

Towards Accessible Shared Autonomous Electric Mobility With Dynamic Deadlines

Guang Wang¹, Zhou Qin², Shuai Wang³, *Member, IEEE*, Huijun Sun,
Zheng Dong⁴, *Member, IEEE*, and Desheng Zhang⁵, *Member, IEEE*

Abstract—Shared autonomous electric mobility has attracted significant interest in recent years due to its potential to save energy consumption, enhance mobility accessibility, reduce air pollution, mitigate traffic congestion, etc. Although providing convenient, low-cost, and environmentally-friendly mobility, there are still some roadblocks to achieve efficient shared autonomous electric mobility, e.g., how to enable the accessibility of shared autonomous electric vehicles in time. To overcome these roadblocks, in this article, we design *Safari*, an efficient Shared Autonomous electric vehicle Fleet mAnagement system with joint Repositioning and charging based on dynamic deadlines to improve both user experience and operating profits. Our *Safari* considers not only the highly dynamic user demand for *vehicle repositioning* (i.e., where to relocate) but also many practical factors like the time-varying charging pricing for *charging scheduling* (i.e., where to charge). To perform the two tasks efficiently, in *Safari*, we design a dynamic deadline-based deep reinforcement learning algorithm, which generates dynamic deadlines via usage prediction combined with an error compensation mechanism to adaptively learn the optimal decisions for satisfying highly dynamic and unbalanced user demand in real time. More importantly, we implement and evaluate the *Safari* system with 10-month real-world shared electric vehicle data, and the extensive experimental results show that our *Safari* achieves 100% of accessibility and effectively reduces 26.2% of charging costs and reduces 31.8% of vehicle movements for energy saving with a small runtime overhead at the same time. Furthermore, the results also show *Safari* has a great potential to achieve efficient and accessible shared autonomous electric mobility during its long-term expansion and evolution process.

Index Terms—Shared autonomous electric mobility, accessibility, fleet management, dynamic deadline, deep reinforcement learning

1 INTRODUCTION

SHARED autonomous electric vehicles (SAEV) are believed to be the future mobility [1] due to their convenience and flexibility for use, the potential to reduce the use of privately-owned vehicles for traffic congestion alleviation, as well as the environmentally-friendly nature for gasoline consumption and carbon footprint reduction. In addition, as an additional social benefit, SAEV fleets are the most effective way to introduce the public to electric cars. Hence, they attracted significant interest from both industrial companies and researchers nowadays.

Although being more flexible and cost-efficient for people with low annual vehicle usage, the SAEVs are also facing

many practical fleet management roadblocks due to both the *supply uncertainty* and *demand uncertainty*. From the supply aspect, we need to consider the vehicle availability and parking availability for better user experience and lower operation costs. From the demand aspect, we need to consider the diverse user behavior, unbalanced spatiotemporal usage, as well as the strict timing requirement. Typically, there are two types of tasks for SAEV fleet management to satisfy the future user demand: (i) *vehicle Repositioning/Relocation*, i.e., deciding where to relocate vehicles, which means proactively dispatching unoccupied SAEVs from one service station to another service station, and (ii) *vehicle Charging*, i.e., deciding where to charge vehicles [2], [3], [4], [5], which means dispatching low-battery SAEVs from service stations to charging stations and moving fully-charged SAEVs from charging stations back to service stations.

In recent years, with the development of ubiquitous and mobile computing techniques, there are an increasing number of works focusing on fleet management for different types of vehicles, e.g., taxi [6], bus [7], [8], [9], ridesharing [10], bikesharing [11], e-scooter sharing [12], carsharing [13], but most of these works focused on vehicle repositioning only [14] or charging issue only [15]. Although some recent works [16], [17] have been done to improve the operational efficiency of fleets by considering both vehicle repositioning and charging scheduling, but these mobility modalities may have different usage patterns from SAEV fleets. Thus, many important practical factors (e.g., user behavior and preference, time-variant charging pricing, and strict timing requirements of user demand) may not be

- Guang Wang is with the Department of Computer Science, Florida State University, Tallahassee, FL 32306 USA. E-mail: guang@cs.fsu.edu.
- Zhou Qin and Desheng Zhang are with the Department of Computer Science, Rutgers University, New Brunswick, NJ 08901 USA. E-mail: {zq58, dz220}@cs.rutgers.edu.
- Shuai Wang is with the School of Computer Science & Engineering, Southeast University, Nanjing, Jiangsu 211189, China. E-mail: shuaiwang@seu.edu.cn.
- Huijun Sun is with the School of Traffic and Transportation, Beijing Jiaotong University, Beijing 10044, China. E-mail: hjsun1@bjtu.edu.cn.
- Zheng Dong is with the Department of Computer Science, Wayne State University, Detroit, MI 48202 USA. E-mail: dong@wayne.edu.

Manuscript received 26 April 2022; revised 14 September 2022; accepted 4 October 2022. Date of publication 10 October 2022; date of current version 5 December 2023.

This work was supported by NSF under Grants 1849238, 1932223, 1951890, 1952096, 2003874, 2103604, and 2140346.

(Corresponding author: Guang Wang.)

Digital Object Identifier no. 10.1109/TMC.2022.3213125

captured by them. More importantly, most of existing works set predefined static schedules for vehicle repositioning or charging scheduling (e.g., set 10 minutes as a time slot to make the management decisions [10], [13], [18]), which potentially lead to two key drawbacks: (i) it is challenging for them to fully satisfy the user demand at intensive and unexpected peaks, so some users cannot have access to the mobility system in time; (ii) it may cause unnecessary decision making during low demand time periods, leading to more energy consumption. Furthermore, most existing optimization methods [16], [19] only considered short-term performance of the fleet management without farsighted views, which potentially causes suboptimal solutions.

To advance existing works, in this paper, we design *Safari*, a new data-driven fleet management system with joint repositioning and charging scheduling for SAEVs to enable shared autonomous mobility accessibility and also improve its operational efficiency, which means whenever users have travel demand, they can have access to the shared vehicles in time. However, it is very challenging to achieve an accessible SAEV fleet considering both vehicle repositioning and charging due to *possible conflicting relationships* (e.g., meet the accessibility of users versus improve the profit of the operator) and many *confounding factors* (e.g., individual user behaviors like spatiotemporal usage preference, time-variant charging pricing, the availability and reachability of SAEVs).

To address these challenges, in *Safari*, we design a *Dynamic Deadline-based Deep Reinforcement Learning* algorithm to learn sophisticated decisions, which has two key advantages for SAEV fleet management: (i) the dynamic deadline strategy helps the system to adaptively satisfy the time-varying unbalanced user demand based on usage prediction and an error compensation mechanism; (ii) the dynamic deadline deep reinforcement learning-based decision making enables a long-term benefit of the system and improves scalability at the cost of less coordination between SAEVs, which causes a small runtime overhead for the real-time requirement.

In particular, the key contributions of this paper include:

- In this paper, we design the first dynamic deadline-based SAEV fleet management system called *Safari*, which integrates two categories of decision-making tasks to guarantee the SAEV accessibility of users: (i) vehicle repositioning (proactively move unoccupied vehicles in one service station to another service station) and (ii) vehicle charging (move low-battery vehicles from service stations to charging stations and move fully-charged vehicles from charging stations back to service stations). *Safari* considers different real-world factors for decision-making, e.g., dynamic demand & supply, availability of SAEVs, reachability to other services stations or charging stations, time-variant charging pricing, etc.
- In *Safari*, we design a dynamic deadline-based deep reinforcement learning algorithm to learn which service station to relocate and which charging station to charge for each SAEV. It has two major components: (i) A prediction-based dynamic deadline mechanism is utilized to adapt to the highly

dynamic demand and supply, where the prediction is performed based on the features we capture from our data-driven observations, and an error compensation mechanism is introduced to make our *Safari* more robust to the prediction error. (ii) Based on the dynamic deadline setting, a deep reinforcement learning module is presented, which enables long-term benefits for the system with a small runtime overhead, and it has the potential to make our system more sustainable.

- More importantly, we implement and evaluate our dynamic deadline-based real-time fleet management system *Safari* for repositioning and charging of SAEVs with multi-source data, including 10-month detailed order records from over 12,000 unique users and the metadata of stations from a shared EV fleet. The experimental results show our *Safari* achieves 100% of accessibility and also effectively reduces 26.2% of the charging costs and reduces 31.8% of vehicle movements. It also has a great potential for efficient fleet management of SAEVs during their long-term expansion and evolution process.

The rest of the paper is organized as follows. Section 2 summarizes related works. Section 3 introduces the SAEV system and its management significance. Section 4 shows the key idea and problem formulation. Section 5 presents the detailed system design. Section 6 evaluates the performance of our *Safari*. The lessons learned are discussed in Section 7. Finally, we conclude this paper in Section 8.

2 RELATED WORK

In this section, we summarize two types of related works, i.e., vehicle repositioning and EV charging.

2.1 Vehicle Repositioning

Owing to the availability of the rich vehicle location information and operation log data, there is a surge number of work on addressing the unbalanced demand and supply problem by vehicle repositioning for different mobility modes, e.g., taxi [20], [21], [22], ridesharing [10], [18], bike-sharing [11], [23], [24], e-scooter sharing [12], and carsharing [25]. *Taxi*: Yuan et al. [26] proposed a recommendation system to help taxi drivers find some locations where they are more likely to pick up passengers quickly and maximize the profit. Zhou et al. [20] presented a spatial network-based Markov Decision Process to recommend taxi drivers to specific regions to maximize the profit in the near future. *Ridesharing*: Lin et al. [18] designed a management system for ridesharing platforms to maximize the gross merchandise volume of the platforms by relocating available vehicles to the locations with a larger demand-supply gap than the current one. Miller et al. [27] presented a positioning method for improving customer quality of service in ridesharing systems. *Bikesharing*: Liu et al. [23] formulated the station-based bikesharing repositioning problem as mixed integer nonlinear programming to minimize the total travel distance.

Different from existing vehicle repositioning work, in this paper, we study an innovative transportation modality, i.e., *electric carsharing*, which has essential differences with

existing mobility modes from both spatial and temporal dimensions. In addition, different fleets have different management modes, so they have different repositioning mechanisms. Moreover, existing vehicle repositioning works rarely consider complicated charging issues and potential impacts of dynamic deadlines.

Even though there are also some works focusing on car-sharing vehicle repositioning, most of them are based on theoretical models with many assumptions. For example, Jorge et al. [28] presented a new mathematical model to optimize the relocation operations that maximize the profitability of a carsharing service and a simulation model to study different real-time relocation policies. Due to the data-driven nature of our work, which reveals some real-world insights for fleet management, and we also verify our algorithm using the real-world data.

2.2 EV Charging Recommendation

In the recent decade, there is also an increase of works on EV charging recommendation [6], [17], [29], [30], [31], [32], [33], which is also related to our work. Wang et al. [17] designed a charging recommendation system to learn the charging policy for reducing the range anxiety of e-taxi drivers. Yuan et al. [34] developed a proactive partial charging recommendation strategy to increase EV utilization. Wang et al. [6] presented a data-driven fairness-aware charging recommendation method to reduce charging idle time of e-taxi fleets. Zhang et al. [35] proposed a framework for intelligently recommending public accessible charging stations to EVs by jointly considering various long-term spatiotemporal factors. Tian et al. [30] presented a real-time recommendation system to reduce e-taxi drivers' charging costs by mining large-scale GPS data and taxis operation data information. Zhou et al. [5] developed a strategy to assign SAEVs to charging stations for maximizing user satisfactions in terms of traveling distance under the budget of solar energy income.

Uniqueness: Different from these works, our paper focuses on a new type of mobility, i.e., shared autonomous electric mobility. One of the most important factor for the success of shared autonomous electric mobility is the accessibility of users, which means whenever users have travel demand, they can have access to the SAEVs in time. Therefore, SAEVs need to be relocated between different service stations to satisfy the highly dynamic and unbalanced spatiotemporal usage distributions. The interactions between service stations and charging stations also make it challenging to manage SAEV fleets efficiently compared to only charge other types of EVs. More importantly, we designed a dynamic deadline-based fleet management strategy, which has not been presented by existing works.

3 PRELIMINARY

3.1 SAEV System and Operation

A typical SAEV fleet operation paradigm is shown in Fig. 1. There are four main parties in the system, i.e., operator, users, SAEVs, and stations. The operator provides a fleet management system to monitor all real-time status information of users, SAEVs, and stations, and performs spatial search. In particular, the real-time location information and

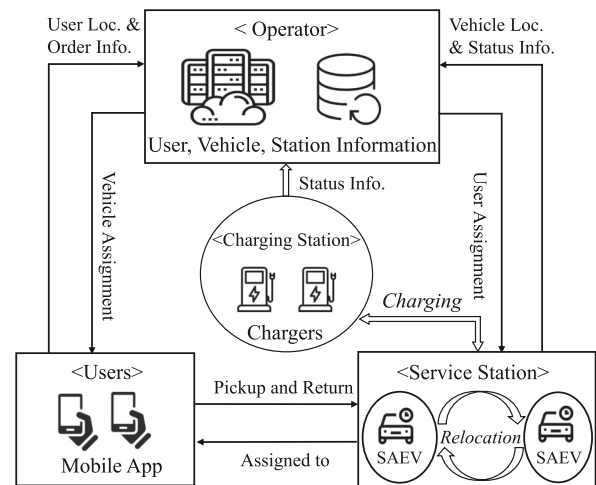


Fig. 1. Shared autonomous EV fleet operation paradigm.

order information of users are recorded and uploaded when they use the mobile app. The real-time location and status information of SAEVs are also periodically uploaded to our servers via communication devices. The transaction information is recorded when users return SAEVs.

In light of the description of this SAEV system, a usage procedure typically includes four steps:

- 1) *Registration.* New users register for using SAEVs via a mobile app. Their demographic information (e.g., age, gender, and occupation) is required when registering for verifying and security purposes.
- 2) *Pickup.* A user send a request via the app. A SAEV will be dispatched to pick the user up once receiving the request from the user. The SAEVs typically have enough energy for the users to use.
- 3) *Using.* Users utilize SAEVs for their purposes.
- 4) *Returning.* Users return SAEVs, and then make the payment via the app. Then the SAEVs will be dispatched to the nearest service station with available parking spots if no other requests.

Ideally, if there are an unlimited number of SAEVs, user demand can be satisfied trivially. However, this assumption is normally not realistic. The operator usually possesses a limited number of SAEVs and parking spots at each service station due to high costs. Thus, some efforts are needed to balance the demand and supply to guarantee the accessibility, which includes two tasks: relocating SAEVs between service stations and driving low-battery SAEVs to charge in charging stations and then distributing fully-charged SAEVs back to service stations for satisfying the future demand. Intuitively, given the unbalanced user demand and supply, *how to effectively guarantee the accessibility of users and decide the optimal locations for SAEVs to relocate and charge* is essential to improve the user experience and increase operating profits for the SAEV operators.

3.2 Problem Significance

Definition 1. *Deadline:* Typically, deadlines in a real time system represent the time at which some specific tasks need to be finished [36]. As shown in Fig. 2, suppose a long time period T is divided into h consecutive intervals, e.g., $\{I_1, I_2, \dots, I_h\} =$

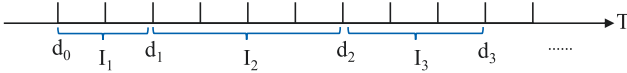


Fig. 2. An illustration of the dynamic deadline.

$\{(d_0, d_1], (d_1, d_2], \dots, (d_{i-2}, d_{i-1}], (d_{i-1}, d_i], \dots, (d_{h-1}, d_h]\}$, where d_0 is the start time of the time period and d_i is the deadline for decision completion. The $(i-1)$ th deadline is the time when we need to finish all vehicle repositioning and charging scheduling to satisfy the user demand by the (i) th deadline. In other word, the vehicle repositioning and charging scheduling performed between the $(i-2)$ th and $(i-1)$ th deadlines will satisfy the user demand arising during the period between the $(i-1)$ th and (i) th deadlines. The dynamic deadline means d_i is not predefined and static, which needs to be learned from the real-world demand and supply and usages. The dynamic deadline generation method will be introduced in Section 5.

In the shared autonomous electric mobility, user experience is directly impacted by the availability of SAEVs. Intuitively, users suffer from bad user experience if they cannot have access to the SAEVs in time when they have travel demand, and operators suffer from low operational efficiency if they frequently relocate the SAEVs because it will cause more energy consumption. Especially, those situations always occur during the rush hours and valley hours if the SAEVs are not managed effectively. Thus, to guarantee good user experience and high operational efficiency, the dynamic usage behaviors need to be taken into consideration. On the one hand, how many SAEVs are picked up and returned at each time slot determines which service station will be jammed or starved, thus impacts how to perform vehicle repositioning and charging scheduling; on the other hand, how to conduct the vehicle dispatching impacts how many available SAEVs and parking spots in each service station, which impacts the future mobility accessibility of users.

Therefore, how to make decisions by considering the future user demand for good user experience and high operational efficiency is a key task of fleet management, which inspires us to (i) characterize the user pickup & return demand using dynamic service deadlines; (ii) design a practical and efficient decision making strategy to improve the user experience and optimize the operating profit while all the deadlines can be met.

4 KEY IDEA AND PROBLEM FORMULATION

4.1 Key Idea of Safari

In this paper, we design Safari, a new data-driven fleet management system for SAEVs to guarantee mobility *accessibility* (i.e., whenever users have travel demand, they can have access to the shared vehicles in time) and improve the operator's *operating profits* (which are highly related to the revenue of the fleet from serving users, costs for charging, and the extra movements of vehicle dispatching) based on joint repositioning and charging scheduling with dynamic deadlines. In addition, Safari also considers different complicated real-world factors.

Fig. 3 shows the framework of our designed Safari, which includes four key modules, i.e., data module, net flow prediction module, prediction-based dynamic deadline generation module, and real-time decision making module by

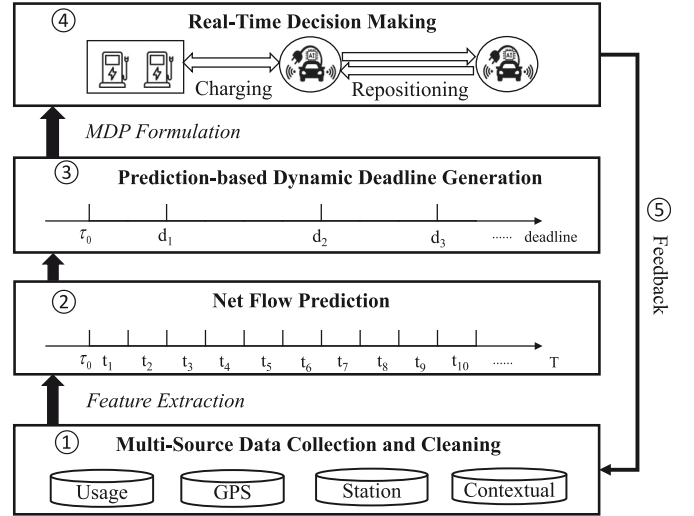


Fig. 3. Framework of the Safari.

considering both vehicle repositioning and charging of SAEVs. (i) We first collect the real-world shared autonomous mobility data and contextual data (e.g., weather and temperature). (ii) Based on these data, we extract features for net flow prediction. (iii) Then we generate the dynamic deadlines based on the net flow prediction and error compensation. (iv) We make decisions for real-time repositioning and charging in order to improve the shared autonomous electric mobility services. (v) It then provides feedback and generates new data to feed to the system. These five steps form a closed loop to constantly improve the shared autonomous electric mobility services.

Due to the sequential decision characteristics of SAEV fleet management, we formulate the problem as a Markov Decision Process (MDP), and then we present a dynamic deadline-based distributed deep reinforcement learning method to achieve our goals, which not only makes the system adaptively satisfy the time-varying unbalanced pickup and return demand based on usage prediction, but also achieves long-term benefits with a small runtime overhead for the real-time requirement.

4.2 Problem Formulation

Formally, we model the SAEV fleet management problem as an MDP $\mathcal{G}$ for N agents, which is defined by a five-tuple $\mathcal{G} = (\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \gamma)$, where $\mathcal{S}$ is the set of states; $\mathcal{A}$ is the action space; $\mathcal{R}$ is the reward function; $\mathcal{P}$ is transition probability functions; and γ is a discount factor. In an MDP, an agent behaves in an environment according to a policy that specifies how the agent selects actions at each state of the MDP. The detailed formulation of the MDP $\mathcal{G}$ in our problem is shown below.

Agent. We consider each unoccupied SAEVs (i.e., does not rent by users) as an agent, and only the unoccupied SAEVs can be scheduled by our system. Although the number of total agents in the fleet is always N , the number of agents in each time interval N_t is changing over time.

State S. The state of a SAEV is defined as a two-dimensional vector indicating its spatiotemporal status. Suppose there are a set of service stations $\{SS\}$ and a set of charging stations $\{CS\}$, so each unoccupied SAEV may be in one of

the service stations or charging stations. We define a local-view state of a SAEV, $s_{t,lo} = [t, l] \in \mathcal{S}_{lo}$, where t is the time index and $t \in ((d-1)th, dth]$, i.e., the time slot between the $(d-1)th$ and the dth deadline, and $l \in \{SS\} \cup \{CS\}$ is the location index (i.e., which service station or charging station the SAEV is in). In this case, the finite local state space $\mathcal{S}_{lo}$ is a Cartesian product of the set of deadlines and the set of service stations + charging stations, i.e., $\mathcal{S}_{lo} = \{D\} \times (\{SS\} \cup \{CS\})$ and the number of states is $|\mathcal{S}_{lo}| = |D| \times |\{SS\} \cup \{CS\}|$. In addition to the local-view state, we also define a global state $s_{t,go}$ to capture the system status includes the number of SAEVs and parking spots availability at each service station and their real-time predicted usages in the next interval, and it also includes the number of unavailable SAEVs at each station in the current and next time interval. The global-view state $s_{t,go}$ will update in each time slot. Finally, the state of each available SAEV k during the time slot t can be represented as $s_t(k) = [s_{t,lo}(k), s_{t,go}(k)] \in \mathcal{S}(k)$.

Action $\mathcal{A}$. The action space of a SAEV k , $\mathcal{A}(k)$ specifies where it should go by the next deadline. We define five types of actions for the SAEV scheduling. (i) $\mathcal{A}_S$: Staying at the current service station; (ii) $\mathcal{A}_R$: Relocating to another service station to satisfy user demand in that station or make parking spots for vehicles that will be returned to this service station; (iii) $\mathcal{A}_C$: scheduling to Charge in a charging station; (iv) $\mathcal{A}_K$: Keeping charging at the charging station; (v) $\mathcal{A}_B$: moving Back to a service station;

The action to take is decided by two factors: (i) *Availability* of SAEVs for users in each service station by deadlines. (ii) *Reachability* to the charging stations or other service stations of SAEVs when performing repositioning or charging scheduling. That means for each service station $s_i \in \{SS\}$: (1) # of pickups $\leq$ # of available SAEVs; (2) # of SAEVs (available or unavailable) + # of returns $\leq$ # of parking spots. Besides, for each SAEV k , it becomes unavailable if its battery level decreases to below a threshold value η (e.g., 30%), which means the low-battery SAEVs should be dispatched/moved to charge in order to satisfy future demand. Since there are enough public charging stations in Beijing, we envision that the number of charging points is sufficient for SAEVs, and SAEVs can always be charged in the nearest available charging stations.

Reward $\mathcal{R}$. Reward usually determines the optimization goal and reflects the immediate sense of the action in a specific state. A typical measurement is to estimate the difference of the accumulated reward between with and without an action. We define three types of immediate rewards here, i.e., positive reward, zero reward, and negative reward, which capture the money transaction.

Specifically, (i) if a SAEV is picked up by a user during certain interval (i.e., $\mathcal{A}_P$), it will have a positive reward, which is equivalent to the money paid by the user; (ii) if a SAEV stays at the current service station (i.e., $\mathcal{A}_S$), it will not be used by users (i.e., no revenue) and also have no charging and moving costs, so the immediate reward is zero; (iii) if a SAEV is relocated from one service station to another service station (i.e., $\mathcal{A}_R$), scheduled from one service station to a charging station (i.e., $\mathcal{A}_C$), or moved back to a service station from a charging station (i.e., $\mathcal{A}_B$), it will have a negative reward due to the energy consumption cost. (iv) if a SAEV is charging in a charging station (i.e., $\mathcal{A}_K$), it will also have a negative reward due to

the charging cost, which is explicitly related to the charging time and charging prices. Implicitly, the charging cost is also related to the previous usages and repositionings since these activities will directly cause the energy consumption of SAEVs. Hence, we define the reward function as follows

$$R_U - C_C - C_D = \sum_{i=1}^m R_u^{(i)} - \sum_{j=1}^n \left(\lambda \cdot T_c^{(j)} \right) - \sum_{k=1}^z \eta \times d_k, \quad (1)$$

where R_U is the total revenue from serving users; C_C is the total charging cost of the electric carsharing fleet; C_D is the total energy consumption cost for dispatching SAEVs, e.g., relocating SAEVs from one service station to another service station and moving SAEVs to charging stations and moving back to services stations; $R_u^{(i)}$ is the revenue from serving i th electric carsharing order; m is the total number of served orders; n is the total number of charges of the fleet. $T_c^{(j)}$ is a three-dimensional vector $T_c^{(j)} = [T_p^{(j)}, T_f^{(j)}, T_o^{(j)}]$ describing the charging time of j th charging event, where $T_p^{(j)}$, $T_f^{(j)}$, and $T_o^{(j)}$ denote the time in peak, flat, and off-peak charging pricing hours of the j th charging event, respectively. Similarly, we also describe the time-varying charging pricing as a three-dimensional vector $\lambda = [\lambda_p, \lambda_f, \lambda_o]$, where λ_p , λ_f , λ_o denote the charging prices during peak, flat, and off-peak hours, respectively (as shown in Fig. 7). η is the energy consumption rate (i.e., energy consumption per kilometer), and we intuitively envision it is a constant value for simplification; d_k is the distance of k th movement; and z is the total number of vehicle movements of SAEVs, so reducing the number of vehicle movements will potentially increase the profit of the SAEV fleet.

Hence, if we can satisfy more passenger demand, then the total revenue from serving users R_U can be increased. We can also improve the profit of the operator by reducing its charging cost C_C for the SAEV fleet and movement cost for dispatching SAEVs C_D .

Probability function $\mathcal{P}$ defines the transition probability between states by taking action $\mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$, e.g., $p(s_{t+1}|s_t, a_t)$ denotes the probability of transition to the next state s_{t+1} given the action a_t in the current state s_t . Our goal is to find a function that maps a state to the best action that each SAEV can take.

Discount factor γ essentially determines how much the agents care about rewards in the distant future relative to those in the immediate future. The value of γ is typically selected from $[0, 1]$, so the final expected reward in the infinite horizon will be convergent and bounded to a finite number. If $\gamma = 0$, the agent will be completely myopic and only learn about actions that produce an immediate reward without considering future rewards.

5 SYSTEM DESIGN

From Fig. 3, we can see that Safari has four key component. In this section, we show the detailed design of each component of Safari.

5.1 Multi-Source Data Module

In this paper, we utilize a real-world shared EV fleet operation dataset collected from the Chinese city Beijing. This

TABLE 1
An Example of the Dataset

Usage Data	Vehicle ID	Order Number	Pickup Time	Pickup Station	Usage Time (min)	Payment (CNY)
	300044kzbch7	1708XXX2831	2017-08-01 00:28:01	Marriott Center	55	19.4
GPS Data	Device ID	Time	Longitude	Latitude	Speed (km/h)	Vehicle ID
	116232100002579	2017-08-01 20:56:41	116.4062079	40.050016	25.0	10001g41q5vn
Service Station	Station ID	Station Name	Longitude	Latitude	# of Parking Spots	Status
	66	Jiajing Tiancheng	116.483295	39.986183	3	Operating
Charging Station	Station ID	Station Name	Longitude	Latitude	# of Chargers	Status
	1	CS1	116.225523	39.909251	10	Operating

dataset is obtained by collaborating with a shared EV operator, who provides us the data for improving its business intelligence. The shared EV fleet is similar to SAEV fleet except for having drivers or not. The time span of the dataset is from January 2017 to October 2017. The dataset includes four different types of data, i.e., vehicle usage data, trajectory data, service station and charging station data. Some key fields of the four types of data are shown below, and an example is in Table 1.

- *Vehicle Usage Data* include all user travel demand records. If a user has successfully accessed the shared EVs, there is a detailed usage record consisting of 26 fields describing vehicles, users, and usage-related information, e.g., the order number, the user ID, user age, gender, workplace & occupation, order time, vehicle pickup and return time and station, the vehicle ID, usage time, and payment, etc. If a user has not successfully accessed the shared EVs, it will have an incomplete record indicating the failure. In our dataset, about 87% of user demand was satisfied by the current system.
- *Vehicle GPS Data* include fields that describe the real-time status of each shared EV, e.g., vehicle ID, time-stamp, and longitude & latitude. The GPS data is collected from users' cellphones, which is uploaded to our server through 4 G/5 G cellular communication. The GPS data is only collected when users are using the shared EVs with the mobile app.
- *Service Station Data* describe the service station information, e.g., the station IDs, the station names, the GPS locations, and the number of parking spots.
- *Charging Station Data* include station IDs, station names, coordinates (i.e., longitudes and latitudes), and the number of charging points in each station. The charging station data is collected and uploaded with CAN bus communication and GPRS.

Based on the multi-source dataset, we conducted a comprehensive data cleaning and data-driven analysis to discover the factors that may impact user behavior. Due to the large size of our shared EV data, it requires significant efforts for efficient management, querying, and processing. Hence, we performed a detailed cleaning process to filter out the error, duplicate, and incomplete order and vehicle GPS data on a high-performance cluster with Spark and Hadoop, which was equipped with 80 TB memory and 20 nodes.

5.1.1 Per-Usage Spatiotemporal Patterns

Figs. 4 and 5 show the per-usage time and distance distributions. We found that about 50% of the per-usage time is longer than 5 hours and 88% of the per-usage time is shorter than 16 hours. There are also some users who use shared EVs for a long time, e.g., about 4% of the per-usage time is longer than 4 days. We found most users use shared EVs for middle-distance trips, e.g., about 75% of trips are longer than 20 km and 80% of trips are shorter than 80 km.

5.1.2 Unbalanced Spatiotemporal Usage Patterns

We then further investigate the fine-grained pickup and return distributions of different service stations at different hours of a day. As shown in Fig. 6, the red circles mean there are more pickups than returns in these service stations, and the aquamarine circles mean there are more returns in these service stations. The size of each circle stands for the absolute value of the difference between the number of returns and the number of pickups, i.e., $|\# \text{ of returns} - \# \text{ of pickups}|$.

We found the pickup and return patterns have significant differences between different hours of a day. For example, there are more vehicle returns in most service stations during late-night hours (e.g., 0:00-1:00), and both the number of pickups and returns are small. In the morning rush

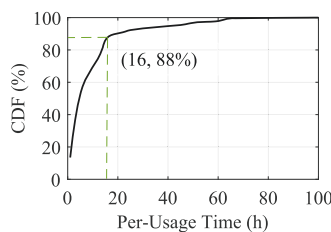


Fig. 4. Per-usage time.

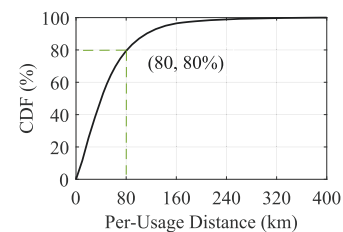


Fig. 5. Per-usage distance.

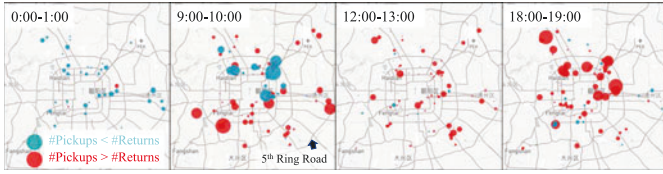


Fig. 6. Spatiotemporal unbalanced usage patterns.

hours (e.g., 9:00-10:00), there are very high pickup demand and return demand in different areas, e.g., more pickups in residential areas and more returns in IT industrial areas. The number of pickups increases during the day time, and it peaks in the evening rush hours (e.g., 18:00-19:00). To meet users' highly dynamic pickup and return demand, there should be enough shared EVs for users to pick up and enough nearby parking spots to return vehicles in each service station. However, this is not easy to be achieved by users themselves, and it is necessary for the operator to dispatch shared EVs between service stations and move low-battery vehicles to charge so that they can be accessed by users in time.

5.1.3 Time-Variant Charging Pricing

We found many cities have time-variant charging pricing, which breaks up 24 hours of a day into several intervals and charges a different price for each interval. For example, the charging rates in Beijing are divided into three types, i.e., off-peak prices (low rates), flat prices (medium rates), and peak prices (high rates), and the corresponding electricity rates are 1.1946, 1.4950, and 1.8044 CNY/kWh, respectively [37]. The time-variant charging pricing of Beijing is shown in Fig. 7. We found the peak price is 51% higher than the off-peak price, which means the charging costs can potentially reduce 51% if the operator charges SAEVs in off-peak charging pricing hours instead of peak charging pricing hours. Hence, the charging prices should also be considered for the SAEV fleet management.

5.2 Data-Driven Prediction Module

We first define the *net flow* and *accumulated net flow* as follows.

Definition 2. We divide a long time period T into a set of small time slots, e.g., we set the time slot length as 5 minutes to capture the more fine-grained pickup and return patterns. We then define the *net flow* of a service station in a time slot as the number of returns (i.e., inflow) minus the number of pickups (i.e., outflow) in this time slot.

For each time slot t_i , we calculate the net flow f_{t_i} of a service station in this time slot based on the real-world order

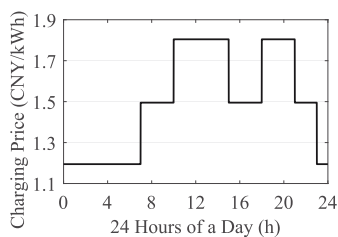


Fig. 7. Time-variant charging pricing in Beijing.

Authorized licensed use limited to: Florida State University. Downloaded on January 04, 2024 at 02:57:36 UTC from IEEE Xplore. Restrictions apply.

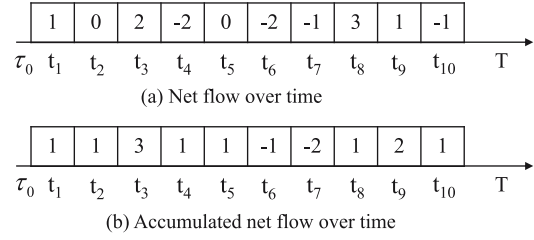


Fig. 8. The (accumulated) net flow of a service station.

records, i.e., the net flow value will be deducted by one if there is a pickup activity and the net flow value will be added by one if there is a return activity.

Fig. 8a shows an example of the calculation of the *net flow* of a service station. In this example, there are 10 small time slots, and the value in each time slot denotes the net flow in this time slot, e.g., the net flow of the service station is 1 in t_1 and -2 in t_4 , which means there are one more returns than pickups in t_1 and there are two more pickups than returns in t_4 . Based on the net flow f_{t_i} of the service station in each time slot, we then calculate the accumulated net flow in multiple time slots, which is defined as follows:

Definition 3. Given a series of time slots $\{t_1, t_2, \dots, t_m\}$, the accumulated net flow by t_i of a service station is defined as the sum of all net flows of previous time slots, i.e., $F_{t_i} = \sum_{j=1}^i f_{t_j}$.

Fig. 8b shows the corresponding accumulated net flow of Fig. 8a. Each value F_{t_i} in time slot t_i is the sum of the net flow from time slot t_1 to t_i . For example, the accumulated net flow in t_3 of the service station is $F_{t_3} = f_{t_1} + f_{t_2} + f_{t_3} = 3$.

Accurate *net flow* predictions $f_{t_i}(i)$ are very important as they directly impact the following dynamic deadline generation and decision making process. Hence, in this paper, we comprehensively consider the factors that may impact users' SAEV usage. Finally, we extracted five categories of basic features (i.e., temporal features, spatial features, historical usage features, user demographic features, and contextual features) that are highly related to users' usage behaviors to predict the *net flow* of each service station in a small time slot more accurately, which include the following 12 features.

5.2.1 Temporal Features $\mathcal{F}_T$

From Fig. 6, we found the number of pickups and returns are closely related to the time factor. Hence, we extract three temporal features: the time of a day (e.g., we divide one day into a set of time slots $\mathcal{F}_{\text{todr}}$ and each time is set to be 5 minutes for a fine-grained prediction), the day of a week $\mathcal{F}_{\text{dowr}}$ and holiday $\mathcal{F}_{\text{hol}}$.

5.2.2 Spatial Features $\mathcal{F}_S$

We found users have different purposes for using shared EVs, which result in different spatial usage patterns, e.g., people who use SAEVs for commuting may pick vehicles in residential areas and then return them in the industrial areas. Hence, we divide the city into seven categories of functional areas based on the method in [38] (i.e., residence, entertainment, business, industry, education, scenery spot, and suburb) to capture the spatial patterns, which forms the functional area feature $\mathcal{F}_{\text{area}}$.

5.2.3 Historical Usage Features $\mathcal{F}_H$

Since the shared EV usages show a weekly pattern, we can consider that the pickups/returns in a time slot are related to the pickups/returns in the same time slot of previous weeks. Hence, we utilize our long-term shared EV operation data to capture the historical usage patterns. We extract the *net flow* of each service station in the same time slot of three previous consecutive weeks (e.g., $\mathcal{F}_{his1}$, $\mathcal{F}_{his2}$, $\mathcal{F}_{his3}$) as the historical usage features.

5.2.4 User Demographic Features $\mathcal{F}_D$

We found that the users who use shared EVs in different regions have different demographic features (e.g., more young male users use shared EVs in areas with many IT companies, and more young female users use shared EVs in areas with many financial companies), so they may have different usage patterns due to their job characteristics. Therefore, the user demographic features are also important for the *net flow* prediction. Finally, we extract the percentage of male users and female users as the gender feature $\mathcal{F}_{gend}$ of each service station, and we utilize the users' average age as the age feature $\mathcal{F}_{age}$ of each service station.

5.2.5 Contextual Features $\mathcal{F}_C$

We also found the contextual features like weather conditions have a great impact on users' shared EV usage behaviors. Hence, we collect meteorology data from the website [39] and extract features for the *net flow* prediction. We identify three contextual features: weather $\mathcal{F}_{wea}$, temperature $\mathcal{F}_{tem}$, and wind speed $\mathcal{F}_{wind}$. Among these features, the weather feature is divided into three categories: sunny (or cloudy), rainy, and snowy. The temperature feature has also three types of values: cold (lower than 15 °C), mild (15 -30 °C), hot (over 30 °C). The wind speed is divided into two categories according to the Beaufort number [40]: light (≤ 3) and heavy (> 3).

After identifying the related features, we then develop an XGBoost [41]-based model to predict the *net flow* of each service station in each time slot based on the long-term real-world shared EV fleet operation data. XGBoost uses a gradient boosting framework and is one of the most effective machine learning models for prediction. Besides, the base model of XGBoost is a decision tree, so it has the potential to show better performance against overfitting and it normally shows the best performance for the problems with small-to-medium structured/tabular data. The predicted *net flow* of the developed model can be represented as

$$\hat{f}_i = \sum_{k=1}^K h_k(\mathbf{x}_i), h_k \in \mathcal{H}, \quad (2)$$

where K is the number of trees; $\mathbf{x}_i$ is the i^{th} input, including the five categories of extracted features (12 in total); $\hat{f}_i$ is the corresponding predicted output, which is learned by a tree ensemble model with a collection $\mathcal{H}$ of K functions h_k . Then the objective function at training round t iteration can be denoted as

$$J^{(t)} = \sum_{i=1}^n (l(f_i, \hat{f}_i)) + \sum_{k=1}^t \Omega(f_k), \quad (3)$$

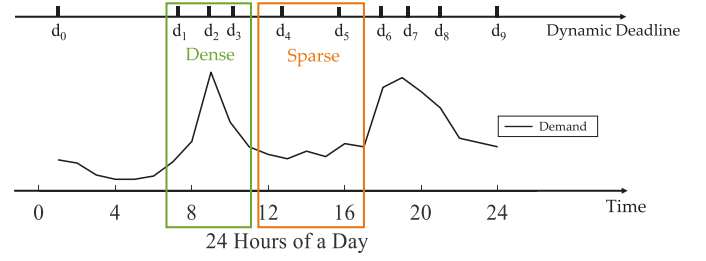


Fig. 9. The key idea of the dynamic deadline setting.

where $l(\cdot)$ is the loss function (e.g., Square loss); Ω is the regularization term (e.g., $L2$ norm), which measures the model complexity.

5.3 Dynamic Deadline Generation Module

As indicated in Fig. 6, the number of pickups and returns are highly dynamic in both spatial and temporal dimensions, so it may be challenging to satisfy the pickup demand and return demand in some usage peak hours and it may also cause high operation costs during low demand hours with only predefined static schedules. Hence, in this paper, we develop a dynamic deadline strategy to address this issue. An illustration of the dynamic deadline setting is shown in Fig. 9. The key idea of our dynamic deadline design is that we set dense deadlines during peak hours and sparse deadlines in other hours, which means we will have more frequent dispatching during peak demand hours and fewer actions during other hours. In this case, the accessibility may be guaranteed in time and the system efficiency can be also improved based on the dynamic deadlines.

Suppose there are N SAEVs $\{EV_1, EV_2, \dots, EV_N\}$ in the SAEV fleet, and n service stations $\{s_1, s_2, \dots, s_n\}$ are deployed across the city to park these EVs. The capacity of each service station is $\{c(1), c(2), \dots, c(n)\}$, and there are $\{v(1), v(2), \dots, v(n)\}$ available SAEVs at each service station in the initial state (e.g., τ_0 in Fig. 8). If the accumulated net flow of service station s_i over the time period T is $F(i) = \{F_{t_1}(i), F_{t_2}(i), \dots, F_{t_m}(i)\}$, and there are $k(i) = \{k_{t_1}(i), k_{t_2}(i), \dots, k_{t_m}(i)\}$ SAEVs that have the battery level lower than a threshold resulting in unavailable at this service station over time, then we can find that there are $F_{t_j}(i) + v(i) - k_{t_j}(i)$ available SAEVs at the service station s_i in the j th time slot. If the predicted number of pickups and returns at s_i in the $(j+1)$ th time slot is $\hat{p}_{t_{j+1}}(i)$ and $\hat{r}_{t_{j+1}}(i)$, and the estimated number of low-battery SAEVs is $\hat{k}_{t_{j+1}}(i)$, so the predicted net flow of s_i in $(j+1)$ th time slot is $\hat{f}_{t_{j+1}}(i) = \hat{p}_{t_{j+1}}(i) - \hat{r}_{t_{j+1}}(i)$ (more detailed prediction process will be elaborated in Section 5.1). Hence, the extra SAEV demand $\hat{v}d_{t_{j+1}}(i)$ and parking spot demand $\hat{p}d_{t_{j+1}}(i)$ of s_i in $(j+1)$ th time slot are

$$\hat{v}d_{t_{j+1}}(i) = \max\{0, -\hat{f}_{t_{j+1}}(i) - (F_{t_j}(i) + v(i) - k_{t_j}(i))\}, \quad (4)$$

$$\hat{p}d_{t_{j+1}}(i) = \max\{0, \hat{f}_{t_{j+1}}(i) + (F_{t_j}(i) + v(i) - c(i))\}. \quad (5)$$

Definition 4. We define the deficiency count of the service station s_i in the $(j+1)$ th time slot as $dc_{t_{j+1}}(i) = \max\{\hat{v}d_{t_{j+1}}(i), \hat{p}d_{t_{j+1}}(i)\}$.

For the n service stations and m time slots, we can obtain a $n \times m$ dimensional station deficiency matrix $D_{n \times m}$.

$$D_{n \times m} = \begin{bmatrix} dc_{t_{j+1}}(1) & \cdots & dc_{t_{j+m}}(1) \\ \vdots & \ddots & \vdots \\ dc_{t_{j+1}}(n) & \cdots & dc_{t_{j+m}}(n) \end{bmatrix}. \quad (6)$$

Each row of the matrix stands for the *deficiency count* of a service station over time. However, if we make a scheduling decision in every time slot, it may cause very high overhead for the fleet management system. Hence, in this paper, we try to set a sequence of dynamic scheduling deadlines, which combine several 5-minute small time slots into an *interval* based on the *deficiency count* of all service stations.

Suppose we finally combine the m time slots into h intervals, e.g., $\{I_1, I_2, \dots, I_i, \dots, I_h\} = \{(\tau_0, d_1], (d_1, d_2], \dots, (d_{i-1}, d_i], \dots, (d_{h-1}, d_h]\}$, where τ_0 is the start time, each interval could consist of different number of small time slots, and d_i is the i th deadline for SAEV decision making. The basic idea of the dynamic deadline is that the *deficiency count* between two consecutive deadlines should not be too small or too large for all service stations (i.e., under a threshold Ψ), which means that there should be a certain amount of demand for extra SAEVs or parking spots in some service stations before making a management decision and the vehicles can finish movements in time, and we should satisfy enough supply for SAEVs and parking spots of the next interval I_{j+1} in all service stations by the deadline d_j .

To achieve the above objective, we combine the maximum entry of each column of $D_{n \times m}$ to obtain a *max deficiency vector* $\mathbf{D}_m = [\max D_{:,1}, \max D_{:,2}, \dots, \max D_{:,m}]$. We first empirically select an initial threshold Ψ based on observations of $\mathbf{D}_m$, then Ψ will also dynamically update with the operation of the fleet management system.

5.4 Prediction Error Compensation Mechanism

However, in practice, it is possible to make inaccurate predictions due to some complicated and unexpected usage behaviors even using advanced prediction methods combined with good feature extraction. Hence, in this paper, we further design an Error Compensation (EC) mechanism to mitigate the influence of inaccurate predictions.

Since the objective of our work is to improve the operational efficiency of the SAEV fleet without sacrificing user experience, our system should guarantee users' pickup and return demand by each deadline. However, if the absolute value of the predicted *net flow* is much smaller than the absolute value of the true value in some service stations, it may cause some users' demand cannot be satisfied. Hence, we add a positive prediction compensation term ϵ , which is an adjustable hyperparameter, in Eqs. (4) and (5) to obtain the following Eqs. (7) and (8) to make the system more robust to prediction errors. The ϵ in Eq. (7) is utilized to compensate for the error that the predicted pickup demand smaller than the real pickup demand. The ϵ in Eq. (8) is utilized to compensate for the error that the predicted parking spot demand smaller than the real parking spot demand.

$$\widehat{vd}'_{t_{j+1}}(i) = \max\{0, -\widehat{f}_{t_{j+1}}(i) - (F_{t_j}(i) + v(i) - k_{t_j}(i)) + \epsilon\} \quad (7)$$

$$\widehat{pd}'_{t_{j+1}}(i) = \max\{0, \widehat{f}_{t_{j+1}}(i) + (F_{t_j}(i) + v(i)) - c(i) + \epsilon\}. \quad (8)$$

Based on the new extra demand of SAEVs $\widehat{vd}'_{t_{j+1}}(i)$ and demand of parking spots $\widehat{pd}'_{t_{j+1}}(i)$, we obtain the new station

deficiency matrix $D'_{n \times m}$ and max deficiency vector $\mathbf{D}'_m$ for dynamic deadlines generation as shown in Section 5.3.

5.5 DRL-Based Decision Making Module

The dynamic deadline is the basis for the following decision making. The purpose of setting the dynamic deadlines is to decide when to make decisions for repositioning and charging. Hence, after determining the deadlines, we then utilize these deadlines as the start and end times of decision making for unoccupied SAEVs. The goal of the real-time repositioning and charging for fleet management is to enable shared autonomous mobility accessibility, which potentially improves the *profit* of the operator without sacrificing user experience. It means that there should always be available SAEVs in a service station when users pick SAEVs up, and there should always be unoccupied parking spots when users return SAEVs. Here, we consider the *revenue* from serving users and *costs* for charging and repositioning as shown in Eq. (1).

The goal of the agent is to learn a policy $\pi(a_t|s_t)$ so as to maximize the expected cumulative future rewards $G_t = \sum_{i=0}^T \gamma^i R_{t+i+1}$ in an episode T starting from time t . To solve an MDP, a common objective is to learn the value functions, including the state-value function $V_\pi(s)$ and the state-action value function (i.e., Q-function) $Q_\pi(s, a)$, where $V_\pi(s) = \mathbb{E}_\pi[G_t|s_t = s] = \mathbb{E}_\pi[\sum_{i=0}^T \gamma^i R_{t+i+1}|s_t = s]$ and $Q_\pi(s, a) = \mathbb{E}_\pi[\sum_{i=0}^T \gamma^i R_{t+i+1}|s_t = s, a_t = a]$, which measures the expected discounted sum of rewards obtained from state s by taking action a at time t and following policy π .

In this paper, we design a dynamic deadline-based deep reinforcement learning (DRL) method (i.e., Dynamic Deadline-based Deep Q-Network (D3QN)) to learn the optimal actions for individual SAEVs, i.e., all available SAEVs sequentially learn which action to take between two consecutive deadlines. The advantage of our D3QN lies in its computational efficiency since we streamline our D3QN training and each individual SAEV independently learns its own optimal policy, which ensures scalability at the cost of less coordination between EVs. This is challenging to be achieved by standard DQN or a multi-agent formulation [13], [18].

We define the optimal Q-function for k^{th} SAEV as the maximum expected return achievable by any policy π_t , which is

$$Q^*(s, a) = \max_{\pi} \mathbb{E} \left[\sum_{i=0}^T \gamma^i R_{t+i+1}^{(k)} | s_t^{(k)} = s, a_t^{(k)} = a, \pi_t \right], \quad (9)$$

which satisfies the Bellman Equation

$$Q^*(s, a) = \mathbb{E}_{s'} \left[R_{t+1} + \gamma \max_{a'} Q^*(s', a') | s_t^{(k)} = s, a_t^{(k)} = a \right]. \quad (10)$$

For the electric carsharing fleet management problem, it is challenging to obtain the Q-function as a table containing values due to the large-scale states and actions. Hence, we use deep neural networks to approximate the Q-value function with parameters θ , which makes $Q(s, a; \theta) \approx Q^*(s, a)$. This can be achieved by updating θ_i at each iteration i to minimize the following loss function:

$$\mathcal{L}_i(\theta_i) = \left(R_{t+1} + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s, a; \theta_i) \right)^2, \quad (11)$$

which is the squared difference between the target Q values $R_{t+1} + \gamma \max_{a'} Q(s', a'; \theta_i^-)$ and the approximate Q values $Q(s, a; \theta_i)$. To make the network updates more stable, we utilize the *experience replay* technique [17], which copies θ to another neural network θ_i^- so that we can fix the Q-value targets temporarily. This forms an input dataset that is stable enough for training. θ_i^- are the parameters from the previous iteration, which are fixed and not updated for learning θ_i . Then we utilize stochastic gradient descent (SGD) with respect to the actual network parameters to minimize this loss. Finally, we obtain the action for each unoccupied SAEV to take.

Algorithm 1. Dynamic Deadline-Based Deep Reinforcement Learning for Decision Making

Input: n service stations $\{s_1, s_2, \dots, s_n\}$; m time slots $\{t_1, t_2, \dots, t_m\} \in [\tau_0, T]$; initial threshold Ψ

- 1 **for** each service station $s_i, i \in [1, n]$ **do**
- 2 **for** each time slot $t_j, j \in [1, m]$ **do**
- 3 Count net flow $f_{t_j}(i)$;
- 4 Count accumulated net flow $F_{t_j} = \sum_{z=1}^j f_{t_z}(i)$;
- 5 Count number of low-battery SAEVs $k_{t_j}(i)$;
- 6 Predict the number of pickups and returns of $(j+1)th$ time slot $\hat{p}_{t_{j+1}}(i)$ and $\hat{r}_{t_{j+1}}(i)$;
- 7 Estimate number of low-battery SAEVs of $(j+1)th$ time slot $\hat{k}_{t_{j+1}}(i)$
- 8 Predict extra SAEV and parking spot demand with error compensation of $(j+1)th$ time slot $\hat{v}d'_{t_{j+1}}(i)$ and $\hat{p}d'_{t_{j+1}}(i)$ as in Eqs. (7) and (8)
- 9 Predict deficiency count of $(j+1)th$ time slot $dc'_{t_{j+1}}(i) = \max\{\hat{v}d'_{t_{j+1}}(i), \hat{p}d'_{t_{j+1}}(i)\}$
- 10 Generate a station deficiency matrix $D'_{n \times m}$
- 11 **for** $j=1$ to m **do**
- 12 Count the maximum entry of each column $D'_{:,j}$
- 13 Generate a max deficiency vector $\mathbf{D}'_m = [\max D'_{:,1}, \max D'_{:,2}, \dots, \max D'_{:,m}]$
- 14 Initialize the interval id $i = 1$
- 15 **for** $j=1$ to m **do**
- 16 **while** $\mathbf{D}'_m[j] < \Psi$ **do**
- 17 $I_i = \text{merge}(\emptyset, t_j)$;
- 18 Deadline $d_i = I_i.\text{end}; i++$
- 19 Obtain h intervals $I = \{I_1, \dots, I_h\}$
- 20 Initialize experience replay memory DD
- 21 Initialize the action-value function Q with random weights
- for** $m=1$ to the maximum iteration number M **do**
- 22 initial the state s_0
- 23 **for** each interval $I_i \in I$ **do**
- 24 Select $a_i = \text{argmax}_a Q^*(s_i, a)$
- 25 Execute action a_i and observe reward R_{i+1}
- 26 Store transition $(s_i, a_i, R_{i+1}, s_{i+1})$ in DD
- 27 Sample a batch of transitions from DD , and then update parameter of the value network θ_i by minimizing the value loss function $\mathcal{L}_i(\theta_i)$ over the batch in Eq. (11) by SGD to get θ_i .

In general, the process of our dynamic deadline-based deep reinforcement learning is shown in Algorithm 1. Line 1 to line 19 are for the dynamic deadline generation, and the

following parts are for decision making based on the dynamic deadlines as shown in line 23.

6 EVALUATION

In this section, we extensively evaluate the performance of our Safari based on the real-world data.

6.1 Evaluation Data and Setting

Evaluation Data. We utilize 10-month shared EV usage data generated by 12,375 unique users in the Chinese city Beijing for evaluation. More than 86,700 usage records are generated during this period. In addition to the vehicle usage data, the evaluation dataset also includes the vehicle GPS data, and metadata of 185 service stations and 226 charging stations. The detailed data information has been introduced in Section 5.1.

Environment Setting. We train and test our prediction models and fleet management decisions on a desktop with 32 GB memory, 1 TB HDD storage, Intel Xeon CPU E5-1660 v3, and a Tesla K40c, installed with the latest Windows 10 and Python coding environment.

For the D3QN-based decision making strategies, we have the following parameter setting: the same three hidden layer Q-network with 128, 64, and 32 nodes from the first to last hidden layer; the activation functions of all hidden units are ReLu, and output layers of the Q-networks use Softmax activation functions. All the experiments are repeated 10 times to ensure the robustness of the results. The batch size of all deep learning networks is set to be 2,000, and we utilize AdamOptimizer with a learning rate of 0.001. For the discount factor, we select $\gamma = 0.99$ based on [17], so the state value is computed within a decaying future horizon.

6.2 Evaluation Results

6.2.1 Comparison of Different Management Strategies

To show the effectiveness of our Safari, we compare it with (i) Ground Truth (GT), which is extracted from our real-world data; (ii) Best-effort [42], which means once there is user demand in a service station, the vehicle repositioning will be performed to satisfy the demand. Once the battery level of a SAEV is below the threshold, it will be scheduled to charge for future user accessibility. (iii) TBA [43], which is a reinforcement learning-based method adopting the REINFORCE rule to update the policy for fleet management. (iv) PPO in [13], which is a state-of-the-art shared EV repositioning algorithm based on DRL.

Since the key focus of the SAEV fleet management is the *accessibility*, which is defined as the percentage of satisfied user demand, we first study the accessibility of different methods. From Fig. 10, we found our dynamic deadline DRL-based Safari achieves the best performance with 100% of accessibility, which means it can satisfy all travel demand of all users, while only 87% of user demand was satisfied by the current system.

6.2.2 Dynamic Deadline versus Static Deadline

In this paper, we designed a dynamic deadline mechanism to achieve a win-win performance, i.e., satisfactory *user experience* and low operational *cost*. We compare Safari

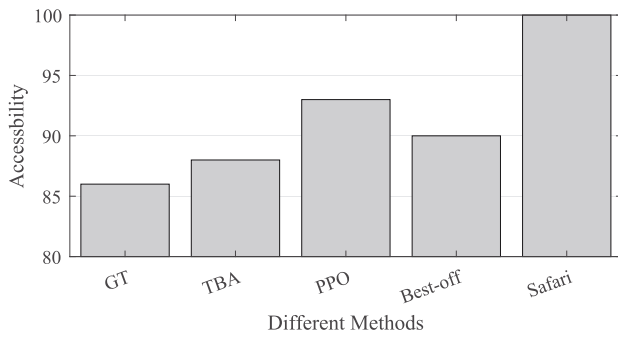


Fig. 10. Mobility accessibility under different methods.

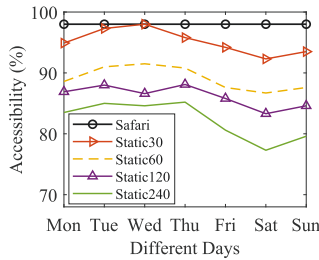


Fig. 11. Accessibility under different methods in different days.

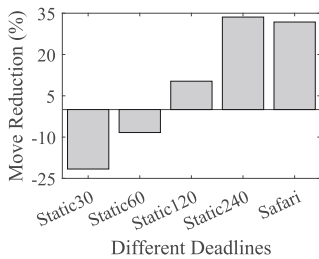


Fig. 12. Percentage reduction of movements.

(i.e., dynamic deadline + DQN) with strategies of different static deadlines + DQN, and we set four typical static deadlines for comparisons, i.e., setting a deadline for every 30 minutes (i.e., DQN + Static30 or Static30 for short), 60 minutes (i.e., Static60), 120 minutes (i.e., Static120), and 240 minutes (i.e., Static240).

Fig. 11 shows the accessibility of Safari and its variants with different static deadlines, which implicitly reflects the user experience, e.g., high accessibility means more user demand can be satisfied since more users can have access to the services in time when they want to use SAEVs. From Fig. 11, we found our dynamic deadline mechanism satisfies all user demand on different days. However, as the duration between two consecutive deadlines becomes larger, more users may not have access to available SAEVs in time, resulting in poor user experience.

Fig. 12 shows the percentage reduction of vehicle movement under Safari and its variants with different static deadlines, which is calculated by (current number of movements - number of movements with different deadline settings)/current number of movements. From Fig. 12, we found Safari reduces 31.8% of movements. With more sparse deadlines, the number of movements becomes less, e.g., Static240 causes fewer movements than Safari, but more demand cannot be satisfied by Static240 as shown in Fig. 11. With more dense deadlines, the overhead becomes

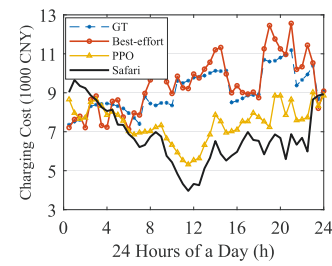


Fig. 13. Monthly charging cost distributions of different methods.

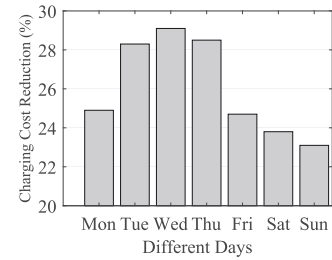


Fig. 14. Reduction of cost of Safari in different days.

higher due to frequent movements for relocating or charging SAEVs. Since the energy consumption is positively correlated with the number of movements and movement distance, the reduction of movements potentially indicates energy cost reduction and profit increase for the SAEV fleet.

Another important factor for SAEV fleet management is the charging cost due to the time-varying charging pricing, so we show the charging cost distribution at different times of a day under different strategies in Fig. 13. We found our Safari can reduce charging costs during most time of the day since it can reduce some unnecessary repositioning and leave some SAEVs to charge in the late-night time, during which the charging price is low. In total, our Safari can reduce about 26.2% of charging costs for the SAEV fleet. Due to the frequent repositioning and charging activities of Best-effort, it causes higher charging costs in some high charging pricing durations, e.g., 13:00-15:00 and 19:00-20:00, which leads to a 3.9% of charging cost increase. Even though PPO is also based on DRL, it sets periodic static deadlines for repositioning, so it also causes higher charging costs.

In addition, we also show the charging cost reduction of Safari on different days of a week, as shown in Fig. 14. We found our Safari shows better performance through Tuesday to Thursday and a little reduction on weekends. One possible reason would be that there is no huge pickup and return peaks from Tuesday to Thursday, which makes less repositioning, so more SAEVs can be charged in low charging price durations. There is highly intensive demand on weekends from 8:00-20:00, which causes more repositioning and charges in daytime to satisfy user demand, so the charging cost reduction is relatively low on weekends.

6.2.3 Prediction Performance

We also compare the XGBoost-based method with other state-of-the-art prediction approaches, including Auto-Regressive and Moving Average (ARMA), and Random Forest (RF), Deep Neural Network (DNN), LSTM, and Graph Convolutional Network (GCN). As we found the

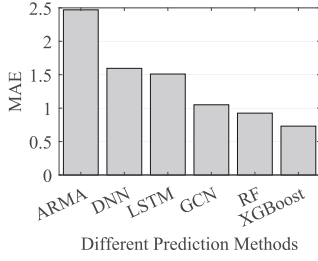


Fig. 15. MAE of different prediction methods.

decision making is directly related to the *net flow* of service stations, so we predict the *net flow* (i.e., # of Returns- # of Pickup) of each service station instead of predicting the number of pickups and returns separately.

We adopt the Mean Absolute Error (MAE) to compare the prediction performance of different methods, which is computed as $MAE = \frac{1}{n} \sum_{i=1}^n |y^{(i)} - \hat{y}^{(i)}|$, where $\hat{y}$ is the predicted value, y is the true value, and n is the total number of predictions.

Fig. 15 also gives evidence of the advantages of the XGBoost-based approach since its MAE is as low as 0.73, which indicates it achieves very high accuracy for the *net flow* prediction of most service stations all the time.

We found some other methods like RF also achieve good performance. One possible reason for the high prediction accuracy is our data-driven feature extraction as we extract 5 categories of features that are highly related to users' pickup and return behaviors based on our data-driven observations. Hence, we further conduct ablation studies on extracted features.

6.2.4 Ablation Studies on Extracted Features and Error Compensation to System Performance

As shown in Fig. 16, we found lacking any category of features, the system may damage user experience due to inaccurate prediction, e.g., only 88.6% of user demand can be satisfied without the temporal features $\mathcal{F}_T$ by Safari (i.e., NoFT). Similarly, without the demographic feature, Safari can only satisfy 93.8% of user demand due to large prediction error.

In addition, although considering all categories of features, it is challenging to achieve 100% of prediction accuracy due to the complicated usage behavior and environment, as shown in Fig. 15. Hence, in this paper, we introduced an error compensation term ϵ to mitigate the influence of the inaccurate prediction because the prediction results will directly impact the dynamic deadline generation and future decision making, e.g., if the prediction error is too large, the supply of each station may not guarantee future user accessibility.

Then we need to decide the value of the error compensation term ϵ . Too small error compensation may cause more users cannot have access to SAEVs in time, while too large error compensation may cause vehicle or parking resources waste and potential inaccessibility of users. In this work, our strategy to decide the value of ϵ is that we round the MAE to the next integer, i.e., $\epsilon = \text{ceiling}(MAE)$. For example, 0.73 is rounded to 1 and 1.1 is rounded to 2. The reason why we utilize the MAE as an indicator is that it can show what is the general prediction deviation and also get rid of the influence of outliers.

Authorized licensed use limited to: Florida State University. Downloaded on January 04, 2024 at 02:57:36 UTC from IEEE Xplore. Restrictions apply.

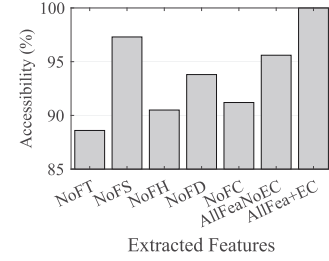


Fig. 16. Impact of feature extraction and EC.

For example, when only one or a small number of prediction values are largely different from the true values, if we consider then them when we decide the value of ϵ , it may cause huge resource waste and potential inaccessibility of users since we need to dispatch more vehicles to these stations or leave more spare parking spaces. In addition, we found that over 96% of prediction deviations are small than $\text{ceiling}(MAE)$ based on our results, so we set $\epsilon = \text{ceiling}(MAE)$ as the error compensation term.

The results show that our data-driven feature extraction combined the error compensation (i.e., AllFea+EC) can make our Safari satisfy all user demand (pickup + return); while it can only satisfy 95.6% of user demand without the error compensation (i.e., AllFeaNoEC). Hence, with the prediction error compensation mechanism, our Safari becomes more robust to prediction uncertainties. It should be noted that the value of the error compensation term ϵ can be updated with the operating of the system.

6.3 Convergence and Runtime Overhead

The training process of Safari is in an offline fashion, and we find that the offline learning process costs an average of 18.6 minutes in our experiments, which is less than the other deep reinforcement learning-based dispatch models, e.g., PPO (35.5 minutes), Static30 (26.3 minutes), and Static60 (22.6 minutes). Fig. 17 shows the training process of our Safari and other RL-based methods. We found that our D3QN converges and achieves the highest reward after about 3,000 steps of training. We also found that the average loss variations reduce as the training steps increase, so the predicted Q-values converge to the target ones quickly in our setting. We found the majority of the overhead is caused by the decision making. Safari should be operated in an online fashion, so we also evaluate the convergence and incurred overhead at runtime. Since our Safari is based on the D3QN, it significantly reduces the overhead of decision making. We found that with 100 SAEVs in the fleet, the

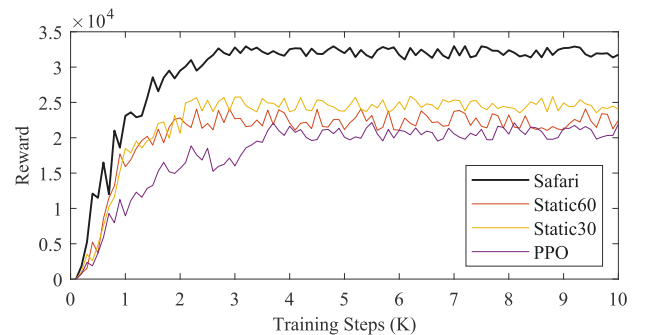


Fig. 17. Learning curves of different methods.

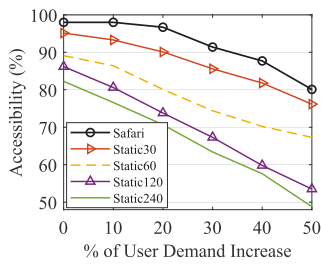


Fig. 18. Accessibility with user demand increase.

maximum overhead for a round of decision making is 0.12 s. As the system scales to 500 and 1,000 SAEVs, the overhead will be progressively larger to 0.8 s and 1.3 s, respectively. Nonetheless, the overhead remains reasonably small and negligible compared to the deadline duration and user pickup time, so it does not affect the real-world operation of the SAEV fleet as Safari still satisfies the accessibility of users by each deadline.

6.4 Performance on Larger Fleets and Higher Demand

It will be a long-term process for SAEV promotion, and the fleet size, the number of users, and the number of usages will be evolving during this process, e.g., the number of monthly usages increased from 113 to 21,750 during the 10 months. The rapid increase of demand potentially leads to more SAEVs needed or more frequent dispatching. Hence, we also investigate the impacts of fleet sizes and user demand on our fleet management system Safari in this paper to verify its generalizability.

In this part, we focus on two questions: (i) can our Safari satisfy higher user demand without adding extra infrastructures (e.g., charging stations and service stations, and SAEVs)? (ii) How to expand a small fleet size to balance the future user demand increase and the number of movements needed? It should be noted that we do not consider new service stations and charging stations in other locations since station siting is another parallel topic of our work, so we leave this for future work.

6.4.1 Impacts of User Demand on Safari Performance

There should be higher user demand in the near future with more and more people utilize SAEV services, and the fleet management system needs to make decisions to adaptively satisfy the new demand. Without loss of generality, we assume there is a different percentage increase of user demand during the evolving process, and the new usage in each service station is proportional to the current usage distribution.

Figs. 18 and 19 show the percentage of satisfied user demand and the percentage of movement reduction with user demand increase, respectively. We found the percentage of satisfied user demand will decrease when user demand increases without adding extra infrastructures (e.g., charging stations and service stations, and SAEVs), but our Safari always achieves the best performance compared to other static deadline-based methods. Under sparse deadline-based methods, the average satisfied user demand will decrease, but they take fewer movements to relocate SAEVs.

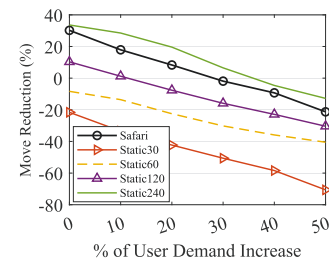


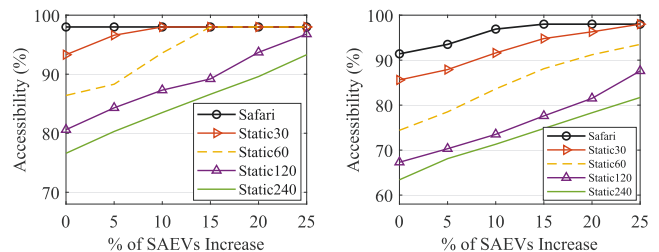
Fig. 19. % movement reduction with user demand increase.

Even though Static240 needs the fewest movements, it satisfies much less user demand than our Safari. In general, our Safari can achieve good performance with higher user demand in the future, so it has the potential to sustainably manage SAEV fleets during their evolving process.

6.4.2 Impacts of Fleet Sizes on Safari Performance

Similarly, operators would also expand their SAEV fleet size in the long run to satisfy higher user demand, and the user demand increase will interact with the fleet size expansion during the long-term evolving process. Hence, we also investigate the impacts of fleet size on system performance. Intuitively, if more SAEVs are added to the fleet, higher user demand can be satisfied, but it may also need more movements to relocate and charge SAEVs, so we also investigate the correlation between fleet size and user demand. Without loss of generality, we initially add the new SAEVs in each service station that is proportional to the number of SAEVs in the early morning, and then the vehicles will be relocated and reach a relatively stable status after a period of time to satisfy user demand.

Figs. 20a and 20b show the accessibility with both the fleet size increase and user demand increase, and there is 10% of user demand increase in Fig. 20a and 30% of user demand increase in Fig. 20b, respectively. We found that more user demand will be satisfied with more SAEVs added to the fleet. Specifically, When there is a 10% of user demand increase, our Safari can satisfy all user demand without adding extra new SAEVs; When there is a 30% of user demand increase, our Safari can satisfy over 93% of user demand without adding extra new SAEVs and it can satisfy all user demand with 15% of new SAEV increase. This result can provide guidance for operators to balance the number of new SAEVs and user satisfaction. If the operator intends to reduce operational costs, fewer new SAEVs should be good for most users. If the operator plans to satisfy all user demand, more new SAEVs will be added to the fleet with higher operational costs.



(a) 10% of user demand increase. (b) 30% of user demand increase.

Fig. 20. Accessibility with fleet size & demand increase.

7 DISCUSSION

In this section, we report some lessons learned from our data-driven study, and then we discuss several potential implications of our work.

7.1 Lessons Learned

- *Data-Driven Insights.* In this work, we revealed a set of new findings of SAEVs based on our data-driven investigation, e.g., (i) The usage patterns of shared EVs are highly dynamic on different regions and different time of a day. (as shown in Fig. 6) (ii) Time-variant charging pricing also impacts the operating profit of the SAEV fleets (e.g., Fig. 7). (iii) Users' usage behavior is highly related to different factors, e.g., spatiotemporal and demographic factors (Section 5.1).
- *Dynamic Deadline-Based Fleet Management.* In this paper, we designed a dynamic deadline mechanism-based deep reinforcement learning algorithm, which shows good performance for real-time fleet management compared to existing predefined static schedules. Our design showed good performance to improve the mobility accessibility of users and SAEV fleet operators' profits (e.g., Figs. 12 and 13) with a small runtime overhead. Our collaborator also has agreed to apply our Safari for their real-world operation.

7.2 Practical Impacts

- *SAEV Fleet Management.* User experience and profitability of operators are two key factors that impact the efficiency of SAEV fleets. From this work, we found our dynamic deadline-based fleet management system Safari can increase the operating profits for SAEV operators and satisfy highly dynamic user demand to guarantee the mobility accessibility, so it has the potential to be reapplied to other cities and enlarge the SAEV fleets. In addition, our system also has a great potential for efficient fleet management of SAEV during its long-term expansion and evolution process, e.g., when there are more vehicles added to the SAEV fleet in the future.
- *Current Fleet Management of Electric Bikes, E-Scooter Sharing, and Electric Ridesharing.* Even though this paper focuses on the SAEV fleets, we believe our joint real-time repositioning and charging scheduling and the dynamic deadline-based DRL method have the potential to be reapplied to other types of EV fleet management, e.g., electric bikesharing, e-scooter sharing, and electric ridesharing. The difference is that extra workers are needed to move e-bikes/e-scooters, and the ridesharing drivers need to perform the repositioning and charging by themselves.

7.3 Impact of Different Factors

7.3.1 Impact of the Number of Charging Stations

The number of public charging stations in the city will affect the charging activities of shared autonomous electric vehicles (SAEVs). For example, when there is a limited number of charging stations, different from the solution in this

work, the SAEVs cannot always be charged in the nearest available charging stations, which means the action $\mathcal{A}_C$ (scheduling to charge in a charging station) will be affected when making decisions, and it may further increase the charging costs. From another point of view, the charging supply is related to the charging supply. A limited number of charging stations is also enough for some other cities with a small number of SAEVs with fewer charging demands. Furthermore, as shown in Figs. 18 and 19, we found the percentage of satisfied user demand will decrease when user demand increases without adding extra infrastructures (e.g., charging stations), which might be similar to reducing the number of charging stations without increasing charging demand, but our Safari always achieves the best performance compared to other static deadline-based methods. Hence, (i) a limited number of charging stations may not affect the proposed solution in other cities with fewer charging demands. (ii) If the number of charging stations cannot satisfy the charging demand of SAEVs, it may increase the charging costs, but our Safari can still achieve better performance than other state-of-the-art methods.

7.3.2 Impact of the Time Slot Period

Based on our data-driven investigation, the per-usage time of shared EVs is typically very long, e.g., most per-usage time is longer than 30 minutes and 50% of usage is longer than 5 hours as shown in Fig. 4. In addition, the times for relocation and charging are also need more than 10 minutes and one hour, respectively. Hence, in this work, we try to set the time slot length as a short duration to capture the more fine-grained pickup and return patterns. However, if the time slot is too short, the usage prediction at each time slot would be not so accurate. After trying several possibilities (1 minute, 5 minutes, 10 minutes, and 20 minutes), we found the system performance is the best when the time slot is set to be 5 minutes. Hence, if we set the time slot to be much less than 5 minutes (e.g., 1 minute), the system is more likely to cause a larger prediction error, and it will also cause a higher computational complexity. The prediction error will further affect future decision making process and accessibility to users. If we set the time slot to be much more than 5 minutes (e.g., 20 minutes), then the merged deadlines will be sparser. In this case, it is challenging to fully satisfy the user demand at intensive and unexpected peaks, so some users cannot have access to the mobility system in time. Hence, we select the 5 minutes after carefully comparison by considering different factors.

7.3.3 Impact of the Wireless Access Technology

Wireless access technology is critical for our work since it facilitates the data collection and information transmission. For example, the GPS data is uploaded to servers periodically through 4 G/5 G cellular communication. The decisions for vehicle repositioning and charging should also be transmitted by wireless communication. Hence, the reliability and stability of wireless communication are important. However, the real-time status or performance of wireless access in specific areas will not have a huge impact on the performance of the proposed algorithm. As shown in the framework of our Safari design, we will conduct data

cleaning to process the raw data to filter out errant GPS records. In addition, we do not require vehicles to respond to the decisions immediately since the repositioning time or charging time should be much longer than the decision's transmission time, which means a short delay due to the wireless communication being tolerable in our system. If there are some wireless network interruptions lasting for a long period, the proposed algorithm will be impacted since information cannot be transmitted at that time.

7.3.4 Impact of the Real-Time Charging Pricing

One contribution of our work is that we consider the time-varying charging pricing for decision making to reduce the charging costs. If the charging rates are the same at different times, we do not need to consider the charging rates, which means we will take the charging scheduling action $\mathcal{A}_C$ immediately when the battery level of a SAEV decreases to below a certain threshold. However, in real-world scenarios, the charging pricing is time-varying with different charging rates in many cities. In our system design, we consider this real-world factor to make the system more practical even though it will also increase the complexity of the design. During the high charging rates duration, if there are enough SAEVs for satisfying user demand, we will keep the low-battery SAEVs at the current location and then move them to charge during low charging rates time, which can reduce the charging costs and operation costs. Hence, the charging pricing is a component of our system design, which can be changed based on the real pricing in different cities, and it will not affect how we obtain the final solution.

8 CONCLUSION

In this paper, we design an effective data-driven fleet management system for accessible shared autonomous electric mobility called Safari, which means whenever users have travel demand, they can have access to the SAEVs in time. Our Safari jointly considers repositioning and charging with dynamic deadlines to improve both the user experience and operating profits. In Safari, we designed a dynamic deadline-based deep reinforcement learning algorithm D3QN to adaptively satisfy the time-varying unbalanced user demand. The dynamic deadlines are learned through usage prediction combined with an error compensation mechanism. Extensive experimental results show that our designed Safari potentially guarantees 100% of mobility accessibility and effectively reduces 26.2% of charging costs and 31.8% of vehicle movements for SAEV fleets, and it also achieves a small runtime overhead.

ACKNOWLEDGMENTS

The authors would like to thank anonymous reviewers for their valuable comments and suggestions.

REFERENCES

- [1] L. D. Burns, "A vision of our transport future," *Nature*, vol. 497, no. 7448, pp. 181–182, 2013.
- [2] C. D. Korkas, S. Baldi, S. Yuan, and E. B. Kosmatopoulos, "An adaptive learning-based approach for nearly optimal dynamic charging of electric vehicle fleets," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 7, pp. 2066–2075, Jul. 2018.
- [3] J. Shi, Y. Gao, W. Wang, N. Yu, and P. A. Ioannou, "Operating electric vehicle fleet for ride-hailing services with reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4822–4834, Nov. 2020.
- [4] S. Schoenberg and F. Dressler, "Reducing waiting times at charging stations with adaptive electric vehicle route planning," *IEEE Trans. Intell. Veh.*, early access, Jan. 06, 2022, doi: [10.1109/TIV.2022.3140894](https://doi.org/10.1109/TIV.2022.3140894).
- [5] P. Zhou, C. Wang, and Y. Yang, "Design and optimization of solar-powered shared electric autonomous vehicle system for smart cities," *IEEE Trans. Mobile Comput.*, early access, Sep. 30, 2021, doi: [10.1109/TMC.2021.3116805](https://doi.org/10.1109/TMC.2021.3116805).
- [6] G. Wang, Y. Zhang, Z. Fang, S. Wang, F. Zhang, and D. Zhang, "FairCharge: A data-driven fairness-aware charging recommendation system for large-scale electric taxi fleets," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 4, no. 1, pp. 1–25, 2020.
- [7] Y. Pan, Q. Chen, N. Zhang, Z. Li, T. Zhu, and Q. Han, "Extending delivery range and decelerating battery aging of logistics UAVs using public buses," *IEEE Trans. Mobile Comput.*, early access, Apr. 13, 2022, doi: [10.1109/TMC.2022.3167040](https://doi.org/10.1109/TMC.2022.3167040).
- [8] G. Wang et al., "Pricing-aware real-time charging scheduling and charging station expansion for large-scale electric buses," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 1, pp. 1–26, 2020.
- [9] C. Wang et al., "Optimizing cross-line dispatching for minimum electric bus fleet," *IEEE Trans. Mobile Comput.*, early access, Oct. 12, 2021, doi: [10.1109/TMC.2021.3119421](https://doi.org/10.1109/TMC.2021.3119421).
- [10] Z. Xu et al., "When recommender systems meet fleet management: Practical study in online driver repositioning system," in *Proc. Web Conf.*, 2020, pp. 2220–2229.
- [11] P. Hulot, D. Aloise, and S. D. Jena, "Towards station-level demand prediction for effective rebalancing in bike-sharing systems," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2018, pp. 378–386.
- [12] S. He and K. G. Shin, "Dynamic flow distribution prediction for urban dockless e-scooter sharing reconfiguration," in *Proc. Web Conf.*, 2020, pp. 133–143.
- [13] M. Luo et al., "Rebalancing expanding EV sharing systems with deep reinforcement learning," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, 2020, pp. 1338–1344.
- [14] C. Gambella, E. Malaguti, F. Masini, and D. Vigo, "Optimizing relocation operations in electric car-sharing," *Omega*, vol. 81, 2018, Art. no. 234.
- [15] E. Biondi, C. Boldrini, and R. Bruno, "Optimal charging of electric vehicle fleets for a car sharing system with power sharing," in *Proc. IEEE Int. Energy Conf.*, 2016, pp. 1–6.
- [16] C. A. Folkestad, N. Hansen, K. Fagerholt, H. Andersson, and G. Pantuso, "Optimal charging and repositioning of electric vehicles in a free-floating carsharing system," *Comput. Oper. Res.*, vol. 113, 2020, Art. no. 104771.
- [17] E. Wang et al., "Joint charging and relocation recommendation for e-taxi drivers via multi-agent mean field hierarchical reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 21, no. 4, pp. 1274–1290, Apr. 2022.
- [18] K. Lin, R. Zhao, Z. Xu, and J. Zhou, "Efficient large-scale fleet management via multi-agent deep reinforcement learning," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2018, pp. 1774–1783.
- [19] A. Di Febbraro, N. Sacco, and M. Saeednia, "One-way car-sharing profit maximization by means of user-based vehicle relocation," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 628–641, Feb. 2019.
- [20] X. Zhou et al., "Optimizing taxi driver profit efficiency: A spatial network-based Markov decision process approach," *IEEE Trans. Big Data*, vol. 6, no. 1, pp. 145–158, Mar. 2020.
- [21] G. Wang, F. Zhang, and D. Zhang, "tCharge-A fleet-oriented real-time charging scheduling system for electric taxi fleets," in *Proc. 17th Conf. Embedded Netw. Sensor Syst.*, 2019, pp. 440–441.
- [22] G. Wang, X. Chen, F. Zhang, Y. Wang, and D. Zhang, "Experience: Understanding long-term evolving patterns of shared electric vehicle networks," in *Proc. 25th Annu. Int. Conf. Mobile Comput. Netw.*, 2019, pp. 1–12.
- [23] J. Liu, L. Sun, W. Chen, and H. Xiong, "Rebalancing bike sharing systems: A multi-source data smart optimization," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 1005–1014.
- [24] H. Zhu et al., "RedPacketBike: A graph-based demand modeling and crowd-driven station rebalancing framework for bike sharing systems," *IEEE Trans. Mobile Comput.*, early access, Jan. 25, 2022, doi: [10.1109/TMC.2022.3145979](https://doi.org/10.1109/TMC.2022.3145979).

- [25] M. Luo, B. Du, K. Klemmer, H. Zhu, H. Ferhatosmanoglu, and H. Wen, "D3P: Data-driven demand prediction for fast expanding electric vehicle sharing systems," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 4, pp. 1–21, 2020.
- [26] J. Yuan, Y. Zheng, L. Zhang, X. Xie, and G. Sun, "Where to find my next passenger," in *Proc. 13th Int. Conf. Ubiquitous Comput.*, 2011, pp. 109–118.
- [27] J. Miller and J. P. How, "Predictive positioning and quality of service ridesharing for campus mobility on demand systems," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 1402–1408.
- [28] D. Jorge, G. H. Correia, and C. Barnhart, "Comparing optimal relocation operations with simulated relocation policies in one-way carsharing systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1667–1675, Aug. 2014.
- [29] G. Wang, S. Zhong, S. Wang, F. Miao, Z. Dong, and D. Zhang, "Data-driven fairness-aware vehicle displacement for large-scale electric taxi fleets," in *Proc. IEEE 37th Int. Conf. Data Eng.*, 2021, pp. 1376–1387.
- [30] Z. Tian et al., "Real-time charging station recommendation system for electric-vehicle taxis," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3098–3109, Nov. 2016.
- [31] G. Wang et al., "sharedCharging: Data-driven shared charging for large-scale heterogeneous electric vehicle fleets," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–25, 2019.
- [32] G. Wang, X. Xie, F. Zhang, Y. Liu, and D. Zhang, "bCharge: Data-driven real-time charging scheduling for large-scale electric bus fleets," in *Proc. IEEE Real-Time Syst. Symp.*, 2018, pp. 45–55.
- [33] G. Fan, Z. Yang, H. Jin, X. Gan, and X. Wang, "Enabling optimal control under demand elasticity for electric vehicle charging systems," *IEEE Trans. Mobile Comput.*, vol. 21, no. 3, pp. 955–970, Mar. 2022.
- [34] Y. Yuan, D. Zhang, F. Miao, J. Chen, T. He, and S. Lin, "p² charging: Proactive partial charging for electric taxi systems," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst.*, 2019, pp. 688–699.
- [35] W. Zhang et al., "Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning," in *Proc. Web Conf.*, 2021, pp. 1856–1867.
- [36] Y. Li et al., "Cooperative service placement and scheduling in edge clouds: A deadline-driven approach," *IEEE Trans. Mobile Comput.*, vol. 21, no. 10, pp. 3519–3535, Oct. 2022.
- [37] B. Bjerv, "Beijing public charging points will implement peak and off-peak electricity prices," 2016. [Online]. Available: <http://www.hmc-hongdu.com/news/20.html>
- [38] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and POIs," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2012, pp. 186–194.
- [39] Tianqihoubao, "Historical weather in shenzhen," 2020. [Online]. Available: <http://www.tianqihoubao.com/lishi/shenzhen/month/201910.html>
- [40] Wikipedia, "Beaufort scale," 2020. [Online]. Available: https://en.wikipedia.org/wiki/Beaufort_scale
- [41] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 785–794.
- [42] Z. Li, Y. Zhang, Y. Zhao, Y. Peng, and D. Li, "Best effort task scheduling for data parallel jobs," in *Proc. ACM SIGCOMM Conf.*, 2016, pp. 555–556.
- [43] F. Borutta, S. Schmoll, and S. Friedl, "Optimizing the spatio-temporal resource search problem with reinforcement learning (GIS Cup)," in *Proc. 27th ACM SIGSPATIAL Int. Conf. Adv. Geogr. Inf. Syst.*, 2019, pp. 628–631.



Guang Wang received the PhD degree in computer science from Rutgers University, New Brunswick, New Jersey. He is an Assistant Professor with the Department of Computer Science, Florida State University. Before that, he was a postdoctoral research associate with the Massachusetts Institute of Technology. He is interested in mobile computing, cyber-physical systems, Big Data analytics, and machine learning. His technical contributions have led to more than 40 peer-reviewed publications in premium conferences and journals, e.g., *MobiCom*, *IMWUT/UbiComp*, *RTSS*, *KDD*, *ICDE*, *WWW*, *AAAI*, *IEEE Transactions on Mobile Computing*, *IEEE Transactions on Vehicular Technology*, *IEEE Transactions on Intelligent Transportation Systems*, and *ACM Transactions on Intelligent Systems and Technology*.



Zhou Qin is currently working toward the PhD degree with the Department of Computer Science, Rutgers University, New Brunswick, New Jersey. His research interests include data-driven applications in smart cities via Big Data analytics and machine learning, such as user behavior modeling, time series prediction, and recommendation. His technical contributions have led to publications in top-tier conferences and journals such as *UbiComp/IMWUT*, *SenSys*, *KDD*, and *ACM Transactions on Sensor Networks*.



Shuai Wang (Member, IEEE) received the BS and MS degrees from the Huazhong University of Science and Technology, China, and the PhD degree from the Department of Computer Science and Engineering, University of Minnesota, in 2017. He is currently a professor with the School of Computer Science and Engineering, Southeast University. His research interests include the Internet of Things, cyber physical systems, data science, and wireless networks and sensors.



Huijun Sun received the PhD degree in traffic and transportation planning and management from Beijing Jiaotong University, China, in 2003. She is currently a professor with the School of Traffic and Transportation Engineering, Beijing Jiaotong University. Her research interests include travel behavior analyses, travel demand management, and so on. She has published more than 60 refereed papers in leading academic journals. She was the Winner of the Education Ministry's New Century Excellent Talents Supporting Plan and Nova of Science and Technology Program in Beijing.



Zheng Dong (Member, IEEE) received the PhD degree from the Department of Computer Science, University of Texas at Dallas, in 2019. He is an Assistant Professor with the Department of Computer Science, Wayne State University. His research interests include real-time cyber physical systems and mobile edge computing. He received the Outstanding Paper Award at the 38th IEEE RTSS.



Desheng Zhang (Member, IEEE) is an Assistant Professor with the Department of Computer Science, Rutgers University. He is broadly concentrated on bridging cyber-physical systems and big urban data by technical integration of communication, computation and control in data-intensive urban systems. He is focused on the life cycle of big-data-driven urban systems, from multi-source data collection to streaming-data processing, heterogeneous-data management, model abstraction, visualization, privacy, service design and deployment in complex urban setting.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.