Learning to Operate an Electric Vehicle Charging Station

Considering Vehicle-grid Integration

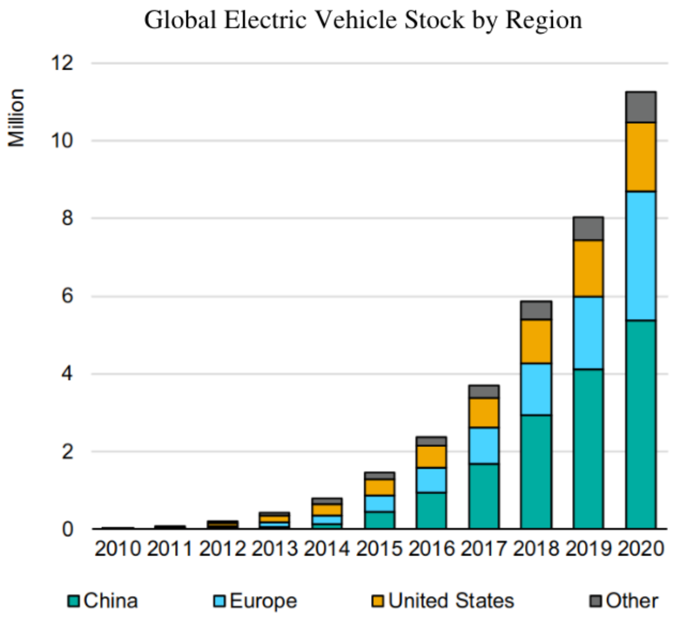
Zuzhao Ye, Yuanqi Gao and **Nanpeng Yu**
The Department of Electrical and Computer Engineering
University of California, Riverside
nyu@ece.ucr.edu

Outline

- Background and Challenges
- Literature Review
- Problem Formulation
- Technical Methods
- Numerical Study
- Conclusion

Background and Challenges

- ❑ Background
  - ➢ In the past decade, the global EV market has been growing exponentially.
  - ➢ To support continued penetration of EVs, it is critical to develop smart charging stations that could satisfy the charging needs in a cost-effective manner.
- ❑ Challenges
  - ➢ How to determine the optimal charging powers?
    - ✓ Stochastic arrival/departure of EVs and varying electricity prices
  - ➢ How to allocate the chargers if there are more EVs than chargers?
    - ✓ Not every parking spot has a charger, varying dwelling time
  - ➢ How to reduce demand charge?



Literature Review: Charging Scheduling

- ❑ Model-based methods
  - ➢ Multi-stage stochastic programming [Kim 2016] [Wang 2018]; Genetic algorithm [Domínguez-Navarro 2019]
  - ➢ Rely on sophisticated designs tailored for specific scenarios; “Curse of dimensionality”
- ❑ Learning-based methods (for a single EV)
  - ➢ Tabular Q [Dimitrov, 2014]; Kernel averaging regression functions [Chiș 2016]
  - ➢ Deep neural networks [Wan 2018]; Focus on the scheduling of a single EV.
- ❑ Learning-based methods (for multiple EVs)
  - ➢ More challenging: dimension of the state space varies with the stochastic arrival of EVs
  - ➢ Learn a collective EV fleet charging strategy [Vandael 2015]
  - ➢ Use feature functions to represent the entire station [Wang 2019] or collected features in state vector [Da Silva 2019] [Tuchnitz 2021]
- ❑ Research Gaps
  - ➢ Demand charge, vehicle-grid integration (only in single EV studies), waiting area.

Problem Formulation

- Consider a charging station with N parking spots, among which N^c are charging spots and N^w are waiting spots.
- The total profit of a charging station Z consists of three components

$$Z = B - C^p - C^l$$

B: Net revenue from charging and discharging EVs

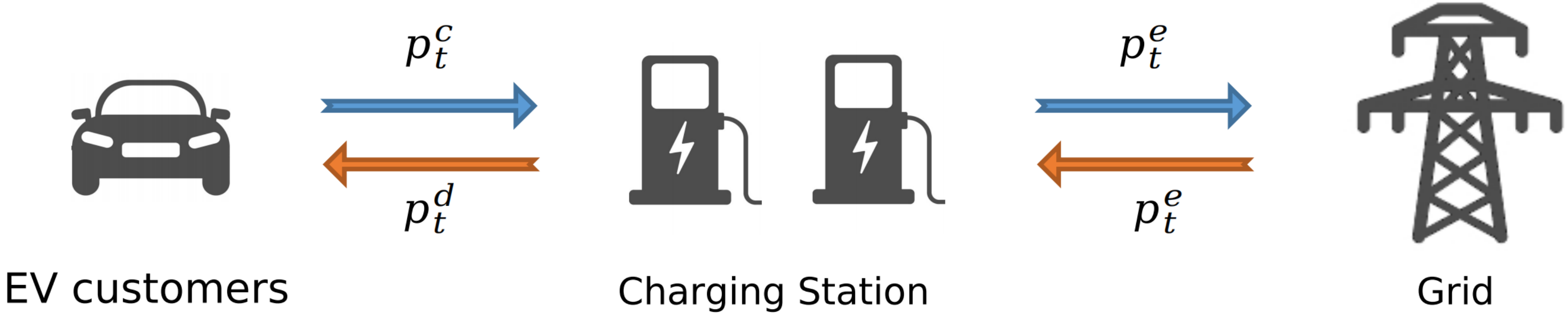
C^p: Penalty if not satisfy EV customers' energy demand upon departure

C^l: Demand charge.

Problem Formulation - Profit

- Suppose the charging power for the ith EV at time t is a_{it}.
- If $a_{it} \geq 0$, the charging station receives p_t^c and pays p_t^e (per kWh)
- If $a_{it} < 0$, the charging station pays p_t^d and receives p_t^e (per kWh)
- The net revenue $B = \sum_{i \in I} \sum_{t \in T} m_t \, |a_{it}| \Delta t$, where

$$m_t = \begin{cases} p_t^c - p_t^e & \text{if } a_{it} \geq 0 \\ p_t^e - p_t^d & \text{if } a_{it} < 0 \end{cases}$$



Problem Formulation - Penalty

- When an EV leaves without being sufficiently charged, a penalty will be imposed to the charging station to compensate the customer.

- Such a penalty shall reflect the gap between the final energy level e_i^{fnl} and the target energy level e_i^{tgt}:

$$c_i^p = \mu\left(e_i^{tgt} - e_i^{fnl}\right)^+$$

- The total penalty for a charging station is the sum of c_i^p over all EVs:

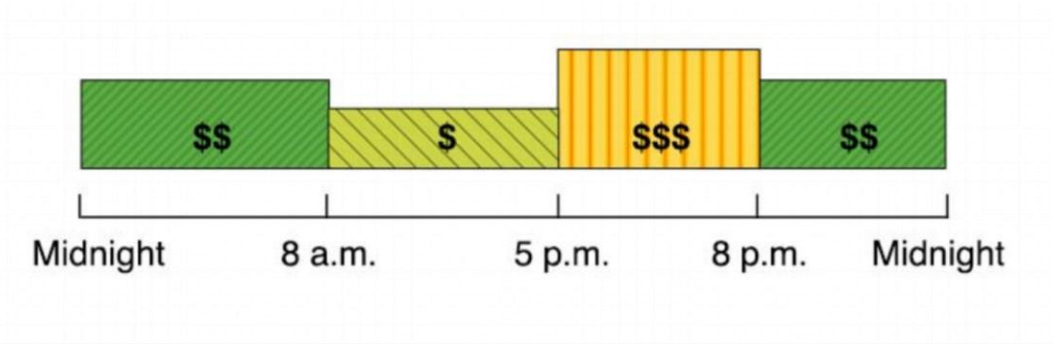
$$C^p = \sum_{i \in I} c_i^p$$

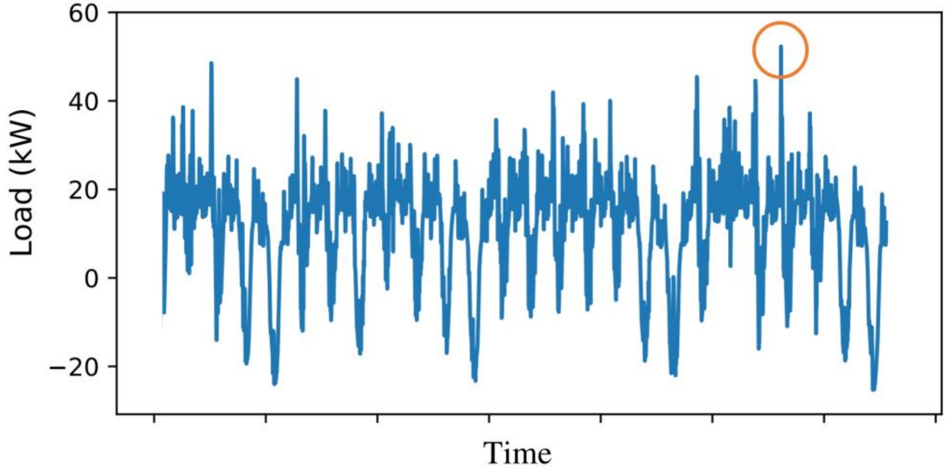
Problem Formulation – Demand Charge

- Suppose the set of time-of-use periods to be H, and $h \in H$ is one of the time-of-use periods.
- p_h^l is the price of each kW of peak demand in time-of-use period h.
- The demand charge is

$$C^l = \frac{T}{T^B} \sum_{h \in H} p_h^l \cdot L_h^+$$

$$L_h = \max_{t \in T_h} \left(\sum_{i \in I} a_{it} \right)$$

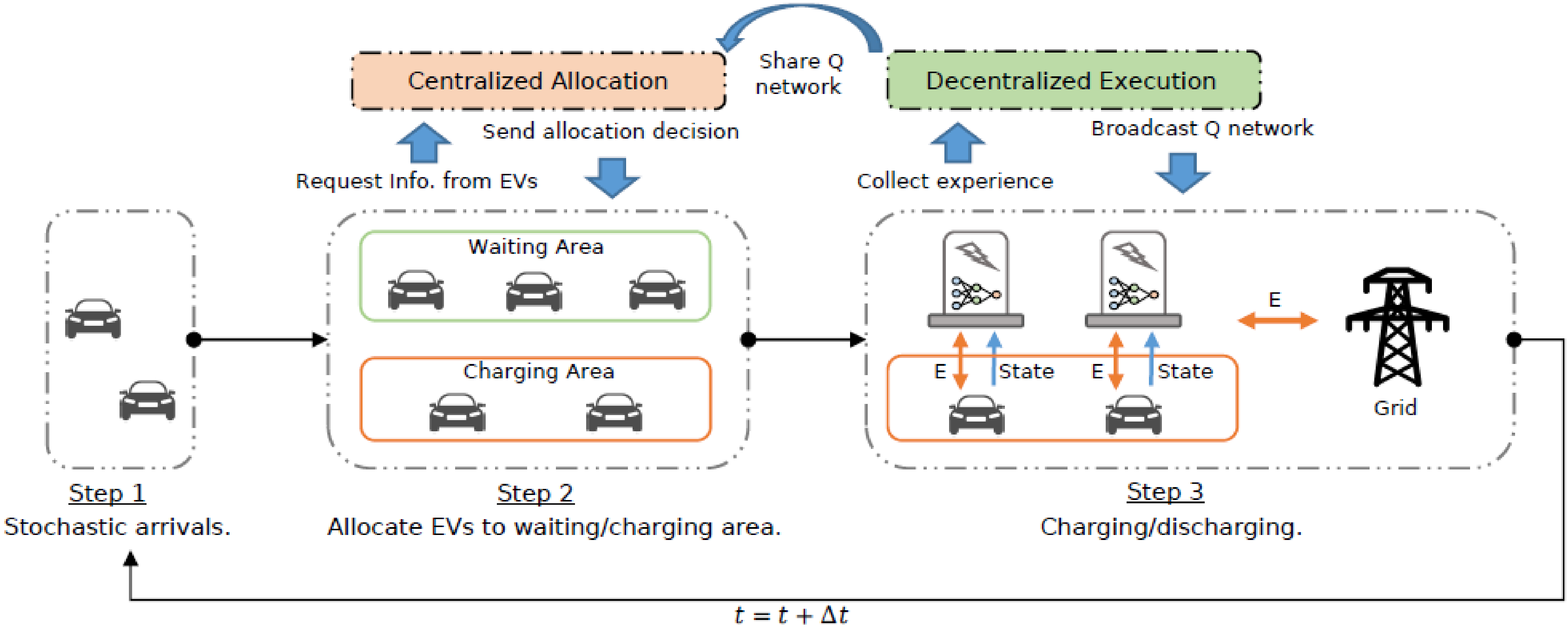




Problem Formulation - Summary

- The charging station scheduling problem is formulated as follows
  - Maximizing charging station total profit
  - Subject to
    - ✓ Power constraints, space constraints, energy constraints
- If all future information of EVs are known, an oracle is able to obtain the global optimal solution of the charging station scheduling problem.
- In practice, it is impossible to have perfect predictions of arriving EVs.
- Instead, model predictive control methods are often developed to handle similar problems.
- However, the effectiveness of MPC-based algorithm relies on accurate prediction of future EV arrivals and does not scale well with the size of the problem.

Overview of Proposed Framework: CADE



- Centralized allocation: Determine whether an EV should be allocated to the charging or waiting area
- Decentralized execution: Each individual charger makes its own decision on output power

Decentralized Execution – MDP Formulation

State Space

- For a charger j, we define its state of environment at time t as:

$$s_{jt} = \{\delta_{jt}, t, t_j^r, e_{jt}, e_{jt}^r, N_t^{EV,w}, E_t^{r,w}, \tilde{h}_t, L_{ht}\}$$

  - ✓ $\delta_{jt} \in \{0,1\}$ indicates if there is an EV connected to charger j.
  - ✓ $t_j^r = t_j^d - t$ is the remaining dwelling time of the EV connected to the charger.
  - ✓ e_{jt} denotes the current energy level of the EV battery.
  - ✓ $e_{jt}^r = e_j^{tgt} - e_{jt}$ is the remaining energy to be charged for the EV battery.
  - ✓ $N_t^{EV,w}$ is the total number of EVs in the waiting area at time t.
  - ✓ $E_t^{r,w}$ is the sum of remaining energy to be charged for all EVs in the waiting area.
  - ✓ $\tilde{h}_t$ is a one-hot encoded vector indicting the current time-of-use period.

MDP Formulation – Action Space

Action

- The action for a charge j at time t is its output power a_{jt}.
- The upper bound of the power $a_{jt}^{upper} = \min\left(a^{max}, \frac{e^{max} - e_{jt}}{\Delta t}\right)$
- The lower bound of the power $a_{jt}^{lower} = \min\left(a^{min}, \frac{e^{min} - e_{jt}}{\Delta t}\right)$
- The feasible action space $A_{jt} = \left\{a_{jt}^{lower}, \ldots, a_{jt}^{upper}\right\}$ with uniform difference Δa between adjacent actions

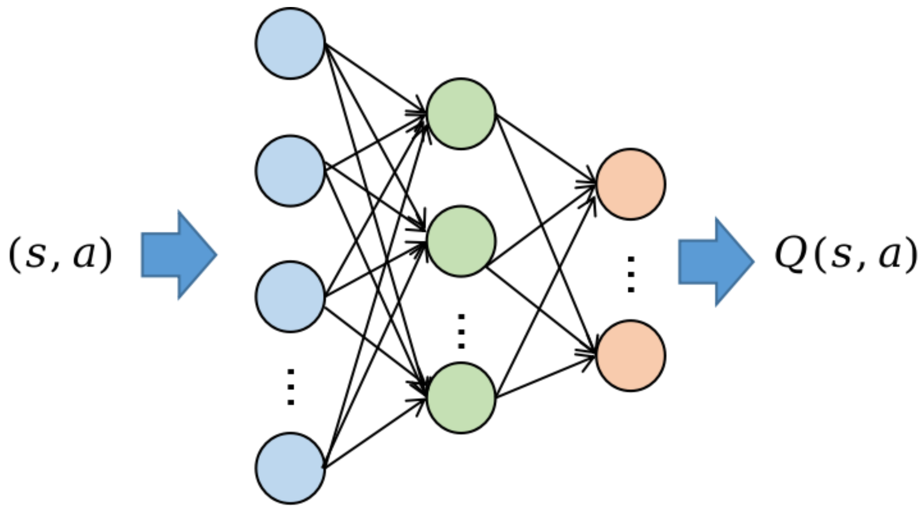
MDP Formulation – Reward

- The reward of a charger: $r_{jt} = r_{jt}^{b} + r_{jt}^{p} + r_{jt}^{l}$
  - Net revenue: $r_{jt}^{b} = m_t \left| a_{jt} \right| \Delta t$
  - Penalty: $r_{jt}^{p} = -\left(c_{jt}^{p} + c_{jt}^{p,w} \right)$
    - ✓ From charging area: $c_{jt}^{p} = \begin{cases} \mu \left(e_j^{tgt} - e_{jt} \right)^{+} & \text{if } \ t^r = 0 \\ 0 & \text{if } \ t^r > 0 \end{cases}$
    - ✓ From waiting area: $c_{jt}^{p,w} = \frac{a^{max} - a_{jt}}{\sum_{j \in J} \left(a^{max} - a_{jt} \right)} R_t^{p,w}$
  - Demand charge: $r_{jt}^{l} = \frac{a_{jt}}{\sum_{j \in J} a_{jt}} R_t^{l}$

Decentralized Execution – Policy Improvement

- Action Value Function

  - The action value functions of the chargers will be learned through a deep Q learning (DQN) algorithm
  - Operation experiences are shared among all chargers
  - Use a deep neural network $Q(s, a; \theta)$ to approximate the action value function and a second one $Q(s, a; \theta^{-})$ to stabilized the learning process



- The parameters are updated by stochastic gradient descent on the loss function:

$$L(\theta) = \mathbb{E}\left[r + \gamma(1 - d) \max_{a' \epsilon A} Q(s', a'; \theta^{-}) - Q(s, a; \theta)\right]^2$$

- The optimal deterministic policy chooses action $a^* = arg \max_{a \in A} Q^*(s, a; \theta)$

Centralized Execution

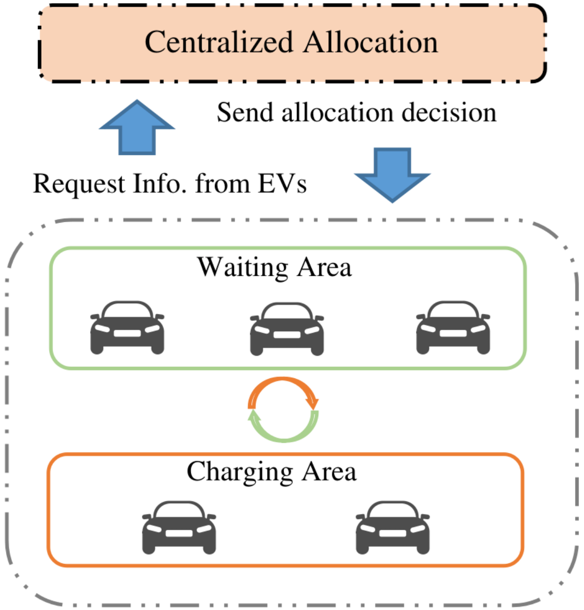
- Connect the kth EV to a charger will create an action value $q_k^c = \max_a Q(s(\alpha_k = 1), a)$.
- When an EV is parked in a waiting spot, it is equivalent to be connected to a charger with zero power output and create an action value $q_k^w = Q(s(\alpha_k = 1), a = 0)$.
- Action values represent expected returns.
- The allocation problem can be solved by finding the EV allocation that maximizes the summation of the action values:

$$\max_{\alpha_k} \sum_{k=1}^{N_t^{EV}} \alpha_k q_k^c + (1 - \alpha_k) q_k^w$$

Subject to:

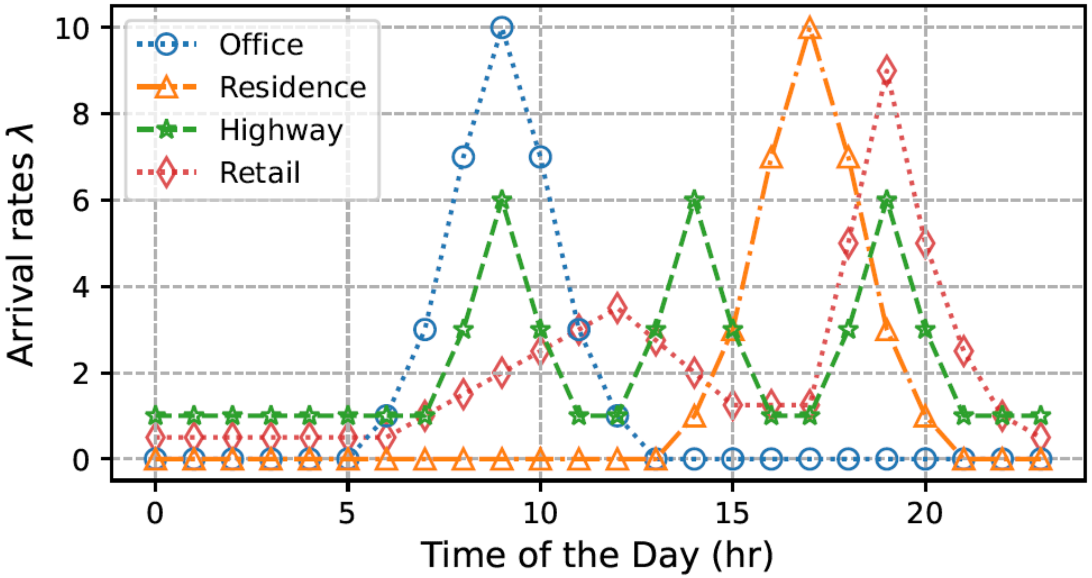
$$\sum_{k=1}^{N_t^{EV}} \alpha_k = \min(N^c, N_t^{EV}),$$

$$\alpha_k \in \{0,1\}, \qquad \forall k \in \{1,2, \dots, N_t^{EV}\}$$



Numerical Study Setup

- The EV arrival patterns
  - Office, residential area, highway, retail stores
- The proposed method does not assume prior knowledge of the specific arrival patterns
- There are three-time-of-use periods in a day
  - On-peak, mid-peak, off-peak
- Set $p^d > p^c$ to compensate the EV customers for battery degradation



Electricity Prices

Time-of-use periods (h)	On-peak 12PM-5PM	Mid-peak 8AM-12PM	Off-peak 5PM-8AM
p^c (\$/kWh)		0.15	
p^d (\$/kWh)		0.16	
p^e (\$/kWh)	0.20	0.10	0.05
p^l (\$/kWh)	2.0	1.0	0.5

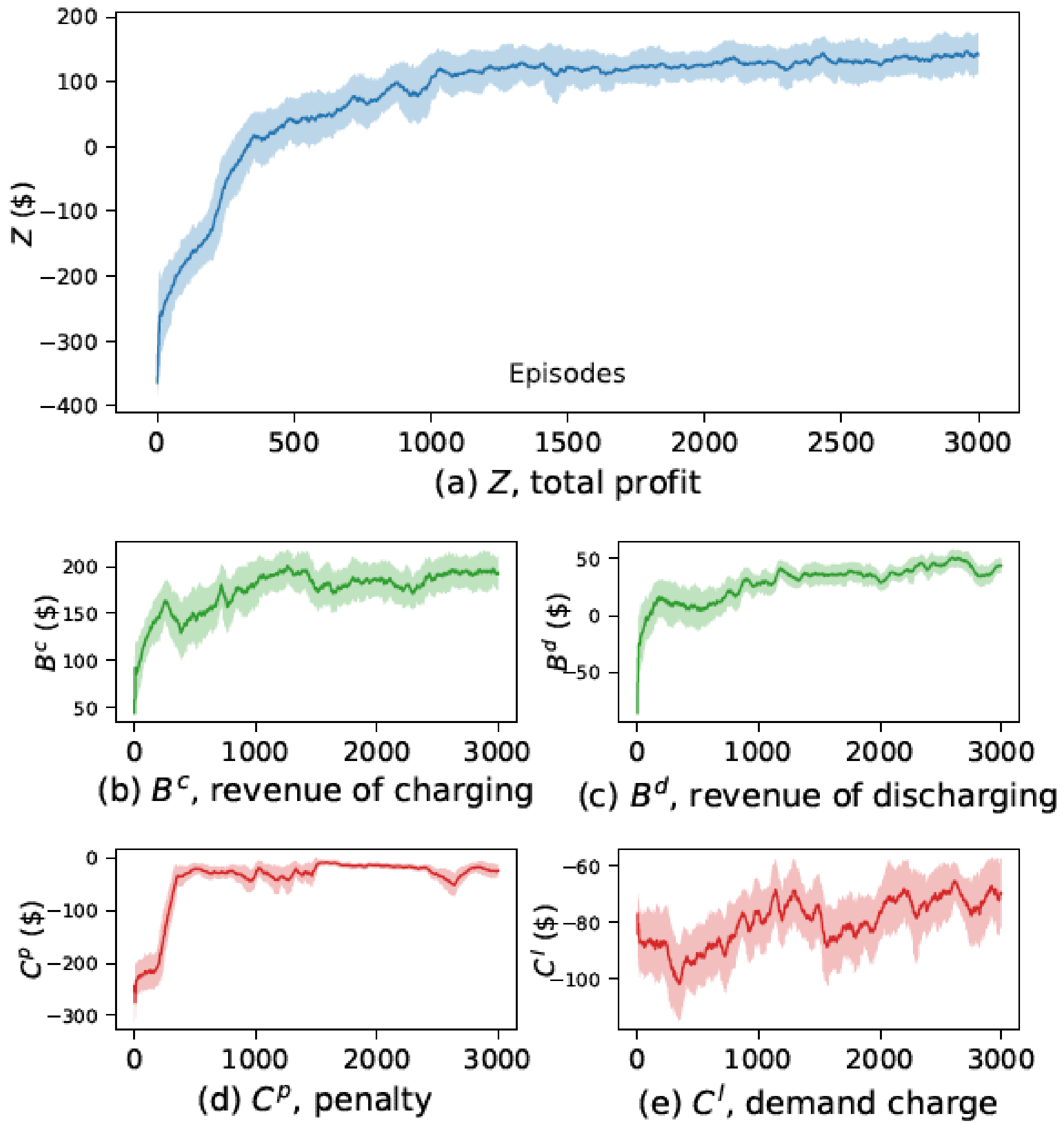
Algorithmic and Hardware Setup

- The neural network parameters
  - Hidden layer size: (256,128,64)
  - Learning rate: 0.01 and scheduled to decrease
  - Discount factor: 1
  - Target network update frequency: 25 episodes
- Hardware
  - CPU: AMD Ryzen 7, 8-socre
  - GPU: NVIDIA RTX 2080Ti
  - Memory: 16 GB

Parameter	Value	Unit
Number of chargers	$10 \sim 100$	-
Number of waiting spots	$5 \sim 50$	-
Battery energy limits	10/100	kWh
Charging power limits	-100/100	kW
Step change in charging power	1	kW
Penalty coefficient	0.2	$/kWh
Time horizon	48	Hour
Time interval	15	Minute
Billing period	30	Day
Initial battery energy level	$\sim N(20, 10^2)$	kWh
Target battery energy level	$\sim N(80, 10^2)$	kWh

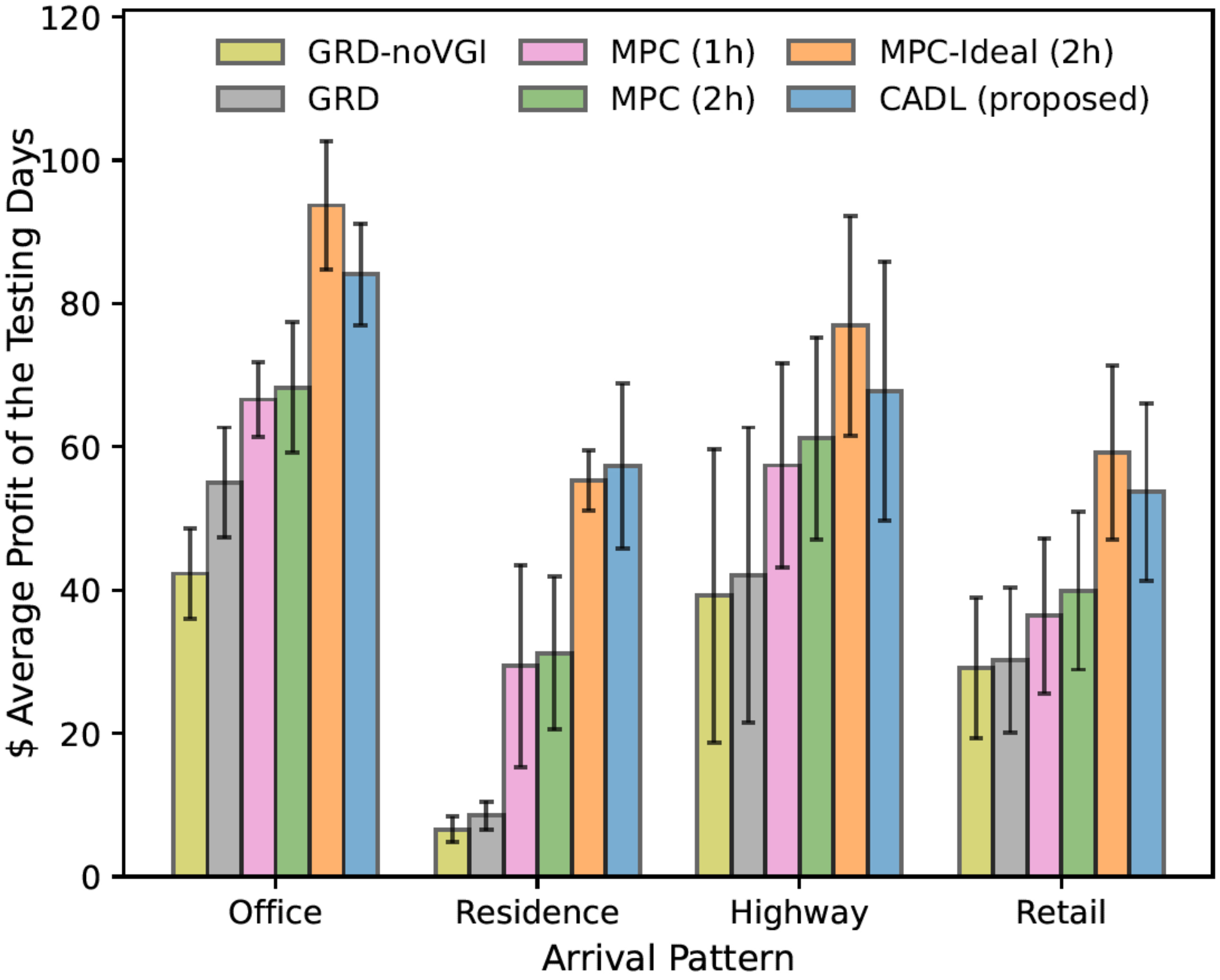
Learning Efficiency

- Learned a good policy at around 1,000 training episodes.
- The net revenues of charging and discharging are improved significantly.
- The penalty term gradually reduces to zero.
- The reduction of demand charge is relatively limited.



Profit Comparison with Baseline Methods

- Selected baseline methods
  - GRD-noVGI; GRD, MPC (1h and 2h look-ahead), MPC-ideal
- GRD (greedy) methods assign EVs based on urgency of demand
- MPC (model predictive control) with forecasted EV arrival
- Proposed framework outperforms all baseline methods except MPC-ideal (with an unfair advantage of having the perfect knowledge of the future)



Scalability Analysis

- The computation time of the propose CADE framework is much shorter than that of the MPC-based method.

Table: Comparison of computation time per step

Station size (N^c/N^w)	10 / 5	20/10	50/25	100/50
CADE	0.7 ms	0.8 ms	1.1 ms	1.7 ms
MPC (1h)	3 s	8 s	313 s	483 s
MPC (2h)	19 s	42 s	506 s	935 s

Conclusion

- A centralized allocation and decentralized execution (CADE) framework is developed to operate a charging station considering waiting area, demand charge, and vehicle-grid integration.
- The key contribution is the synergistic combination of reinforcement learning algorithm with optimization method.
- Comprehensive numerical studies with different EV arrival patterns show that the proposed CADE framework outperforms state-of-the-art baseline algorithms.
- The scalability analysis shows that the CADE framework is more computationally efficient than the baseline model-based control algorithm.

Contact Information

Nanpeng Yu, nyu@ece.ucr.edu

Thank You

- Financial support: University of California, Office of the President and Institute of Transportation Studies

- UCR Team Members: Zuzhao Ye and Yuanqi Gao

- Publication: Z. Ye, Y. Gao, and N. Yu, "Learning to operate an electric vehicle charging station considering vehicle-grid integration," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, July, 2022.