Smart Charging of Electric Vehicles

Oskar Triebe (triebe@stanford.edu), Will McCloskey (mcclow12@stanford.edu), Justin Luke (jthluke@stanford.edu)

Summary

Smart charging of electric vehicles (EVs) can support renewables integration and reduce costs, while providing quality charging services to customers.

- We contribute a new open-source OpenAI Gym
- We train and evaluate several reinforcement-learning algorithms. The current status-quo controller is a strong local minimum. Cost and charge completion are to be traded off. Cost is a proxy for renewables integration.

Best overall: Linear policy gradient algorithm

- [single station] Reduce charging cost by 18% at expense of 7% lower charge completion.
- [multi-station with transformer limit] Achieved 6% higher charge completion at 4% higher cost.

Methods

EV Charging OpenAI Gym Environment

- Simulates charging facility using historical charging and price data
- Plug-and-play with different controllers
- Highly customizable environment configuration

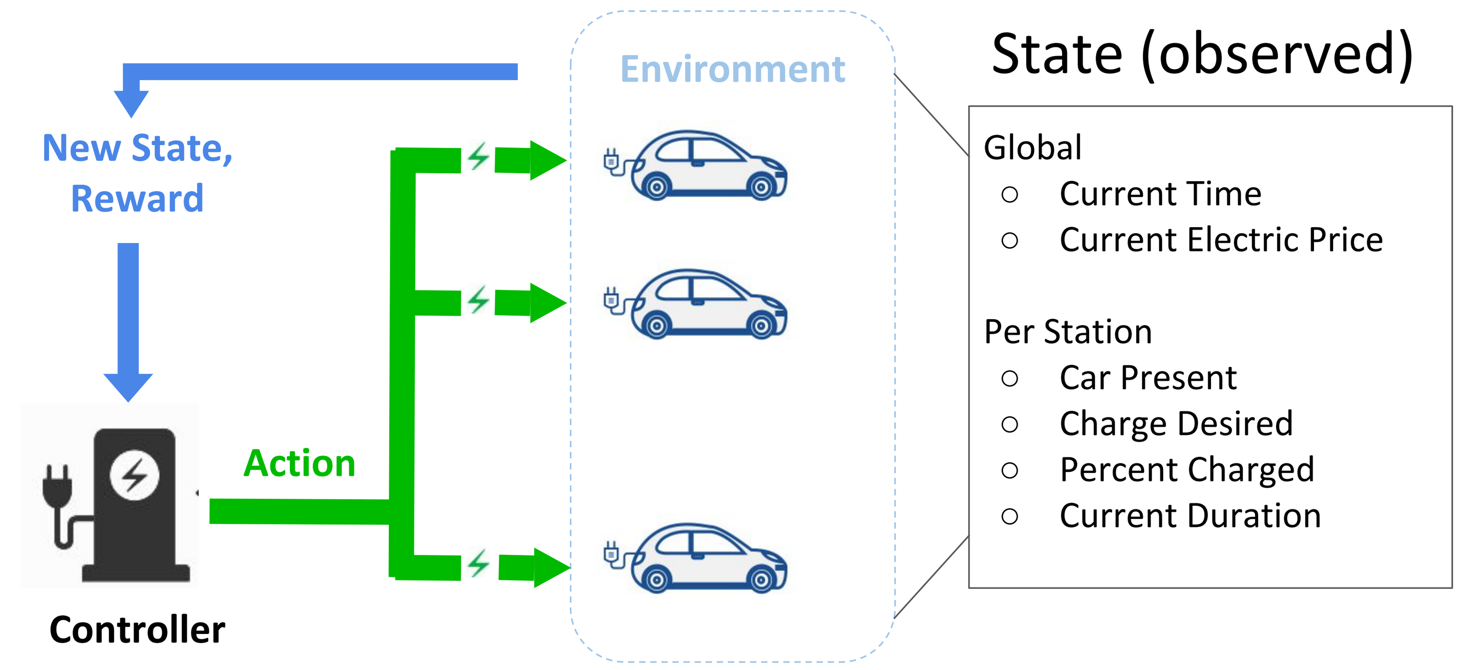
Reward components:

$$R(t) = \vec{w}^T (r_{charge}(t), -r_{cost}(t), -r_{CT}(t))$$

$$r_{charge}(t) = \frac{r_{max}}{t \cdot TC} \cdot \sum_{i=1}^{N} I_i(t) E_i(t)$$

$$r_{cost}(t) = \frac{r_{max}}{t \cdot TC} \cdot price(t) \cdot \sum_{i=1}^{N} E_i(t)$$

$$r_{TC}(t) = \exp\left(\log r_{max} \cdot \min\left[1, \ \frac{1}{TC} \cdot \max\left(0, \ \sum_{i=1}^{N} P_i(t) - TC\right)\right]\right)$$



Models

We designed a controller framework to train and evaluate models:

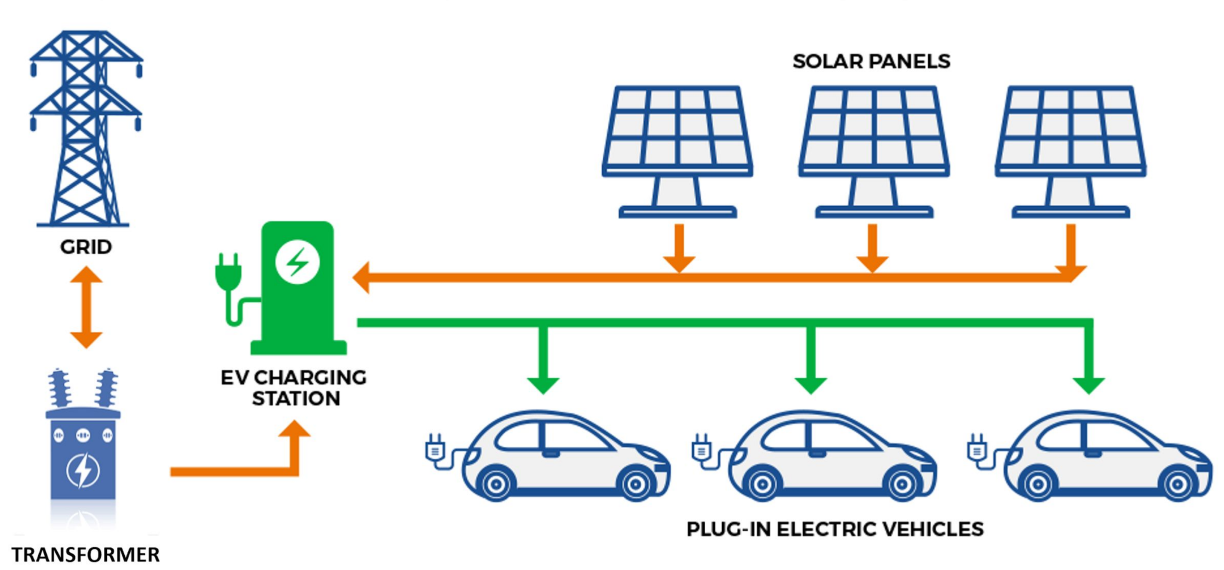
- Baseline (status quo) charges until full
- Deep (Double) Q-learning
- Deep (Double) SARSA
- Policy Gradient with discrete/continuous actions and long or short training schema
- NN sizes: linear (), nano (64, 32, 16), small (128,128,64), default (512, 512, 256)

We set charge reward to be delayed until departure, cost penalty to be immediate and gamma 0.9 to 1.

Background

Motivation

- World: 3 million EVs today → 125 million by 2030 (IEA)
- CA: 5 million EVs by 2030, 100% renewables by 2045
- Problem: EV load peaks problematic for electric grid
- Goal: Shift EV load to peak solar (via cost minimization)
- Challenges: Unknown car departure. Respect facility's transformer limit. Designing reward function appropriately.
- No past work in Deep RL for smart charging with real data



Data

Training Setup

- Size: 1,250,000 charging sessions (80% train, 20% eval)
- Episode length: session duration (Single); 3 days (Multi)
- Episode randomly sampled (fixed seed)

Characteristics

- Workplace stations in the Bay Area (2016 to 2018)
- Use data from busy stations at ZIP 95014.
- Station ID, start time, energy charged (e) and duration (d)
- Clean: 5 min < d < 1 day ; 1.0 kWh < e < 100 kWh.
- Price data (simulated): Inverse to solar PV supply

Results

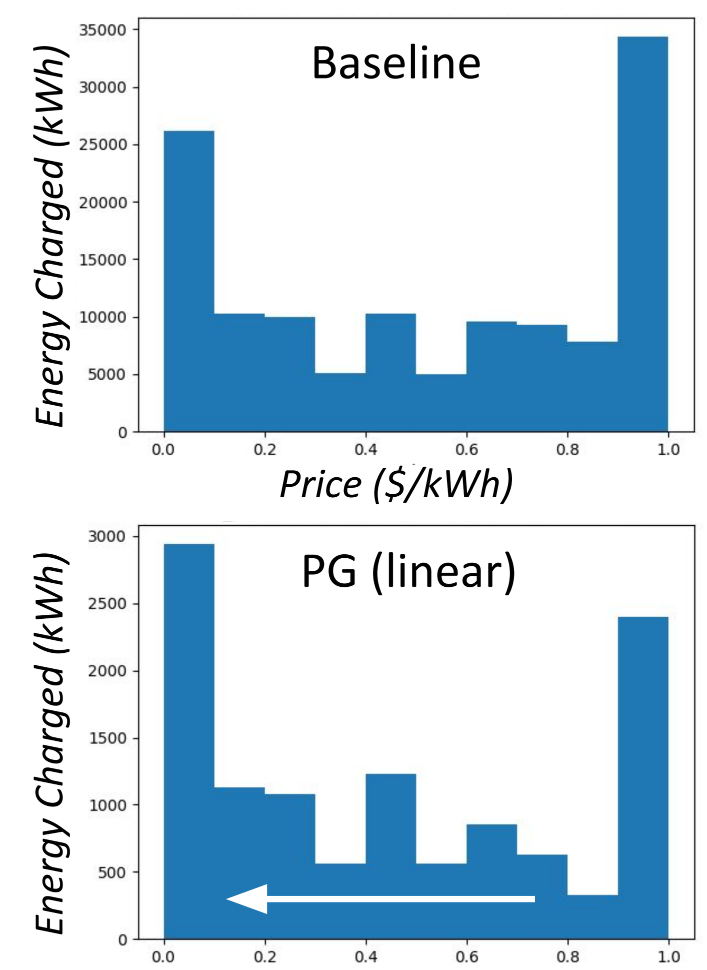
Single Station Environment

Controller	Reward	Energy (total %)	Cost (per kWh)
BaselineZero	0.000 ± 0.0	0.000	0.000
BaselineOne	3.650 ± 0.1	**1.000**	0.565
Random	2.888 ± 0.1	0.784	0.557
LinearQN	3.372 ± 0.1	0.855	0.447
DeepQN_{nano}	3.329 ± 0.1	0.809	0.375
DeepQN_{small}	3.357 ± 0.1	0.824	0.390
DeepQN	3.445 ± 0.1	0.888	0.478
PG_{linear}	$\mathbf{3.668} \pm 0.1$	0.938	0.460
PG_{nano}	3.650 ± 0.1	1.000	0.565
$\text{PG}_{nano-long}$	3.510 ± 0.1	0.850	0.367
PG_{small}	3.650 ± 0.1	1.000	0.565
PG	3.650 ± 0.1	1.000	0.565
PGC_{linear}	3.589 ± 0.1	0.954	0.520
PGC_{nano}	2.902 ± 0.1	0.726	0.437
$\text{PGC}_{nano-long}$	2.794 ± 0.1	0.639	**0.262**
PGC_{small}	3.650 ± 0.1	1.000	0.565
PGC	3.650 ± 0.1	1.000	0.565
Q_{MLP}	2.685 ± 0.1	0.723	0.551
Sarsa_{MLP}	2.690 ± 0.1	0.730	0.564
$\text{Q}_{DoubleMLP}$	1.636 ± 0.1	0.460	0.620
$\text{Sarsa}_{DoubleMLP}$	0.818 ± 0.1	0.213	0.505

Multi Station Environment

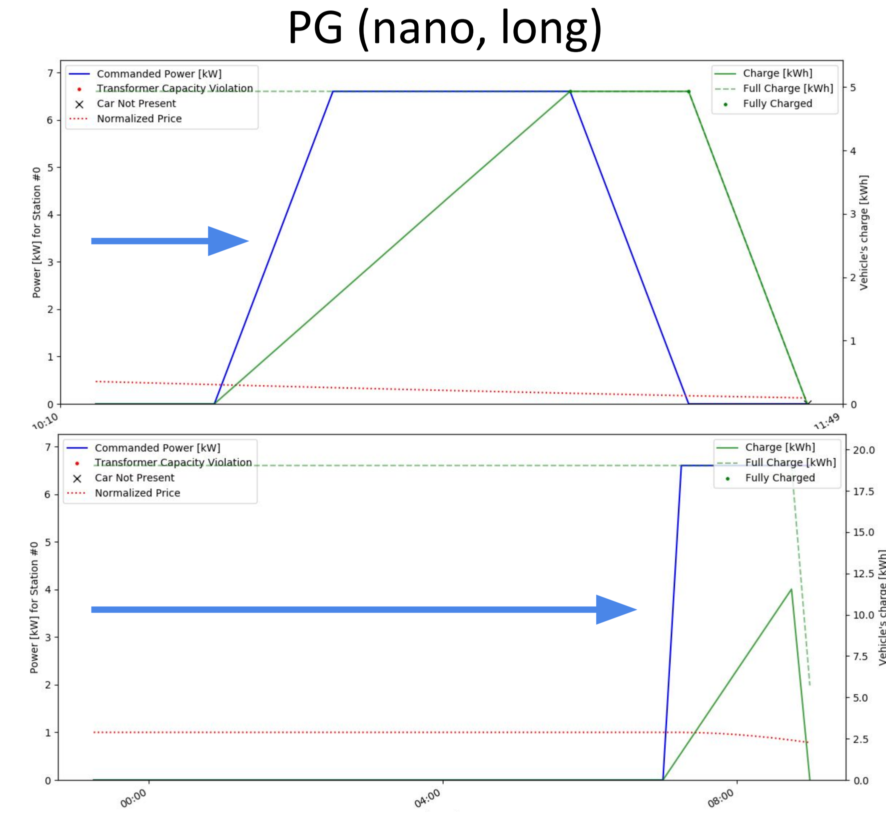
Controller	Reward	Energy (total %)	Cost (per kWh)
BaselineZero	0.000 ± 0.0	0.000	0.000
BaselineOne	32.43 ± 0.3	0.794	0.589
Random	26.64 ± 0.3	0.653	0.587
LinearQN	15.24 ± 0.2	0.324	**0.381**
DeepQN_{small}	13.21 ± 0.2	0.314	0.551
PG_{linear}	$\mathbf{33.75} \pm 0.4$	**0.844**	0.615
PG_{small}	21.32 ± 0.3	0.540	0.615

Charging Cost distribution



Controller Behavior

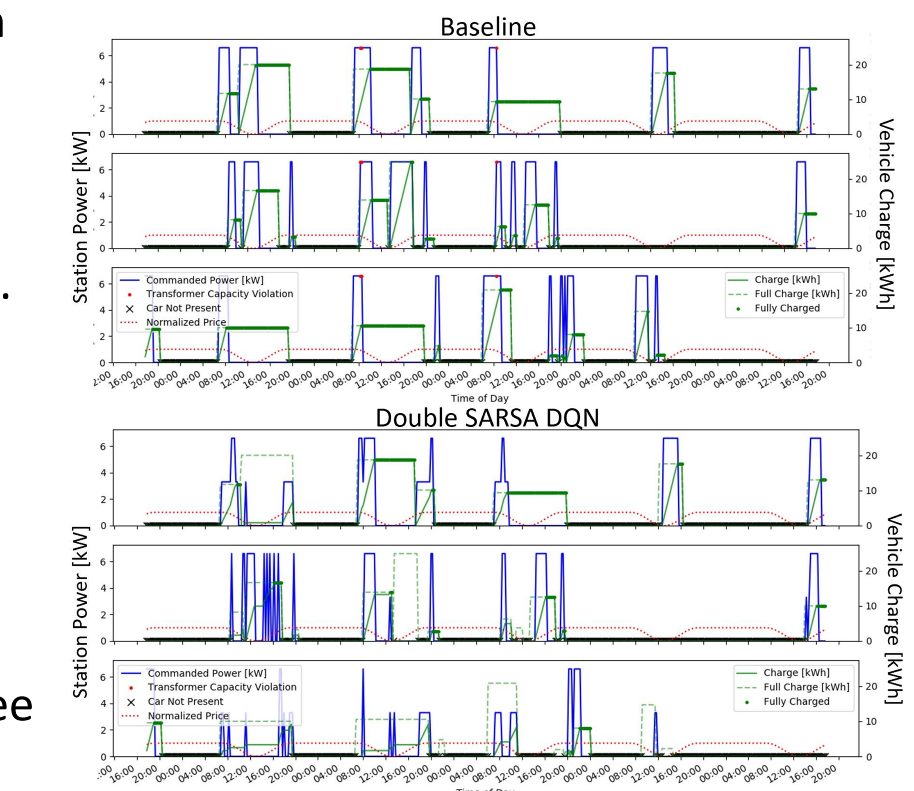
PG (nano, long)



Discussion

As shown in Table "Single Station Environment", most algorithms do better than the baseline in terms of price. An example is shown in Figure "Cost Distribution". However, this comes at a trade-off with the total energy supplied. Delaying charge for better prices runs the risk of the car leaving before being fully charged. An example of this trade-off can be seen in Figure "Controller Behavior". PG Linear has the greatest average reward and balance between the two objectives in both single and multi station settings.

The controllers tested have also shown success in managing peak demand on the facility transformer. As shown on the figure on right, Baseline violates the 14kW transformer capacity whereas Double SARSA DQN has learned to distribute power between the stations over time. A challenge for future work aims to achieve all three objectives satisfactorily.



Conclusion

Our results show reinforcement learning based charging of EVs has potential to increase renewables integration and prolong lifetime of grid assets. Unlike prior work, this project jointly leverages the capabilities of deep-learning to extract higher-order features and is validated on a large-scale EV charging dataset. Future work will explore multi-agent frameworks, the benefits of coordination with on-site stationary storage and local solar PV arrays, and vehicle-to-grid charging scenarios.

Acknowledgments: We thank the SLAC GISMo Lab for providing access to the dataset and our great CS 234 teaching staff for their mentorship.