Transactive Electric Vehicle Agent: A Deep Reinforcement Learning Approach

Swastik Sharma, Swathi Battula, Sri Niwas Singh

{swastiks21, swathi, snsingh}@iitk.ac.in

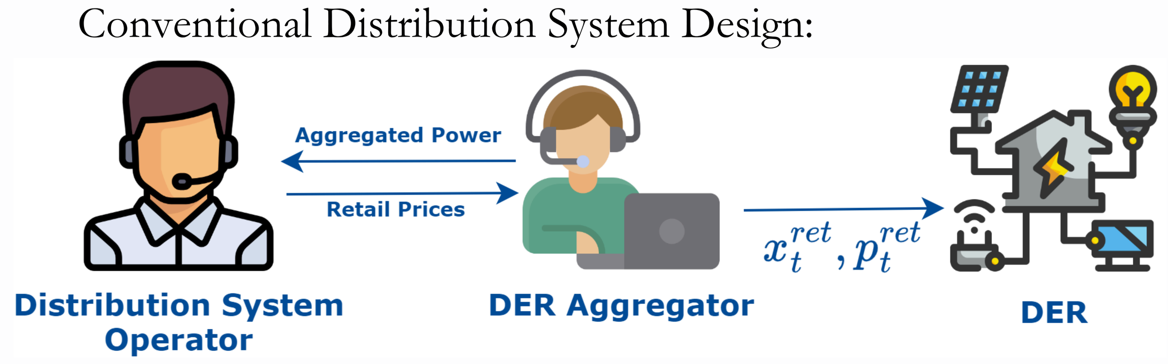
Indian Institute of Technology Kanpur, India, 208016

Paper ID: 24PESGM0011

MOTOVATION

- Research on Transactive Energy Systems (TES)-based distribution system design is gaining traction due to its numerous advantages, including fair compensation for participating resources.
- Few studies focus on bid-based TES designs where Distributed Energy Resources (DERs) can express their goals and constraints through price- and quantity-based bids/offers [1].
- Existing literature often derives bids/offers using model-based stochastic optimization models with several limitations.
- **When deriving price-sensitive bids/offers from DERs in a model-based derivation, inherent non-linearities in the problem are typically linearized.**
- **Optimization techniques are often time-intensive, hindering their real-time decision-making and large-scale simulation application.**
- **Establishing retail prices in a bid-based TES design for distribution systems is complex. Predicting these prices is challenging for an agent due to their dependence on various factors, including DER agents' bids/offers, which are intricately linked to the market clearing process.**

Conventional Distribution System Design:



Bid-based TES-based Distribution System Design:

Research Contributions:

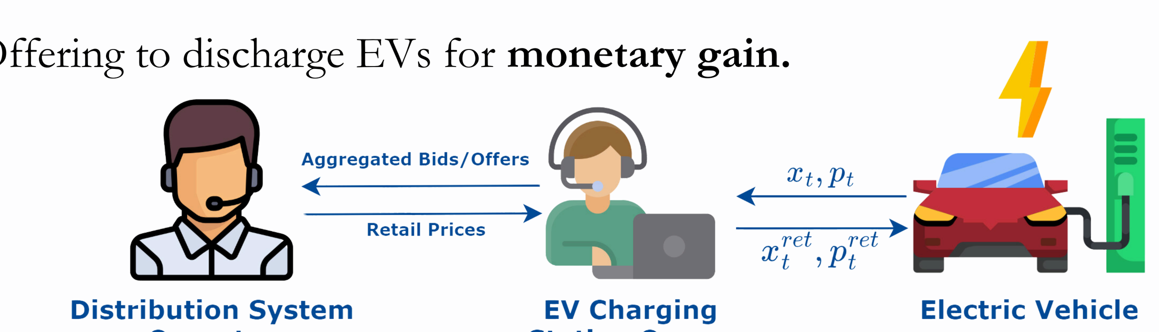
- **Designing a model free Deep Reinforcement Learning (DRL) based Electric Vehicle (EV) agent to bid/offer on behalf of the EV customer.**
- **Modelling the problem with Partially Observable Markov Decision Process (POMDP) based formalism for robust decision making under uncertainty.**
- **Assessment of the proposed agent on real world data.**

PROBLEM DESCRIPTION

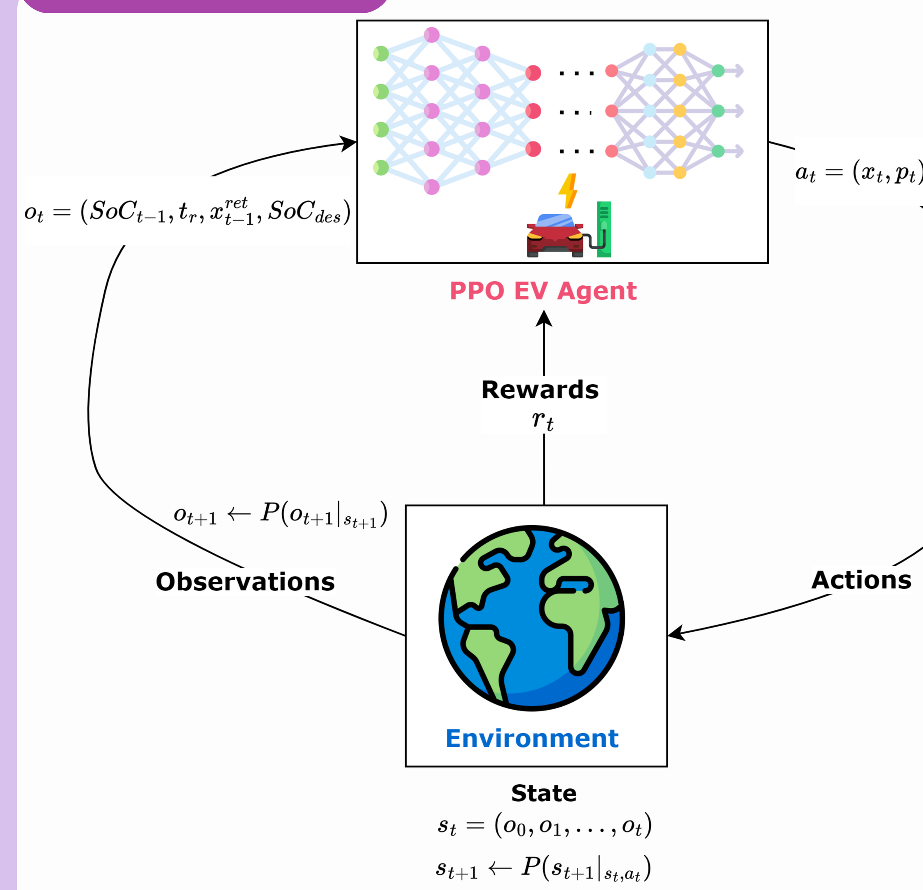
Bids/offers placed by the DRL-based EV agent go through a hierarchical aggregation aggregated by the EV Charging Station first and then by DSO to participate in the WPM.

The aim is to build a DRL-based EV agent which can bid/offer on behalf of the customer in a bid-based TES design while meeting the customer's goals subject to constraints.

- Placing bids to achieve a specified **desired SoC** by the end of the plug-in period.
- Offering to discharge EVs for **monetary gain.**

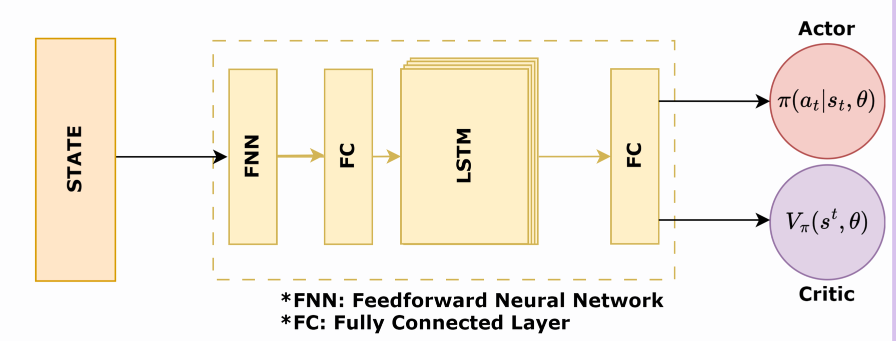


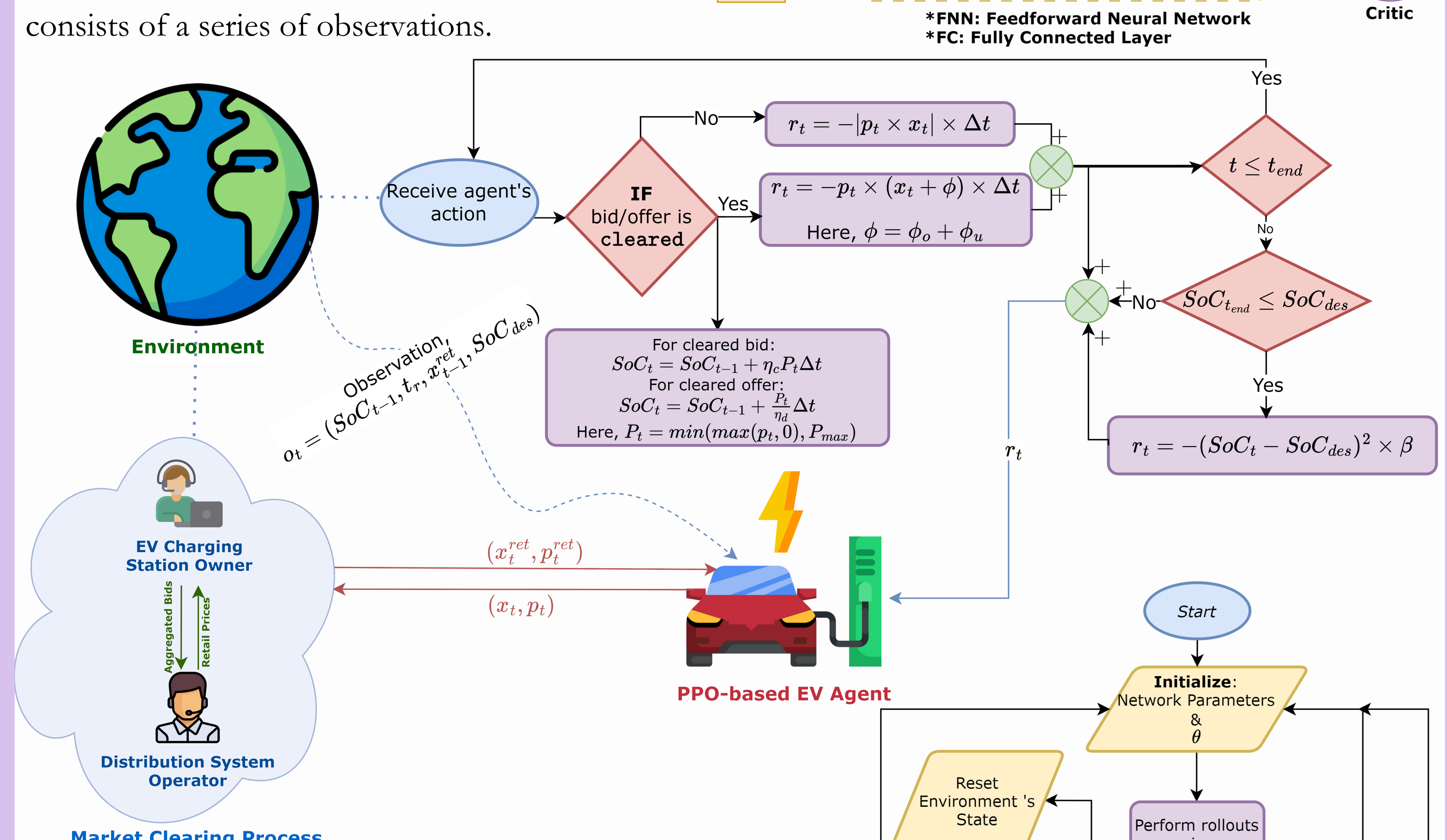
METHODOLOGY



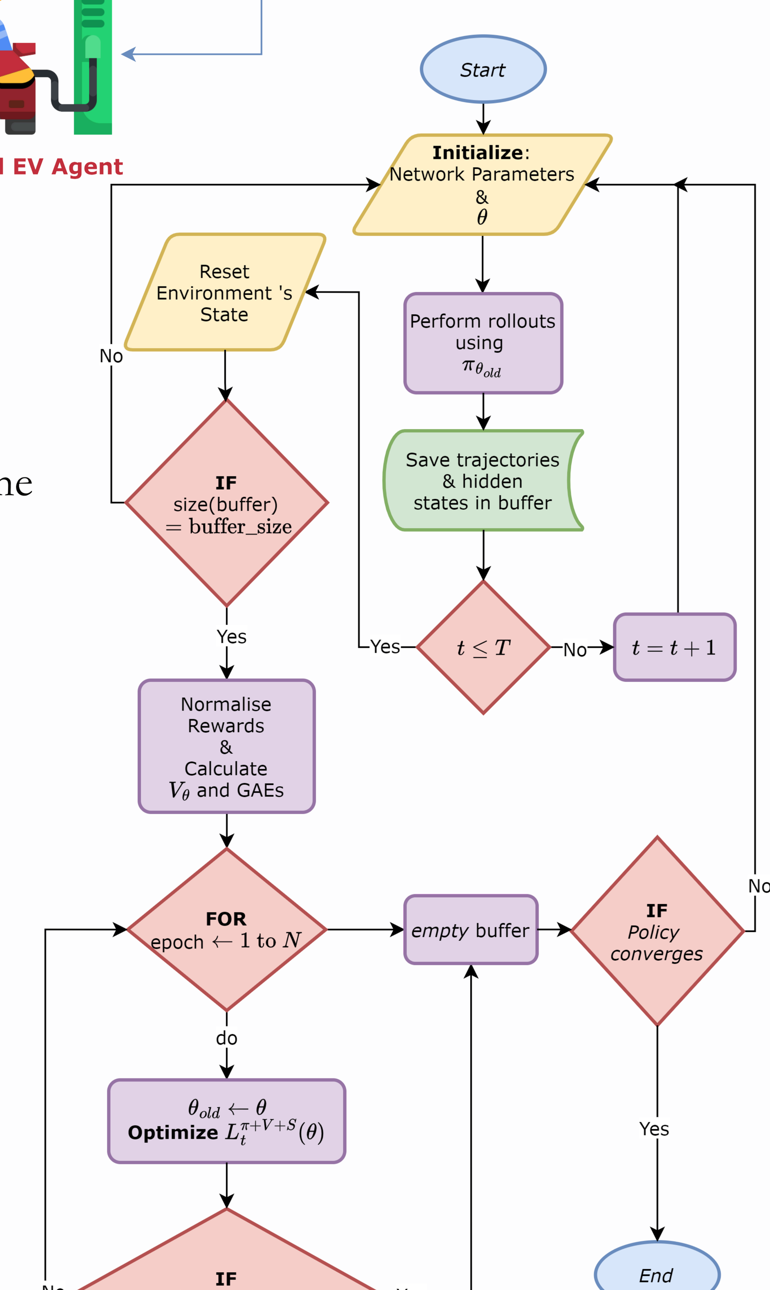
We employ a policy gradient algorithm, Proximal Policy Optimization (PPO), to manage the continuous and high-dimensional state space. PPO utilizes an actor-critic architecture, where the actor determines the actions to be taken, and the critic evaluates the performance of these actions. PPO is highly sample-efficient and can help the agent learn near-to-optimal policy with lesser interactions with the environment compared to the other state-of-the-art DRL algorithms.

For the neural network architecture, we use Long Short-Term Memory networks (LSTMs). Their usage helps preserve information about long-term dependencies, enabling the extraction of valuable insights from historical data, which consists of a series of observations.





The adjacent flowchart depicts the working of the proposed PPO algorithm:



For performance enhancement of PPO algorithm, we use the following to increase the learning efficiency of the agent since the environment is complex and non-stationary:

1. **Early Stopping based on Kullback-Leibler (KL) Divergence:**

$$D_{KL}(\pi_{\theta_{old}}|\pi_\theta) \leq \beta_{KL}$$

2. **Normalisation of Rewards:**

$$r_i' = \frac{r_i - r_\mu}{r_\sigma + \epsilon_r}$$

RESULTS

The algorithm specific and environment related parameters are given below:

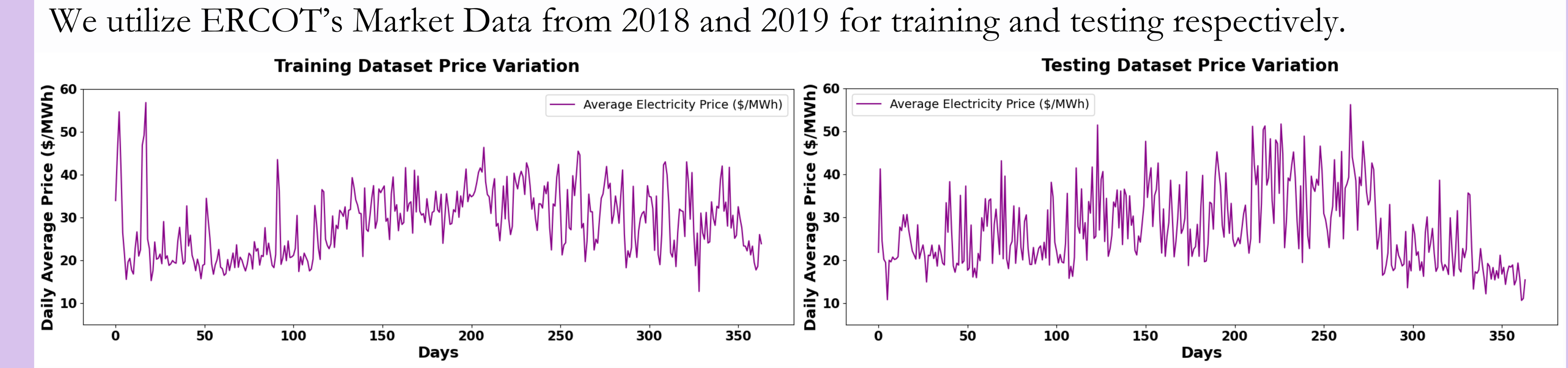
PPO Hyperparameter Settings

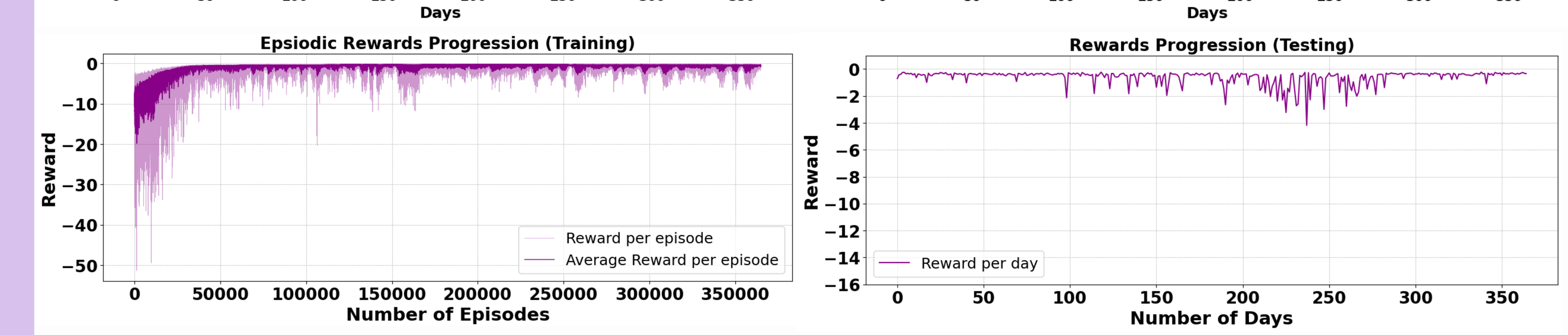
Hyperparameter	Value	Hyperparameter	Value
α	3e-4	GAE's λ	0.1
Hidden Layers	4	N	25
Hidden Size	128	β_{KL}	0.015
ϵ	0.1	buffer_size	3840
γ	0.99	c_v	0.5
		c_s	0.01

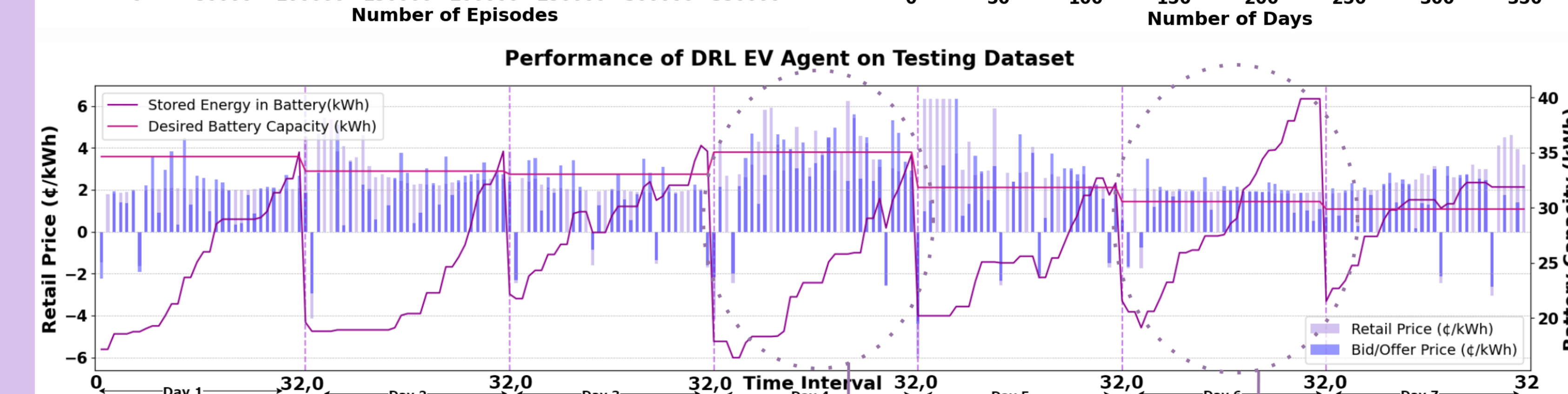
Environment Related Parameters

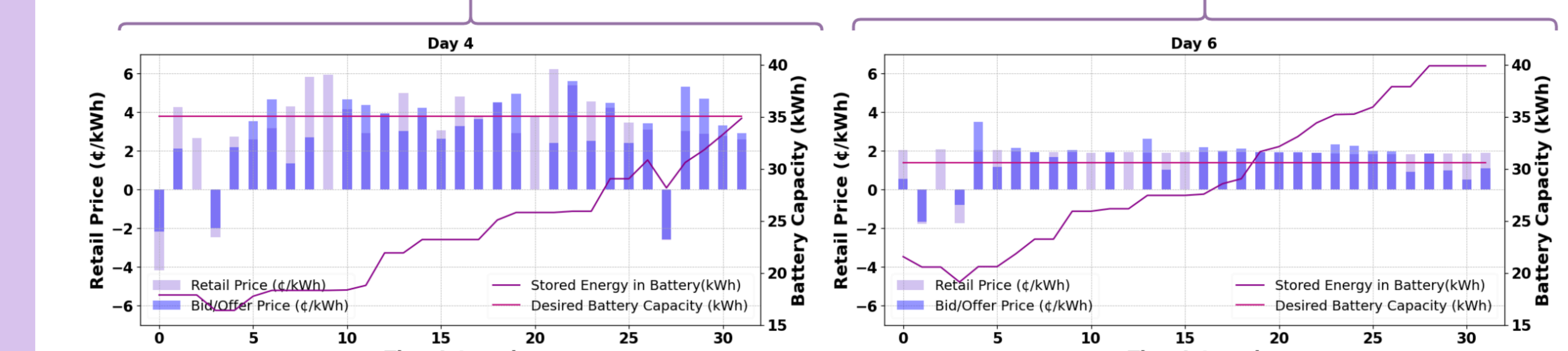
Parameter	Value	Parameter	Value
η_c	0.95	ϕ_u	2.5 ¢/kWh
η_d	0.95	P_{max}	20 kW
B^{cap}	40 kWh	β	80
ϕ_0	10 ¢/kWh	Δt	0.25 h

We utilize ERCOT's Market Data from 2018 and 2019 for training and testing respectively.









The results demonstrate the policy's convergence for placing bids and offers aligned with user goals and constraints, adapting to dynamic price fluctuations and varying charging requirements. Notably, the agent learns not to place discharge offers when prices are constant or low, instead charging to an $SoCt \geq SoC_{des}$. When $SoCt \geq SoC_{des}$ or prices are higher, it benefits by offering to discharge.

CONCLUSIONS

- The research presents a DRL-based EV agent that empowers customer to participate in a bid-based TES-based design for distribution system.
- POMDP formalism helps in enabling the agent make better decisions under uncertainty.
- Testing on real-world data shows the agent converges to a user-aligned policy, adapting to dynamic prices and varying charging needs.

ACKNOWLEDGEMENT

This work was supported by Ministry of Education (MoE), Government of India, through Prime Minister's Research Fellowship Grant.

REFERENCES

[1] S. Battula, L. Tesfatsion, and Z. Wang, "A Customer-Centric Approach to Bid-Based Transactive Energy System Design," IEEE Transactions on Smart Grid, vol. 11, no. 6, pp. 4996–5008, 2020

[2] Dai, Y., Ouyang, H., Zheng, H. et al. Interpreting a deep reinforcement learning model with conceptual embedding and performance analysis. Appl Intell 53, 6936–6952 (2023).