A Reinforcement Learning Approach to Plan Charging Stations for Shared Electric Vehicles

Qiming Ye[1,2], Prateek Bansal[1,3,*], Bryan T. Adey[1,2]

[1] Adaptive Mobility, Infrastructure and Land-use Module, FCLG, Singapore-ETH Centre, Singapore
[2] Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Zürich, Switzerland
[3] Department of Civil and Environmental Engineering, National University of Singapore, Singapore
qiming.ye@sec.ethz.ch prateekb@nus.edu.sg adey@ibi.baug.ethz.ch
* Corresponding author

Extended abstract submitted for presentation at the 12th Triennial Symposium on Transportation Analysis conference (TRISTAN XII)
June 22-27, 2025, Okinawa, Japan

March 4, 2025

Keywords: Shared electric vehicles, Charging infrastructure planning, Reinforcement learning, Stochastic sequential decision problem, Deep deterministic policy gradient

1 INTRODUCTION

The massive adoption of shared electric vehicles (SEVs) necessitates the availability of publically accessible fast chargers, compensating for SEVs' onboard battery capacity limitations and supporting SEVs' opportunistic charging needs. Using these fast chargers, SEVs can quickly restore their SoC levels, ensuring minimal interruption to their working schedules (Cilio & Babacan, 2021). Planning fast charging stations over multiple years in response to evolving charging demand patterns and induced demand due to the improved charging service is challenging (Hu *et al.*, 2024, Moniot *et al.*, 2022). Existing models predominantly used deterministic optimisation methods to address this problem, excluding temporal dimension in the decision-making (Ye *et al.*, 2024, Lokhandwala & Cai, 2020). The inherent uncertainties in ride-hailing demand were oversimplified by prior studies. Meanwhile, a large number of existing studies did not include fleet size as a decision variable in their models or modelled it deterministically.

The approach proposed in this paper improves on the state-of-the-art methods used to develop charging station deployment plans over time. Using real-world Singapore taxi data, the model achieved a 61.47% reduction in fast chargers (21,347 fewer chargers), a 496% increase in charging time per charger, and a 20.52% reduction in waiting time per request. Principal contributions include: *First*, this is the first multi-phase framework for long-term fast charging station planning. We formulated a stochastic sequential decision problem and employed the reinforcement learning method to solve it. In specific, the deep deterministic policy gradient (DDPG) algorithm, representing a model-free, policy gradient-based, and off-policy RL method, was applied. *Second*, the model incorporates fleet size decisions which evolve over time. *Third*, an agent-based model was incorporated to simulate complex SEV fleet operations involving vehicle assignment, charging dispatch, queuing for changing at the stations, and repositioning. The uncertainties considered in our model include (i) the ride-hailing demands and the induced demand, (ii) the SEV service capability for the new phase, and (iii) SEVs' charging demands in space and time.

2 METHODOLOGY

In an agent-based transport model with discrete time steps, each second is denoted as $t \in \mathrm{T}^p$, where p represents decision phases (years) $p \in \mathrm{P}$ in a planning horizon P. The candidate charging stations are

$f \in \mathrm{F}$. Let D^p represent all ride-hailing requests d in time step T^p of phase p. A transport network company (TNC) operates SEVs $v \in \mathrm{V}^p$ to serve these requests. SEVs are assigned to charging stations when stationed or when their SoC drops below a threshold (BL). The planning authority decides the number of chargers add_f^p to install at each station, with a_f^p as the rate of remaining capacity at station f allocated for new chargers in phase p, where $add_f^p = \lfloor a_f^p \cdot (cap_f^p - num_f^p) \rfloor$.

The objective function 1 maximises the total reward r_f^p at phase p, comprising two immediate rewards. The first immediate reward is the effective charging service time ratio rv_f^p. As expressed in Equation 2, it captures the time a SEV spends on charging ($chr_{v,f}^{t,p}$) compared to the time spent on the entire charging process ($que_{v,f}^{t,p} + chr_{v,f}^{t,p}$) involving queueing, where $chr_{v,f}^{t,p} \in \mathbb{N}_{\{0,1\}}$ and $que_{v,f}^{t,p} \in \mathbb{N}_{\{0,1\}}$. The second immediate reward maximises the capacity utilisation efficiency ratio rf_f^p defined in Equation 3 as the future available space (cap_f^p-num_f^p) for fast charger deployment at each station, divided by the station's total capacity (cap_f^p). Variable num_f^p indicates the cumulatively deployed chargers. Equations 4 and 5 represent the ride-hailing demand growth using a sigmoid growth curve, and the induced demand $\Delta|\mathrm{D}^p|$, which is a function of the charging service level of the preceding phase. TNC constantly calibrate the fleet size $|\mathrm{V}^p|$ based on the demand and ride-hailing completion rate g-g' pair-wise, following Equation 6. Constraint 7 defines a piecewise probability for opportunistic charging of SEVs.

Figure 1 illustrates the proposed RL-integrated planning framework: **Action** (a_f^p) represents allocated for new chargers. **State** (s_f^p) captures the total pick-up and drop-off events and the charging behaviours. **Reward** (r_f^p) follows the objective function. The RL model was trained over 200 epochs.

$$\max_{a_f^p \sim \mu} r_f^p = \mathop{\mathbb{E}}_{f \sim \mathrm{F}} \left[\sum_{k=0}^{|\mathrm{P}|-p-1} \gamma^k \left(rv_f^{p+k} + rf_f^{p+k} \right) \right] \quad \forall p \in \mathrm{P}, f \in \mathrm{F}, k \in \mathbb{N}_{[0,|\mathrm{P}|]} \tag{1}$$

$$rv_f^p = \sum_{t \in T^p} \sum_{v \in V_f^{t,p}} \frac{chr_{v,f}^{t,p}}{que_{v,f}^{t,p} + chr_{v,f}^{t,p}} \tag{2}$$

$$rf_f^p = \frac{cap_f^p - num_f^p}{cap_f^p} \tag{3}$$

$$|\mathrm{D}^p| = \frac{|\mathrm{D}^{p=|\mathrm{P}|}|}{1 + e^{-\alpha p}} + \Delta|\mathrm{D}^p| \quad \forall p \in \mathrm{P} \tag{4}$$

$$\Delta|\mathrm{D}^p| = \sum_{g \in \mathrm{I}} rnd(1,2) rv_{f=g'}^{p-1} d_{g,g'}^{p-1} \quad \forall p-1 \in \mathrm{P}, g, g' \in \mathrm{F} \tag{5}$$

$$|\mathrm{V}^p| = \beta \left(\mathrm{D}^{p-1}, \sum_{g,g' \in \mathrm{I}} \frac{com_{g,g'}^p}{com_{g,g'}^p + drp_{g,g'}^p} \right) \tag{6}$$

$$prb_v^t = \begin{cases} 1, & if \quad soc_v^t \leq BL \\ 1 - \frac{soc_v^t - BL}{BP - BL}, & otherwise \end{cases} \tag{7}$$

3 RESULTS

Figure 2(A) shows the relationship between the phase-based ride-hailing demand, SEV fleet size, and the average action decisions ($\bar{a}_f^p$) on the deployment of the fast chargers. Initially, deployment decisions are cautious, clustered around 0.012. Over time, they expand, reaching 0.099 to 0.252 by year five, reflecting a more assertive strategy to meet rising demand. This trend, also seen in Figure 2(B), indicates the model's ability to balance immediate needs with long-term growth, prioritising sustained development over rapid short-term deployment.

As charging demand grows yearly, the supply of fast chargers scales accordingly. Figure 3(A) compares natural demand growth (black line) with growth that includes induced demand (other lines), illus-

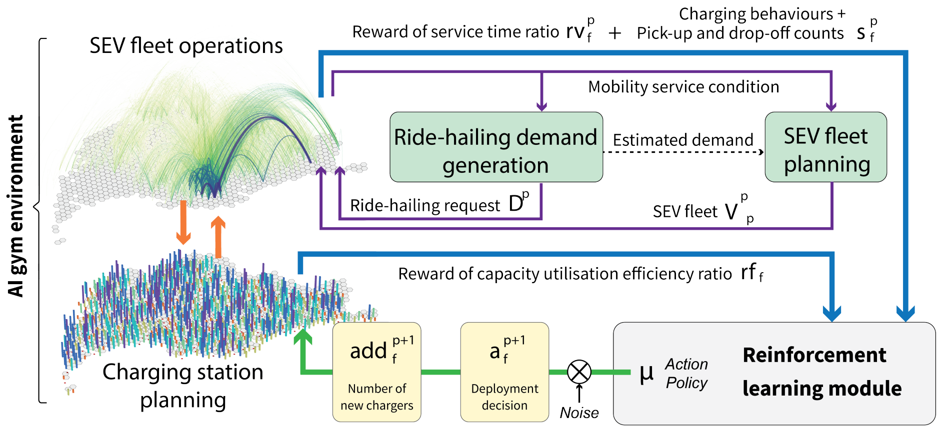

Figure 1 – *A reinforcement learning-integrated multi-phase fast-charging station planning framework*

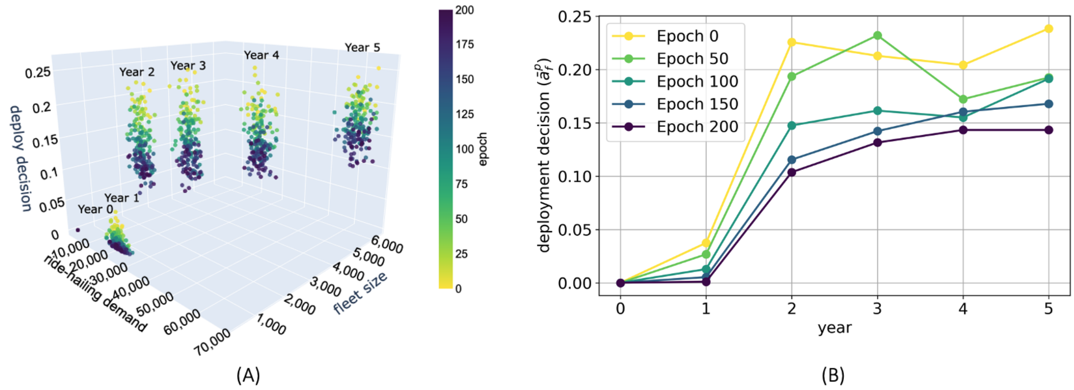

Figure 2 – *Fast charger deployment decisions per station: (A) Patterns under all ride-hailing demand and SEV fleet sizes. (B) Patterns of featured training epochs*

trating how the model captures the feedback loop between infrastructure and demand. Figure 3(B) shows that by Epoch 200, only about 60 chargers are needed in the fifth year—61.47% of the initial 98 chargers in Epoch 0—to meet similar demand levels. This reflects the model's efficiency in meeting demand with fewer resources. By reducing the number of chargers per station, the model increased charging service time per charger, rising from 353 seconds in Epoch 0 to 1,332 seconds in Epoch 200 (Figure 3(C)), indicating more intensive charger use in later epochs. Figure 3(D) shows improved queue management, reflecting a more effective allocation of fast chargers to meet demand.

Figure 4(A) shows a steady increase in total reward convergence, indicating effective optimisation of fast charger deployment. The sub-rewards convergence patterns in Figure 4(B) reveal improved efficiency, with fewer chargers deployed while maintaining a stable service time ratio. This trade-off reflects a strategy prioritising long-term scalability and flexibility, balancing efficient capacity use with operational stability.

4 DISCUSSION

This study formulated a stochastic sequential decision-making framework and employed a reinforcement learning (RL) method-integrated planning framework to approximate the optimal sequence of configurations for the locations and numbers of new fast chargers. The proposed RL framework offers a dynamic, data-driven approach to SEV charging station planning. It adapts more effectively to evolving demand and operational uncertainties than classical integer programming and robust optimisation, which rely on predefined scenarios or deployment strategies. Our scalability analysis demonstrates that the compu-

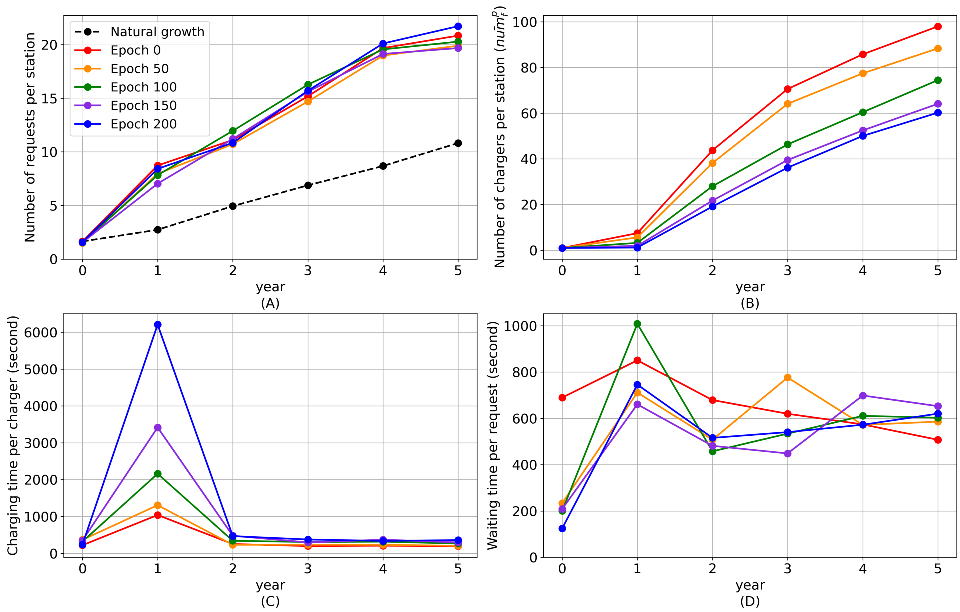

Figure 3 – *Charging station performances: (A) Charging requests per station (B) Number of chargers per station. (C) Total charging time per charger. (D) Queueing time per charging request.*

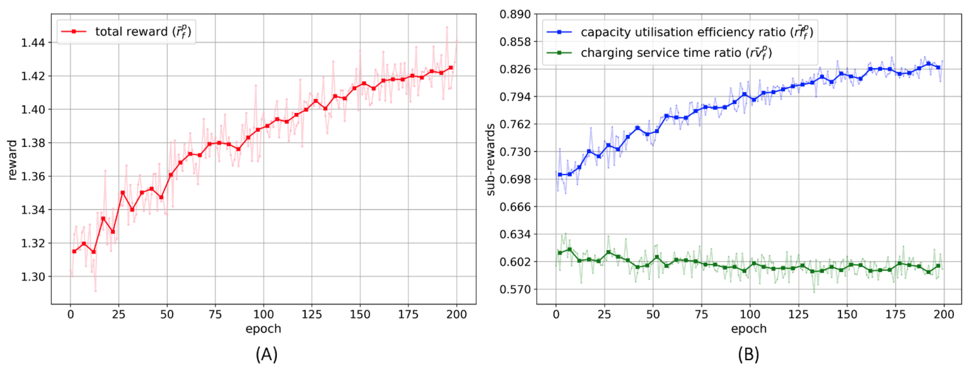

Figure 4 – *Total rewards attained per training epoch*

tational cost of the model increases with higher spatial resolution, but remains feasible for large urban applications. Simulating one epoch with a 5000m grid resolution (9 grids) using 100% of a day's Singapore taxi trip data takes 428s, while finer resolutions require 1135s for 2000m (1135 grids), 1922s for 500m (935 grids), 3362s for 200m (3362 grids), and 5717s for 100m (23,259 grids).

References

Cilio, Luca, & Babacan, Oytun. 2021. Allocation optimisation of rapid charging stations in large urban areas to support fully electric taxi fleets. *Applied Energy*, **295**, 117072.

Hu, Dandan, Huang, Liu, Liu, Chen, & Liu, Zhi-Wei. 2024. Locating and sizing charging station in multi-period to promote electric vehicles adoption in urban areas. *Energy Reports*, **11**, 3581–3598.

Lokhandwala, Mustafa, & Cai, Hua. 2020. Siting charging stations for electric vehicle adoption in shared autonomous fleets. *Transportation Research Part D: Transport and Environment*, **80**, 102231.

Moniot, Matthew, Ge, Yanbo, & Wood, Eric. 2022. Estimating fast charging infrastructure requirements to fully electrify ride-hailing fleets across the United States. *IEEE Transactions on Transportation Electrification*, **8**(2), 2177–2190.

Ye, Qiming, Bansal, Prateek, & Adey, Bryan. 2024. Multi-period charging infrastructure planning under uncertainty: Challenges and opportunities. *Sustainable Cities and Society*, 105908.