E-Scooter Sharing: Leveraging Open Data for System Design

Alessandro Ciociola[a], Michele Cocca[a], Danilo Giordano[b], Luca Vassio[a], and Marco Mellia[a]
[a]Department of Electronics and Telecommunications, Politecnico di Torino, first.last@polito.it
[b]Department of Control and Computer Engineering, Politecnico di Torino, first.last@polito.it

***Abstract*—With the shift toward a Mobility-as-a-Service paradigm, electric scooter sharing systems are becoming a popular transportation mean in cities. Given their novelty, we lack of consolidated approaches to study and compare different system design options. In this work, we propose a simulation approach that leverages open data to create a demand model that captures and generalises the usage of this transportation mean in a city. This calls for ingenuity to deal with coarse open data granularity. In particular, we create a flexible, data-driven demand model by using modulated Poisson processes for temporal estimation, and Kernel Density Estimation (KDE) for spatial estimation. We next use this demand model alongside a configurable e-scooter sharing simulator to compare performance of different electric scooter sharing design options, such as the impact of the number of scooters and the cost of managing their charging. We focus on the municipalities of Minneapolis and Louisville which provide large scale open data about e-scooter sharing rides. Our approach let researchers, municipalities and scooter sharing providers to follow a data driven approach to compare and improve the design of e-scooter sharing system in smart cities.**

***Index Terms*—open data, demand model, scooter sharing, electric vehicle, data driven optimization.**

I. Introduction

Urban mobility presents a number of non-trivial challenges both for researchers and regulators. Some of these challenges are related to sustainability and pollution: in EU, for example, urban mobility accounts for 40% of all CO2 emissions of road transport and up to 70% of other pollutants comes from transport systems.[1] The needs to reduce emissions and congestions, along with the rising of the sharing economy, moved several policy makers in promoting micro-mobility services in cities. These services refer to lightweight, often electric-powered vehicles rented for short trips and typically operating at low speeds.

In this context electric scooters (e-scooters) represent a sustainable and cheap alternative to reduce the number of private vehicle trips [1] and consequently traffic congestion [2] and land use [3]. Indeed, e-scooters are among the fastest growing electric micro-mobility means. The number of companies offering e-scooters to rent and the number of cities where the service is available keep growing. Indeed, the e-scooter usage is nowadays comparable with popular car ride-sharing services like Uber and Lyft [4].

Since 2017 several e-scooter companies started their services in many cities in North-America and Europe. Internet-of-Things technologies, paired with accurate GPS tracking, allow the providers to track the position of the e-scooters and monitor users' trips. These data can be used to understand the impact and the utilization of e-scooter in the smart city mobility ecosystem. In this direction, municipalities started offering open data to let other players study alternative solutions.

In this work we are the first - to the best of our knowledge - to study the service sustainability of e-scooters systems from the point of view of a provider. Notice that the peculiarities of this novel scenario call for new approaches (see Section II for a discussion). In this work, we consider the municipalities of Minneapolis and Louisville as use cases.

First, we need to understand how, when and where e-scooters are used by the users, i.e., the mobility demand. For this purpose we rely on open data. Open data typically shares coarsely aggregated data for privacy reasons. This challenges its usage, and calls for ingenuity to appropriately pre-process data with spatio-temporal disaggregation techniques to increase resolution and derive a flexible - albeit realistic - demand model. For this, we combine Poisson processes for customers' arrivals, and Kernel Density Estimate to model the spatial demand [5]. To allow other researchers to reproduce and extend our results, we make our demand models available upon request.

Afterwards, we leverage the constructed demand model to run simulation studies to compare different fleet management policies, with a focus on battery charging strategies. For this we extend our simulator implemented in [6] to support e-scooters scenarios. The simulator allows us to model system parameters such as the operative area granularity, vehicles characteristics, fleet size, users' preferences or fleet management policies. It simulates the search, rental, and return of e-scooters by customers, and the battery consumption and charging operations needed to maintain the fleet. As performance metrics, we mainly focus on satisfied trips, i.e., the fraction of customers' requests that the system can accommodate; and the fleet management cost, proportional to the time workers have to spend to reach and charge the e-scooter battery, assuming a battery swap policy.

The results show that with a spatio-temporal disaggregation coupled with Poisson process and the Kernel Density Estimate we can create a reliable demand model to perform accurate simulations. Our findings show how e-scooter operators should carefully evaluate the best trade-off to balance the users' satisfaction and the fleet management costs. In particular, we

[1]https://ec.europa.eu/transport/themes/urban/urban_mobility_en

show (i) the impact of the size of the fleet, (ii) the impact of the choice of when to swap/charge the batteries, (iii) the implications of using workers or asking the users' cooperation for charging operations.

Results show that the very heterogeneous demand calls for a large number of e-scooters. Similarly the fleet management operations have a high cost due to many battery swap operations. Furthermore, reducing the time for workers to reach the e-scooters and change their batteries has a fundamental impact for reducing cost. Alternatively, directly involving the users in the charging process would further reduce costs, becoming a key design decision.

The paper is organized as follow. In Sec. II we discuss existing works about e-mobility and charging solutions. In Sec. III we describe and characterize the used open datasets. In Sec. IV we introduce the spatio-temporal disaggregation techniques to create our demand models, as well as the simulation assumptions and performance metrics. In Sec. V we show results of our methodology for the cities of Louisville and Minneapolis. Finally in Sec. VI we summarize the paper and present future directions.

II. Related work

Impact of e-scooters in urban mobility is an emerging research topic. The seminal works [7] tested in 2011 the benefits of e-scooters on commuters. Since then, few other studies have tried to gauge the impact of e-scooters on mobility. For instance, authors of [8] present an extensive market analysis emphasizing the possible growth in the usage of e-scooters and raising the problem of how to handle the charging process in presence of large fleets. As a possible solution, authors of [9] propose a model where a MILP formulation clusters together the e-scooters that need to be charged. Similarly, authors of [10] study the benefits of electric fleet (of e-scooters and e-bikes) in last mile delivery for big players in Milan. They are among the first to exploit real data - albeit collected from a very limited deployment (less than 75 vehicles). Authors of [11] offer a first users' habits characterization collecting the daily trips of 38 users, pointing out how the leisure component is relevant for e-scooters. More recent works ([12], [13]) compare micro-mobility services (dockless bike, e-bike and e-scooters) using data exposed by providers. The results confirm that users prefer e-scooters to cover trips shorter than 1.6 km. Moreover the e-scooter daily patterns do not match the commuting patterns. In [14] the authors show that the number of bookings per hour is higher in good weather condition. These characteristics reinforce the need of specific models and tools to study this new type of mobility.

To the best of our knowledge, our work is the first to present a holistic approach to study and compare different system design options, leveraging large open data. We follow a similar approach as in our previous work [15], [16] where we analyze customers' mobility demand patterns in a free floating car sharing system. Here we revisit our methodology in the context of micro-mobility.

Considering shared electric vehicle systems in general, the major challenge is the battery charging process. The battery swap appears the most suitable approach for e-scooters but its study has never been explicitly targeted so far in scientific literature. Early studies about e-buses [17] focus their attention on the management of possible battery switch station and their placement. Other models focus on optimizing the charging process for large vehicles taking into account electric network constraints and system degradation [18], or considering the distance travelled to reach a battery switch station [19]. In a recent work [20], authors optimize battery switch stations considering costs of energy, equipment degradation and energy demand variability.

Few studies analyze the battery swap process applied to shared vehicles. Authors of [21] proposes a mixed integer programming formulation to maximise the satisfied trips in an electric station-based car sharing system, minimizing at the same time the number of battery swaps. Authors of [22] propose an optimal schedule for EV battery swap at stations minimizing travel distance and electrical usage. Differently from our work, all these models do not fit the e-scooter scenario because they do not consider small vehicles and small batteries, hence they do not allow local swap of the batteries.

III. Data collection and characterization

In this section we describe the datasets and characterize the system usage focusing on the most important metrics that would impact the design of an e-scooter sharing system.

A. Dataset description

We focus our study on two cities in the US, namely Louisville and Minneapolis, where their municipalities make available data about all the e-scooter rides performed by the customers using any of the e-scooter sharing providers present in each city.[2] To protect the riders privacy and do not leak any company-specific strategy, data do not contain any identifier of the company, or vehicle, or customer. Furthermore, data is aggregated and/or fuzzed following NACTO guidelines in order to make the user tracking impossible. [3] This challenges the direct usage of the open data, and calls for ingenuity to derive suitable models.

In our cases, each trip exposes information describing the trip *duration*, *distance*, *starting* and *ending* position, and the time when the trip *started*. Different quantisation applies. For Louisville, starting and ending position are encoded with GPS *coordinates* rounded at 3 decimals (approximately 80 m bins); trip duration is given with a precision of one minute, and the starting timestamp is rounded to the closest 15 minutes period. Minneapolis data expose similar information but even more aggregated. Origins and destinations position are defined by *street IDs* so that each trip refers to an entire street length rather than precise coordinates. Timestamps are rounded to the closest 30 minutes period. This rounding are essential to

[2]Datasets are available at: https://data.louisvilleky.gov/dataset/dockless-vehicles, and http://opendata.minneapolismn.gov/search

[3]National Association of City Transportation Officials https://nacto.org/

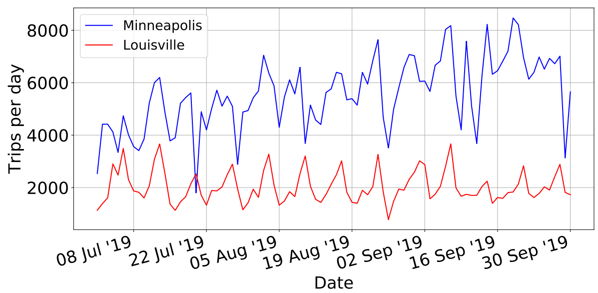


Fig. 1: Time series of trips per day

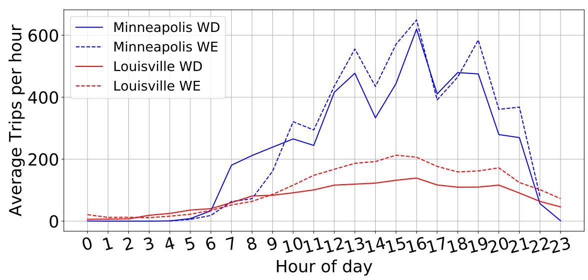


Fig. 2: Average trips per hour in weekends (WE) and working days (WD)

protect the users' privacy, but they complicate the extraction of useful insights from the data. The granularity of rides duration, distance, day and hour of the day still allow us to extract useful patterns about e-scooter usage over time. However, the absence of e-scooter identifier, precise coordinates and timestamps makes impossible to track how each e-scooter moves in the city. Thus we cannot simply reply the same trace in a simulator as done for car sharing services (e.g., in [15]).

B. Dataset characterization

First we provide a data characterization to let understanding the scenarios we are facing. In Fig. 1 we report the number of total recorded trips (i.e., rentals) for each day over the months of July, August and September 2019. More than half a million and 180 k trips have been recorded in Minneapolis and Louisville, respectively. Interestingly, while Louisville shows a repetitive weekly pattern with peaks over weekend but without any specific trend, Minneapolis exhibits an increasing trend. The different number of daily trips justifies the difference in size among the cities, with Minneapolis having more than twice as much the e-scooters in Louisville (see Table I).[4] Some sudden falls are related to bad weather conditions that affects the willingness of customers to rent an e-scooter [14].

To analyze how the demand is distributed during the hours of the day, in Fig. 2 we report the average number of trips per hour per weekday (solid line) and per weekend (dashed line). As expected, Minneapolis exhibits more trip per hour than Louisville. At night we observe a negligible number of trips,

[4]As no vehicle ID is present the maximum number of vehicles is extracted from Louisville service description[2] and Minneapolis official website http://www.minneapolismn.gov/publicworks/trans/WCMSP-212816

TABLE I: Dataset characteristics (Jul. 1st to Sep. 30th, 2019)

City	Trips	Fleet Size	Trip duration [min]		Trip distance [km]		Operative Area
			Avg	Med	Avg	Med	$[km^2]$
Minneapolis	511 k	2 000	13'30"	8'00"	0.95	0.45	268
Louisville	187 k	850	13'50"	7'55"	1.78	1.20	83

with Louisville showing slightly higher figures probably due to a more vivid nightlife. During weekdays we observe a high utilization during central hours of the day (12:00 to 17:00) rather than during commuting hours. This drastically differs from what commonly observed for other shared transportation means like car sharing [15] where utilization peaks during commuting time. Regarding weekends, Louisville confirms the higher utilization with about 30% more trips than during weekdays. This result highlights the importance of a correct characterization of different transportation means usage - which results fundamental to study system design alternatives.

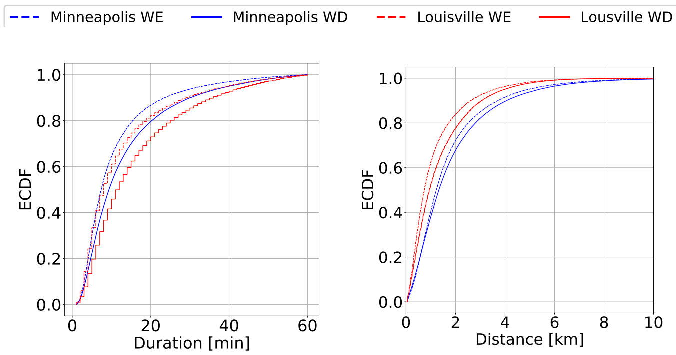
We now focus on the characterization of two important metrics: (i) trip duration, (ii) and trip covered distance. These metrics are fundamental to understand the e-scooter availability and battery discharge properties. Fig. 3a reports the Empirical Cumulative Distribution Functions (ECDFs) of the trip duration for each city during weekdays and weekends. The similarity in the duration is striking, with both Minneapolis and Louisville trips lasting longer during the weekdays than the weekends. Recall that Louisville dataset exposes time duration with a minute granularity which causes the quantisation seen in the ECDF. Overall, trip duration is very short, with the majority of the trips lasting less than 13 minutes. This reflects on the trip distance, as seen in Fig. 3b. Observe that almost 90% the trip lasts less than 4 km, and more than 60% are shorter than 2 km. These results confirms the typical usage of e-scooters [12], [13]. Notice also the different service area size of Minneapolis and Louisville which allows for longer trips in the former. Table I provides a summary of the data.

Considering spatial characterization of the demand we observe that most of trips are confined in few relatively small neighborhoods. Fig. 4 show heatmaps to intuitively gauge this effect. Here, we divide the service areas of each city in 200 m x 200 m cells. Then we count the number of trips originating in each cell during the three months. We use a decimal logarithmic scale. The heatmap shows how concentrated trips are, with few hotter (in red) cells that accounts for 4 orders of magnitude more trips than those cells with few trips (in blue).

Overall, these informations are fundamental to generate a demand model to compare different system designs.

IV. System model and simulator

In this section, we first describe the spatio-temporal disaggregation methodology that we employ to generalize the trips present in the open data. Second we detail how we use them to generate our demand model. Finally we use the demand model to determine the occurring trips and feed our mobility simulator.



(a) Trip duration

(b) Trip distance

Fig. 3: ECDFs for trips duration and distance in weekends (WE) and working days (WD)

A. Spatio-temporal disaggregation

Assume we have a dataset $\mathcal{D}$ of trips recorded during a given period of time. Each trip $i \in \mathcal{D}$ is defined by a discrete start time $a_s(i)$, i.e., with time rounded with a granularity ΔT (of 15 or 30 minutes in our case). To provide an estimation of the time instant in which the trip started, we assume a local stationary process, and simply extract a new timestamp $t_s(i)$ from a uniform distribution in range $\left[a_s(i) - \frac{\Delta T}{2}, a_s(i) + \frac{\Delta T}{2}\right]$. This allows to get back to a continuous-time trace of events.

Considering the spatial information, origin $o(i)$ and destination $d(i)$ positions may be already associated to spatial coordinates, albeit rounded. First, we compute the distribution of distance between $o(i)$ and $d(i)$ which will be useful to generate trip distances later. Second, we obtain, for each $(o(i), d(i))$ pairs, the trip duration from the open data.

Origin and destination information might be aggregated into different *geometries* $o_{id}(i)$ and $d_{id}(i)$. We have to employ a spatial disaggregation methodology to derive possible coordinates. In Minneapolis case, $o_{id}(i)$ and $d_{id}(i)$ are segments representing streets and we randomly select two coordinates along the entire street (with a uniform probability). We obtain thus a possible origin $o(i)$ and destination $d(i)$ coordinates for each trip i.

At the end of this pre-processing step, we have a new disaggregated trace where each trip in the dataset is characterized by its start time, and initial and final coordinates.

B. Demand model

The goal of the demand model is to generalize the trace generated from the original open data. For this, we model the demand in time by using modulated Poisson processes - a common accepted model for i.i.d. service requests of a very large population [23]. For space, we generalize the demand using Kernel Density Estimation (KDE) [5]. KDE gives us the possibility to smooth the point process of a trace over a multi-dimensional space while maintaining the origin/destination correlation.

1) Time Modeling: We assume that the inter-arrival time of trips follows an exponential distribution with rate that depends on the type and hour of the day. To account for the highly periodical rate as seen in Fig.1, here we distinguish between weekday and weekends. We consider 24 time bins of 1 h each (48 periods in total), where the Poisson arrival rate reflects the average rate of requests in the original dataset. This allows to scale the overall demand by introducing a global scaling factor. Not reported here for brevity, we compare the number of trips in the simulated and the disaggregated trace. As expected, there is a very good match (relative percentage residuals for total trips between 0.6% and 1.3% for Louisville, 0.8% and 3.4% for Minneapolis)

2) Spatial Modeling: Given an hour and a day, we want to generate origin and destination of a request according to the specific demand model as exhibited in the disaggregated trace. For this, we leverage KDE to estimate the joint probability distribution of the origin and destination positions of a trip. Given our scenario, this is fundamental to further smooth our discrete events.

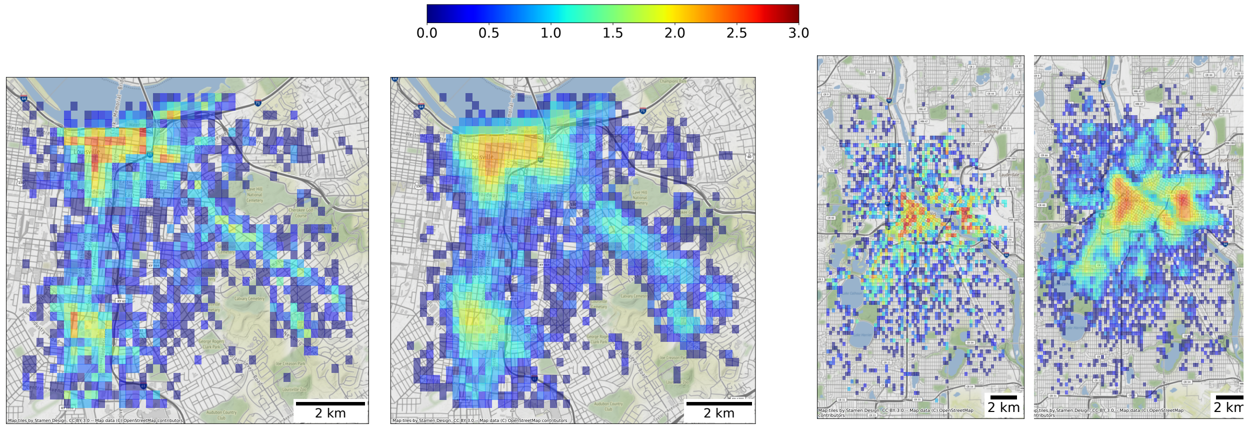
To ease the KDE computation and the simulation process, we divide the whole city area into contiguous squared zones of side 200 m and map the trips to this grid.

Then, for each of the 48 time bins we fit a separate KDE based on the origin-destination zone grid, obtaining a four dimensional problem (2 coordinates for origin and 2 coordinates for destination). In this way, we obtain 48 models summarising the spatial mobility habits of the users in time. Here, we consider a Gaussian kernel [5] and set the bandwidth matrix of the KDE to the 4 x 4 identity matrix. Given the 200 m x 200 m zoning, this corresponds to a bandwidth selection of 200 m for each coordinate. On the one hand a smaller bandwidth would not help us to generalize the demand. On the other hand, a bigger bandwidth would reduce the granularity of city zoning, leading to a reduced precision in incorporating spatial patterns.

In a nutshell, we use KDE as a spatial data smoothing tool, able to capture mobility patterns from the trips in the disaggregated trace while reducing the impact of the original open data aggregation. This is also very effective to cope with the fine grained spatial quantisation that is needed to model the demand of e-scooter sharing systems. To show how effective this is, in Fig. 4 we report the demand in each zone before and after applying the smoothing procedure for Louisville (Fig. 4a) and Minneapolis (Fig. 4b). To ease the readability we report only the demand in the peak hour. Looking at the demand before the smoothing, most of the trips are concentrated in a few areas with large differences also between nearby cells - resulting in a very noisy picture. Most popular zones do not change with the smoothing, but we observe a redistribution of the requests among neighboring zones. In a nutshell, trips are no more concentrated in single cells but rather in larger areas.

C. Mobility and charging simulator

Our goal is to simulate a fleet of e-scooters that move within the city. The simulator uses the demand model to generate



(a) Louisville: demand in open data (left) and in the peak hour model (right).

(b) Minneapolis: demand in open data (left) and in the peak hour model (right).

Fig. 4: Heatmap of the number of trips starting form each zone in a decimal logarithmic scale (the legend report the exponent). The warmer the color, the higher the value

mobility requests. During the simulation we track each e-scooter over time saving information about its location and battery state.

We use an event-based simulator. The simulator has a set $\mathcal{S}$ of e-scooters. At any time t, each e-scooter $s \in \mathcal{S}$ is characterised by its location $P(s,t)$ and state of charge $c(s,t) \in [0, C]$, where C is the maximum battery capacity. As previously, we use a 200 m x 200 m grid. At $t = 0$, e-scooters are placed at random proportionally to the spatial demand, with uniform random charge $c(s, 0) \in [C/2, C]$.

The model generates `trip-request` event i at time t_i according to the Poisson model. It extracts the origin and destination coordinates $\hat{o}(i)$ and $\hat{d}(i)$ from the KDE, and associates the trip duration $\hat{f}(i)$ and distance $\hat{l}(i)$ according the CDF extracted from the original open data. The latter allows us to compute the eventual energy consumption assuming simple proportionality, i.e., $e(i) = k \cdot \hat{l}(i)$. We obtain k from the e-scooter characteristics. When the i-th `trip-request` event fires, the simulator checks if there is any e-scooter s with enough battery $c(s, t_i) \geq e(i)$ available in the same zone or 1-hop neighbors (the 8 adjacent zones in the grid). This is equivalent to assume that customers are willing to rent an e-scooter that is within the same or at neighboring zone from where they are walking at most approximately up to 300m to get it.

If more than one e-scooter exists, the simulator picks s^*, the one having the highest $c(s, t_i)$. It then schedules a `trip-end` event at time $t_i + \hat{f}(i)$. Otherwise, it marks the request as *unsatisfied*. In both cases, it schedules the next `trip-request` event at time $t_i + negexp(\lambda(t_i))$, being $\lambda(t_i)$ the current request rate. When the j-th `trip-end` event fires at time t_j, the simulator picks the e-scooter s^* used for this trip, updates its battery charge $c(s^*, t_j) = c(s^*, t_j) - e(j)$, makes s^* available in position $\hat{d}(j)$, and checks if a charging process is required. That is, it checks if $c(s^*, t_j) < \alpha \cdot C$, being $\alpha \in [0, 1]$ a threshold. If so, it triggers a `charging` event.

The `charging` operation can be performed either by the e-scooter provider through a *battery swap* operation, or by volunteers through *battery charging* operation.

System battery swap: the e-scooter provider manages the charge events by means of a workforce of N worker-equivalent. Battery charge requests are modeled with a FIFO queue, with N parallel servers as follow:

- *Charge request arrival*: If there is a free server, the request gets service immediately. Otherwise, the request gets queued and waits to be processed by a worker.
- *Service time*: the battery swap entails two service operations: *Reach time*, i.e., the time it takes the worker to physically reach the e-scooter; and the *Swap time*, i.e., the time it takes the worker to complete the battery swap operation.

We model the reach time and swap time as negative exponential distributions with average T_{reach} and T_{swap}.

Volunteer charging: We model the possibility that volunteers may contribute to fleet energy management, as done by some companies that remunerate people to handle the charging of e-scooters. When a charge is needed, a volunteer may be found with probability $w \in [0, 1]$. w models people willingness to contribute to the system. If found, we assume the volunteer brings the e-scooter at home and plugs it for charging. We assume the *charging time* to be a Gaussian random variable with average T_{charge} and standard deviation σ_{charge}. The *charging time* is a random variable as it includes the whole process of taking the e-scooter home, charging it, and bringing it back to the streets - in the same location as before for simplicity.

D. Performance metrics

To compare system performance and gauge the impact of parameters, we consider two fundamental metrics:

i) The *Satisfied Demand* measures the percentage of trip requests that can be satisfied due to the presence of e-scooters with enough energy in the trip origin zone.

ii) The *Swap Time* measures the total man-time needed to handle the battery swap operations.

The simulator also breaks down the satisfied demand to distinguish between i) no e-scooter is available, and ii) e-scooters do not have enough energy to complete the trip request. Similarly, it maps events to the city maps to observe the city areas where most of these events occurs.

V. Results

Here we present simulation results obtained starting from the original open data, from which we first generate a dis-aggregated trace, and then extract the trip request model as described above. We use the model to run simulations to gauge the impact of system design choices. In particular we study the impact of:

- $|S|$: the e-scooters fleet size;
- α: the battery threshold that triggers a charging operation;
- N: the provider workflow size;
- T_{reach}: the average time to reach the e-scooter;
- w: volunteers' willingness to handle charging.

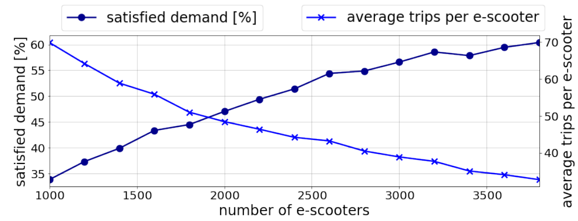
We assume an homogeneous fleet of e-scooters having a $C = 425\,\text{Wh}$ battery capacity and $k = 11\,\text{Wh/km}$ energy efficiency, based on average characteristics present on the market.

A. Impact of fleet size

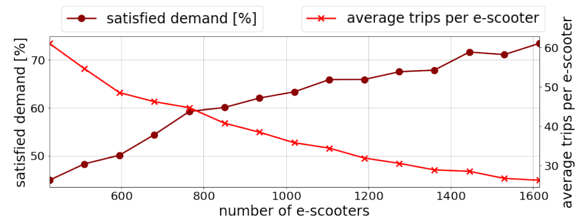
We first evaluate the impact of the fleet size on the satisfied demand. We consider $w = 0$ and $N = |S|$, i.e., system takes care of the charging, with enough workers to immediately perform the battery swap. To consider ideal scenario, we fix $T_{reach} = T_{swap} = 0$. We choose $\alpha = 0.2$ for Louisville and $\alpha = 0.4$ and Minneapolis - so to guarantee their maximum distance trips.

We report results in Fig. 5a and Fig. 5b for Minneapolis and Louisville, respectively. They show the percentage of satisfied demand (left y-axis - blue curves) and the average monthly number of trips performed by each e-scooter (right y-axis - red curve). Fleet size varies around the currently available number of e-scooters - 2 000 in Minneapolis and 850 in Louisville.

The average number of monthly trips per e-scooter decreases with $|S|$, while the bigger the fleet size - the higher the probability to find an e-scooter in the desired origin zone - the higher the percentage of satisfied demand. For Minneapolis, the currently available 2 000 e-scooters can satisfy less than 50% of the demand. Notice the sub-linear growth, hinting that spatial heterogeneity calls for possible relocation policies. For instance, for Louisville results are better, with 60% of satisfied trips with 850 e-scooters. Doubling the fleet size would increase of just about 15% the satisfied demand.



(a) Minneapolis



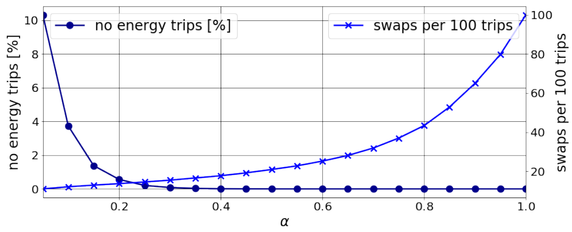
(b) Louisville

Fig. 5: Percentage of satisfied demand and average number of trips per e-scooter per month

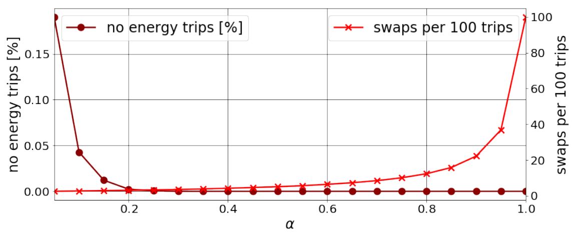
B. Impact of charging threshold

Next, we evaluate the impact of the battery threshold α that triggers charging events. In the one hand, the lower the α, the less frequently e-scooter need to be charged. On the other hand, if α is too low, we may cause users' discomfort and loose revenues as the probability to find an e-scooter with not enough energy would increase. If eventually taken, that e-scooter would suddenly run out of the battery before reaching the desired destination. Here we set $|S| = 2\,000$ for Minneapolis and $|S| = 850$ for Louisville. Again, we assume the ideal charging policy with $T_{reach} = T_{swap} = w = 0$ and $N = |S|$.

Fig. 6 reports the percentage of trips in which the user would run out of battery (left y-axis) and the percentage of trips that require a charging at the end of a trip (right y-axis). The latter represents the charging cost for the system. Starting from this (red curve), observe how the cost linearly increases up to α around 0.5, after which quickly grows to 100%. Indeed, when α approaches 1, every e-scooter needs to be charged at the end of each trip. Looking at the fraction of trips that would not have enough energy to complete them (blue curves), we observe a sudden growth for values of α approaching 0. That is, if we allow the e-scooter battery to reach a very low level, the probability of not completing the trips increase. Minneapolis shows the strongest impact with up 10% of the trips resulting impossible (for $\alpha = 0$). Instead, Louisville exhibits a negligible fraction even for very low α. This is due to the shorter distance than Minneapolis - see Fig. 3b. These results clearly highlight a trade-off between impossible trips (and loss of revenues) and number of charging events (and costs). Our model and simulator allows one to explore this in details.



(a) Minneapolis



(b) Louisville

Fig. 6: Percentage of impossible trips caused by insufficient battery level (left scale) and number of needed battery swaps (right scale) by changing battery swap threshold α

C. Impact of charging policy

We evaluate the cost that the provider faces for the charge operations based on two different charging scenarios: with ($w > 0$) and without ($w = 0$) the users' cooperation. Here we fix $T_{swap} = 5$ min for the operator, and $T_{charge} = 4$ h, and $\sigma_{charge} = 30$ min for the volunteers. We set $T_{charge} = 4$ h and $\sigma_{charge} = 30$ using an average time needed to charge an e-scooter with similar characteristics. Given the ease of the Louisville case with respect to Minneapolis seen in the previous sections, here we just report the case of Minneapolis, with $\alpha = 0.3$. First, we evaluate the cost when the charging operations are performed only by the workers ($w = 0$). For this we run simulations with 2 000 e-scooters and evaluate how many workers are needed to satisfy as much demand as possible. We define a worker as an always available resource (24 hours a day) that perform only one battery swap operation a time. Since we model the time to reach the e-scooter (t_{reach}) as a stochastic variable, we also evaluate its impact in the charging cost.[5]. Intuitively, when few workers are present, or when t_{reach} is too high, an increase in the charging FIFO queue happens, causing e-scooters to be not available and decreasing the satisfied demand.

In Fig. 7 we evaluate the percentage of satisfied demand while increasing the number of workers simultaneously available in the system, with different values of t_{reach}. With small t_{reach} (15 minutes), we can see how with 8 workers we reach the highest satisfied demand as in the best case scenario (Fig. 5a). The increase of the reach time cause a drop in the satisfied demand down to 30% when $t_{reach} = 60$ minutes, even when 14 workers are present. This strong dependence of

[5]Given our policy that only 1 battery swap operation is allowed per event, if two discarded e-scooters are close to each other we consider two reach time

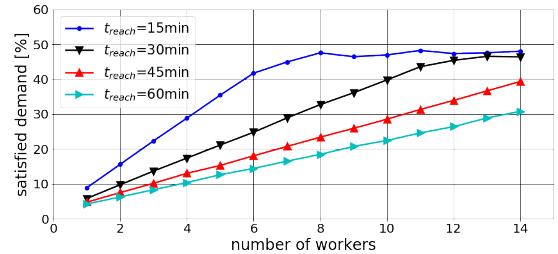


Fig. 7: Minneapolis - percentage of satisfied demand, varying number of workers and average reach time

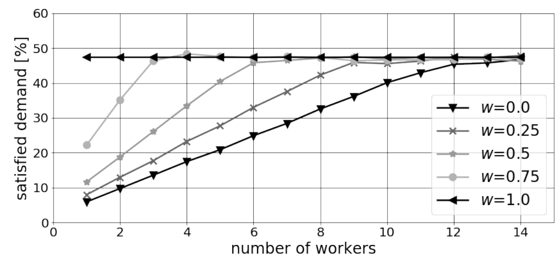


Fig. 8: Minneapolis - percentage of satisfied demand, varying number of workers and users willingness ($t_{reach} = 30$ minutes)

satisfied demand with respect to number of workers suggests to employ strategies to reduce as much as possible t_{reach}. For example, each worker could be assigned to service a limited area of the city.

Finally, we evaluate how the users' help reduces the charging cost for the operator. For this, we consider the same scenario as before, choosing $t_{reach} = 30$ minutes and evaluating different users' willingness (w). Intuitively, the more volunteers help the less workers are needed to perform a battery swap operation. At the same time, due to the longer time for the charge operation by the user, i.e., 4 hours, other effects may appear like a decrease in the satisfied demand due to several e-scooters being under charge at the same time. In Fig. 8 we show the impact on the satisfied demand by changing users' willingness with different number of workers. As a reference we also include the curve with $w = 0$ (same as in Fig. 7). Despite users' recharges are generally longer, there is a limited impact concerning the availability of scooters, and therefore satisfied demand. With a willingness $w = 0.5$ we can see how the number of workers needed to reach the maximum possible feasible trips halves from 12 in Fig. 7 to 6 in Fig. 8. With $w = 1$, the management of the batteries is completely taken care by volunteers. Interestingly, the longer unavailability due to longer charging time has negligible impact on the satisfied demand.

In Fig. 9 we show the total time employed by workers to perform the battery swap on a daily basis. When $w = 1$, workers are not needed - hence the total average daily time is 0 hours. When $w = 0$, there are no volunteers, and the system needs up to 250 hours of cumulative daily work to reach the

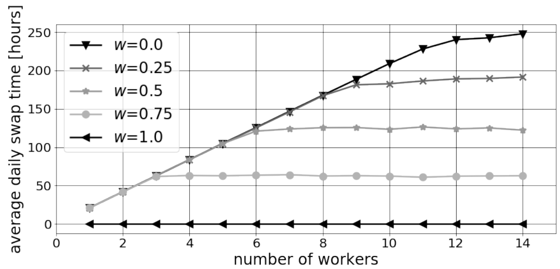


Fig. 9: Minneapolis - Cumulative time needed by the workers to perform the management of the batteries, varying number of workers and users' willingness ($t_{reach} = 30$ minutes)

maximum satisfied demand. w reduces the number of charging events the system has to handle, thus the time spent.

The results show how e-scooter operator should carefully evaluate the best trade-off between using workers or asking the users' cooperation based on its cost for workers and for encouraging users' cooperation.

VI. Conclusion and Future Work

In this paper we proposed a methodology to translate open data describing e-scooter sharing usage into a demand model able to capture and generalize the usage of this transportation mean in a city. We first converted coarse granularity data into a detailed trace. Then, we leveraged modulated Poisson processes and KDE to model the demand over time and space. Thanks to a flexible data-driven simulator, we compared different system design options to evaluate the impact of different e-scooters fleet management strategies.

Our findings show that the design of an e-scooter system asks for different trade-off to balance the users' satisfaction and the management costs. Results show that in order to satisfy the demand and avoid the users to run out of battery we need a large number of e-scooters and battery swap operations. Furthermore, we have analyzed different policies for managing the batteries. Reducing the time for workers to reach the e-scooters and change their batteries has a fundamental impact for reducing cost. Moreover, involving the users to contribute in the charging process might further reduce costs.

We believe that our approach can be useful for researchers, municipalities and e-scooter sharing providers to compare and improve the design of e-scooter sharing system in smart cities. Our ongoing efforts are focused on three directions: (i) extend our methodology to evaluate the detailed economic aspects of the different options, (ii) evaluate how to improve the demand model merging contextual data (as in [24]) and (iii) analyzing possible relocation operations to increase the satisfied demand by using solutions (such as genetic algorithms we used in [25] to improve the users' satisfaction).

References

[1] M. Lee, J. Y. Chow, G. Yoon, and B. Y. He, "Forecasting e-scooter competition with direct and access trips by mode and distance in new york city," *arXiv preprint arXiv:1908.08127*, 2019.

[2] J. Hollingsworth, B. Copeland, and J. X. Johnson, "Are e-scooters polluters? the environmental impacts of shared dockless electric scooters," *Environmental Research Letters*, vol. 14, 2019.

[3] PBOT, "2018 e-scooter findings report," https://www.portlandoregon.gov/transportation/article/709719, 2019.

[4] J. Fong, P. McDermott, and M. Lucchi, "Micro-mobility, e-scooters and implications for higher education," https://upcea.edu/wp-content/uploads/2019/05/UPCEA_Micro_Mobility-White-Paper-May-2019.pdf, 2019.

[5] R. T. García, M. F. Lopez, J. C. Pérez Sánchez, and R. Pérez Sánchez, "The kernel density estimation for the visualization of spatial patterns in urban studies," in *International Multidisciplinary Scientific GeoConference SGEM*, 2015.

[6] A. Ciociola, D. Markudova, L. Vassio, D. Giordano, M. Mellia, and M. Meo, "Impact of charging infrastructures and policies on electric car sharing systems," in *IEEE ITSC*, 2020.

[7] J. D. Bishop, R. T. Doucette, D. Robinson, B. Mills, and M. D. McCulloch, "Investigating the technical, economic and environmental performance of electric vehicles in the real-world: A case study using electric scooters," *Journal of Power Sources*, vol. 196, 2011.

[8] C. Sachs, S. Burandt, S. Mandelj, and R. Mutter, "Assessing the market of light electric vehicles as a potential application for electric in-wheel drives," in *IEEE EDPC*, 2016.

[9] M. Masoud, M. Elhenawy, M. H. Almannaa, S. Q. Liu, S. Glaser, and A. Rakotonirainy, "Heuristic approaches to solve e-scooter assignment problem," *IEEE Access*, vol. 7, 2019.

[10] R. Nocerino, A. Colorni, F. Lia, and A. Lue, "E-bikes and e-scooters for smart logistics: environmental and economic sustainability in pro-e-bike italian pilots," *Transportation research procedia*, vol. 14, 2016.

[11] C. Hardt and K. Bogenberger, "Usage of e-scooters in urban environments," *Transportation Research Procedia*, vol. 37, 2019.

[12] A. Y. Chang, L. Miranda-Moreno, R. Clewlow, and L. Sun. Trend or fad? https://www.sae.org/binaries/content/assets/cm/content/topics/micromobility/sae-micromobility-trend-or-fad-report.pdf. [Online; accessed 24-June-2020].

[13] G. McKenzie, "Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in washington, dc," *Journal of Transport Geography*, vol. 78, 2019.

[14] J. K. Mathew, M. Liu, and D. M. Bullock, "Impact of weather on shared electric scooter utilization," in *IEEE ITSC*, 2019.

[15] M. Cocca, D. Giordano, M. Mellia, and L. Vassio, "Free floating electric car sharing: A data driven approach for system design," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, 2019.

[16] ——, "Data driven optimization of charging station placement for ev free floating car sharing," in *IEEE ITSC*, 2018.

[17] Y. Zheng, Z. Y. Dong, Y. Xu, K. Meng, J. H. Zhao, and J. Qiu, "Electric vehicle battery charging/swap stations in distribution systems: comparison study and optimal planning," *IEEE Transactions on Power Systems*, vol. 29, 2013.

[18] M. R. Sarker, H. Pandžić, and M. A. Ortega-Vazquez, "Optimal operation and services scheduling for an electric vehicle battery swapping station," *IEEE Transactions on Power Systems*, vol. 30, 2014.

[19] X. Zhang and G. Wang, "Optimal dispatch of electric vehicle batteries between battery swapping stations and charging stations," in *IEEE Power and Energy Society General Meeting*, 2016.

[20] G. Battapothula, C. Yammani, and S. Maheswarapu, "Multi-objective optimal scheduling of electric vehicle batteries in battery swapping station," in *IEEE PES ISGT-Europe*, 2016.

[21] E. S. Rigas, S. D. Ramchurn, and N. Bassiliades, "Algorithms for electric vehicle scheduling in mobility-on-demand schemes," in *IEEE ITSC*. IEEE, 2015.

[22] P. You, S. H. Low, W. Tushar, G. Geng, C. Yuen, Z. Yang, and Y. Sun, "Scheduling of ev battery swapping—part i: Centralized solution," *IEEE Transactions on Control of Network Systems*, vol. 5, 2017.

[23] A. K. Menon and Y. Lee, "Predicting short-term public transport demand via inhomogeneous poisson processes," in *ACM CIKM*, 2017.

[24] M. Cocca, D. Teixeira, L. Vassio, M. Mellia, J. M. Almeida, and A. P. Couto da Silva, "On car-sharing usage prediction with open socio-demographic data," *Electronics*, vol. 9, no. 1, p. 72, Jan 2020.

[25] M. Cocca, D. Giordano, M. Mellia, and L. Vassio, "Free floating electric car sharing design: Data driven optimisation," *Pervasive and Mobile Computing*, vol. 55, pp. 59 – 75, 2019.