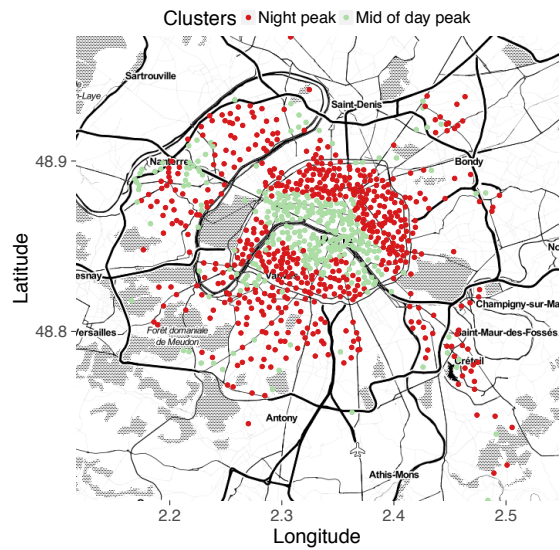


Fig. 5. Map of stations coloured by cluster. In Paris, 304 stations feature a night peak, 231 a mid of day peak. In the suburbs, only 86 stations show a mid of day peak, while 279 feature a night peak.

these “rides without customers”), thus it has to be optimised as much as possible. To this aim, being able to characterise the demand is crucial. Our clustering results, for example, tell the operator that it would not be effective to move cars between stations belonging to the same cluster, because their demand is synchronised (they observe the highest pickup rates at roughly the same time). Instead, redistributing cars between stations in different clusters is a smart and simple approach because the two stations have demand peaks at opposite time of the day.

C. Parking time, drop-off and pickup rates

In this section we want to characterise statistically the distribution of some important metrics of a car sharing system, namely the parking time, drop-off rate, and pickup rate. To this aim, we take the following approach. First, we identify the set of candidate distributions (among well-known distributions) using the skewness-kurtosis graph, then we use Maximum Likelihood Estimation (MLE) to fit the candidate distributions to the observations. Afterwards, we verify that these fitted candidate distributions are actually plausible hypotheses (i.e., that they cannot be rejected) using the Cramér-von Mises goodness-of-fit test. In case more than one distribution is found to be plausible, we pick the best one using the loglikelihood ratio test.

Let us start with the parking time, defined as the time interval between the drop-off and subsequent pickup of a vehicle⁴. Parking times are very important from the operations point of view, in particular for electric car sharing, since they correspond to the time available to the operator for recharging the vehicle. Longer parking times make the recharging process easier but the fact that vehicles remain parked for a long

⁴Since in our trace we do not have information on the vehicle ids, we make the assumption that vehicles are picked up according to a first-in-first-out policy.

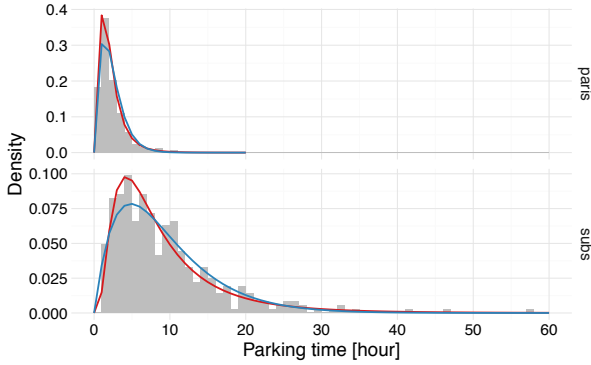


Fig. 6. Parking times with Gamma (blue) and Lognormal (red) fit.

time also means smaller revenues for the operator. We plot in Figure 6 the histogram of parking times, distinguishing between Paris and the suburban area. We observe that there is a big difference between them, with parking times in suburban areas reaching very high values (much more than one day in some case). The skewness-kurtosis graph of parking times (omitted for space reasons) tells us that the observations (both for Paris and the suburban area) fall between a Gamma and a Lognormal behaviour. Both hypotheses are not rejected when applying the Cramér-von Mises goodness-of-fit test but, according to the loglikelihood ratio test, the Lognormal distribution provides the best fit for both Paris and the suburbs (with p -value equal to $2.8 \cdot 10^{-8}$ and 0.057, respectively).

Another important property of a car sharing system is the drop-off rate, i.e., the rate at which vehicles are dropped-off at stations. We compute this rate as the inverse of the mean interarrival time of vehicles for each station and we plot its density in Figure 7. The drop-off rate is generally higher for stations in Paris than for stations in the suburban area, and the distribution is different in the two cases. The Normal distribution is singled out by the skewness-kurtosis graph as the only plausible fit for Paris, and the Cramér-von Mises test is not able to reject this hypothesis. We do not perform the likelihood ratio test because there are no other candidate distributions. The best fit for the suburban case is more challenging. The skewness-kurtosis graph identifies the Gamma distribution as the closest well-known distribution but its distance from the observations is not negligible. As expected, the fitting barely passes the Cramér-von Mises test. The behaviour of the pickup rate mirrors that of the drop-off rate, so we do not add further considerations about it. We summarise the results of the MLE best fit for parking times, drop-off and pickup rate in Table I.

TABLE I
RESULTS OF THE MLE FIT (α AND β CORRESPONDS RESPECTIVELY TO THE SHAPE AND RATE OF THE GAMMA DISTRIBUTION).

	Paris	Suburbs
Parking time	$\ln \mathcal{N}(\mu = 0.63, \sigma = 0.66)$	$\ln \mathcal{N}(\mu = 1.98, \sigma = 0.74)$
Drop-off rate	$\mathcal{N}(\mu = 0.86, \sigma = 0.28)$	$\Gamma(\alpha = 3.35, \beta = 6.30)$
Pickup rate	$\mathcal{N}(\mu = 0.86, \sigma = 0.28)$	$\Gamma(\alpha = 3.31, \beta = 6.25)$

As a final remark, please note that, for all the three metrics

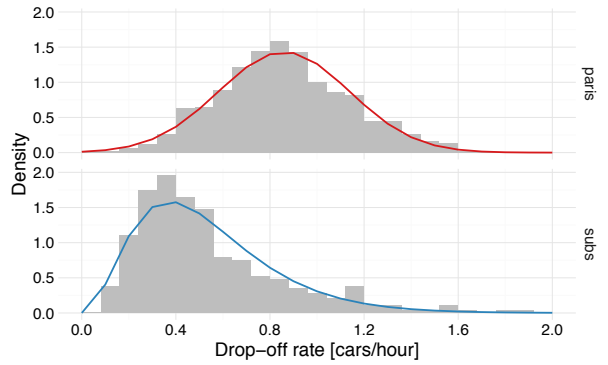


Fig. 7. Drop-off rates with the Normal (red) and Gamma (blue) fit.

above, we only reported how they vary depending on the area under study (Paris or suburbs) because the cluster to which stations belongs to, and the weekend/weekdays temporal split have little or no effect on the distribution. This is an important feature of the system that can be exploited, e.g., when modelling it for planning recharging or vehicle redistribution. To this aim, the network of car sharing stations could be modelled as a network of queues [16]. Each station is treated as a queue, and its arrival rate and service rate can be either estimated from the dataset or drawn from the distribution of drop-off and pickup rates of the relevant area (Paris/suburbs). The latter strategy allows us also to include additional stations that were not in the dataset, for the purpose of experimenting with different sizes of the system.

D. Put your money where your demand is

In this section we propose a simple classifier that links together the average pickup rate and the average car availability of a station in order to find out what are the stations on which the operator may be losing money. Please note that this kind of analysis would have been straightforward if the dataset had contained the billing details of car sharing journeys, but these data are rarely shared openly by the operator (recall that we downloaded data that were publicly available on the Internet).

In a car sharing system, it is reasonable to assume that the revenue at a station is directly proportional to the number of journeys originating at the station, since there is a monetary transaction between the customer and the car sharing operator every time a new vehicle is picked up. For this reason, we assume that the pickup rate at a station is a strong indication of how much money is made thanks to this station⁵. Thus, a station with a low average pickup rate in general should raise a red flag to the operator. However, these stations may be imposed by agreements with local authorities. What is critical for the operator is that the resources (vehicles) assigned

⁵We acknowledge that even with a low pickup rate the station could be remunerative for the operator: this happens if all the (few) trips departing from the station are long trips. Since the billing system is typically designed in order to discourage using the car for a long time, in this work we assume that the pickup rate is indeed a good indicator of how remunerative a station is for the operator. However, as future work we plan to refine this preliminary classifier to take into account additional parameters, such as the length of trips.

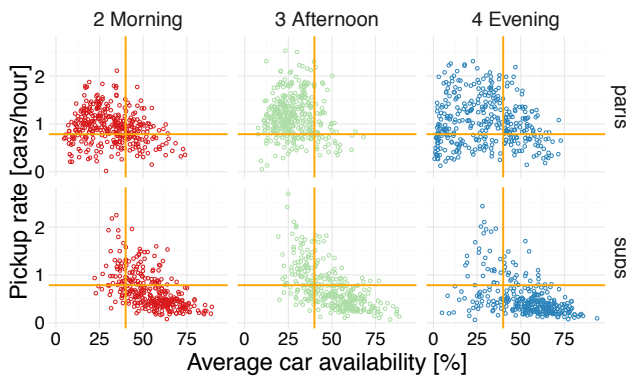


Fig. 8. Performance of stations measured as a function of pickup rate and average car availability.

to these stations are proportional to the journeys that they generate. Otherwise, vehicles get stuck at a station where nobody is going to pick them up, thus damaging the service in more profitable areas. A good indication of the quantity of resources blocked by a station is provided by the average car availability. Thus, based on these considerations, it is then natural to identify as problematic those stations for which the pickup rate is low and at the same time car availability is high.

Figure 8 shows a scatterplot of average pickup rate against average car availability (each point corresponds to a station), divided into morning (6am-12pm), afternoon (12pm-6pm), evening (6pm-12am). For convenience of analysis we set thresholds for the pickup rate and the capacity equal to the 50th percentile of their aggregate distribution, corresponding to a pickup rate of 0.8 and an availability of 40%. From the figure we observe a clear, opposite trend between Paris and the suburban areas. In Paris, many stations fall in the top-left corner, meaning high pickup rate and low average availability. This is where the operator is making money: cars are picked up frequently and only a bunch of them remain idle at the stations. Vice versa, in suburban areas the vast majority of stations fall in the bottom-right corner: a lot of idle vehicles, with a low pickup rate. The operator is possibly losing money on these stations, which thus become the ideal candidates for actions such as vehicle redistribution.

V. CONCLUSION

This paper has sought to demonstrate the insights offered to the management of car-sharing systems by the collection and analysis of car sharing system data. In particular, using as input the pickup and drop-off events at the stations of a large electrical car sharing operator, we have characterised the capacity and the utilisation of the stations in this system, highlighting different daily, weekly, and geographical usage patterns. We have also grouped stations based on the way they are used by the car sharing customers, discovering a dichotomy (night vs mid-of-day availability peak) that characterises the system and that can be exploited for optimising redistribution operations. Using this dataset, it was also possible to characterise statistically the demand at each station in terms

of pickup and drop-off rates, as well as the parking times (which are essential for optimising the recharging of EVs). Finally, we have discussed a simple classifier that, exploiting the metrics that we have investigated in the paper, is able to uncover stations which may not be profitable for the car sharing operator. As future work we intend to refine this preliminary classifier in order to include additional parameters and socio-economical data that are available for the city under study. Furthermore, we plan to exploit the rates studied in Section IV-C to model the car sharing system as a network of queues.

ACKNOWLEDGMENT

This work was partially funded by the ESPRIT project. This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 653395.

REFERENCES

- [1] R. Botsman and R. Rogers, *What's mine is yours: how collaborative consumption is changing the way we live*. Collins London, 2011.
- [2] R. Hampshire and C. Gaites, "Peer-to-peer carsharing: Market analysis and potential growth," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2217, pp. 119–126, 2011.
- [3] S. Shaheen and A. Cohen, "Mobility and the Sharing Economy: Impacts Synopsis – Spring 2015," Transportation Sustainability Research Center, University of California, Berkeley, Tech. Rep., 2015.
- [4] S. Weikl and K. Bogenberger, "Relocation strategies and algorithms for free-floating car sharing systems," *Intelligent Transportation Systems Magazine, IEEE*, vol. 5, no. 4, pp. 100–111, 2013.
- [5] B. Boyaci, K. G. Zografos, and N. Geroliminis, "An optimization framework for the development of efficient one-way car-sharing systems," *European Journal of Operational Research*, vol. 240, no. 3, pp. 718–733, February 2015.
- [6] M. Nourinejad, S. Zhu, S. Bahrami, and M. J. Roorda, "Vehicle relocation and staff rebalancing in one-way carsharing systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 81, pp. 98–113, 2015.
- [7] K. Uesugi, N. Mukai, and T. Watanabe, "Optimization of vehicle assignment for car sharing system," in *Knowledge-based intelligent information and engineering systems*. Springer, 2007, pp. 1105–1111.
- [8] D. Jorge, G. H. Correia, and C. Barnhart, "Comparing optimal relocation operations with simulated relocation policies in one-way carsharing systems," *IEEE Transactions on Intelligent Transportation Systems*, 2014.
- [9] G. H. de Almeida Correia and A. P. Antunes, "Optimization approach to depot location and trip selection in one-way carsharing systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 233–247, 2012.
- [10] B. Schwiager, P. Victorero-Solares, and D. Brook, "Global carsharing operators report 2015," Team Red, Tech. Rep., 2015.
- [11] O. O'Brien, J. Cheshire, and M. Batty, "Mining bicycle sharing data for generating insights into sustainable transport systems," *Journal of Transport Geography*, vol. 34, pp. 262–273, 2014.
- [12] A. Sarkar, N. Lathia, and C. Mascolo, "Comparing cities' cycling patterns using online shared bicycle maps," *Transportation*, pp. 1–19, 2015.
- [13] R. Trasarti, F. Pinelli, M. Nanni, and F. Giannotti, "Mining mobility user profiles for car pooling," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 1190–1198.
- [14] P. Santi, G. Resta, M. Szell, S. Sobolevsky, S. H. Strogatz, and C. Ratti, "Quantifying the benefits of vehicle pooling with shareability networks," *PNAS*, vol. 111, no. 37, pp. 13 290–13 294, 2014.
- [15] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, p. 12, 2012.
- [16] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi, *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*. John Wiley & Sons, 2006.