Learning-Based Optimal Charging Discharging Strategy for Electric Vehicles Under Vehicle-to-Grid Scheme

Julie Maeng

Introduction

Recent advances in electric vehicles (EV) technologies have raised the significance of vehicle- to-grid (V2G) schemes in the smart grid domain, which allows bidirectional flows of energy and information between consumers and suppliers. In the V2G scheme, each vehicle is viewed as a potential energy storage system (ESS) that can provide surplus energy to the grid. Thus, it is essential to intelligently manage charging and discharging according to electricity prices and users' needs.

Research Methodology

This paper formulates the individual EV charging problem as a Markov Decision Process (MDP) without a defined transition probability. The Deep Deterministic Policy Gradient (DDPG) is a model-free deep reinforcement learning method which is composed of actor and critic parts, and is based on the Deterministic Policy Gradient (DPG).

Experimental Settings

Notations	
p_t	Price of electricity at time t
e_t	Energy level of EV at time t
s_t	State at time t
a_t	Action at time step t
r_t	Reward at time step t
d_t	Travel distance required for driving event
l_t	Travel energy required from driving event
κ_t	penalty for untraveled distance
C	Cost of battery
E_{max}	Maximum battery capacity
p_{min} , p_{max}	Minimum and maximum value of electricity price
$P_{driving}$, $P_{station}$	Probabilities of driving and station event
T	End of time step

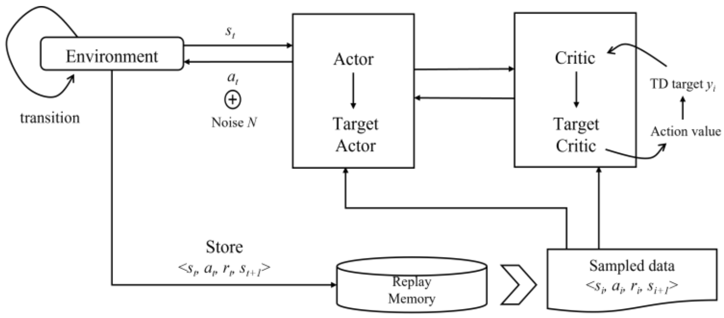
Battery degradation

The battery capacity curve expressed as a function of battery cycles in order to represent the degradation behavior of the EV battery within the simulator. The LFP battery reaches 80% of the initial capacity at about 1000 charge cycles. In this study, tracking the shape of the curve was most prioritized.

Driving Distance

In order to consider some realistic factors for the scheduling scenario, the driving distances traveled by each EV are accounted in the simulation process.

Structure of the DDPG algorithm



Baseline Models

Random Operation: Assuming the maximum amount of charging energy is A, the random operation selects the charging amount of EV at random, ranging from $-A$ to A.

Heuristic-Based: The heuristic-based approach aims to reflect the heuristic behaviors of users that may be affected by the price of energy, since it is assumed that users are aware of real-time electricity prices in the V2G scheme. In this model, the user chooses to discharge energy when the current price is higher than a certain baseline, in order to seek profit. If the price is lower than the set threshold, the user decides that it is more profitable to charge electricity to save costs.

Results and Analysis

The values and distributions of parameters used in the simulated environment are referenced from related experiments. Meanwhile, electricity prices used for testing were obtained from KEPCO's (Korea Electric Power Corporation) EV charging tariff, valid since April 2022.

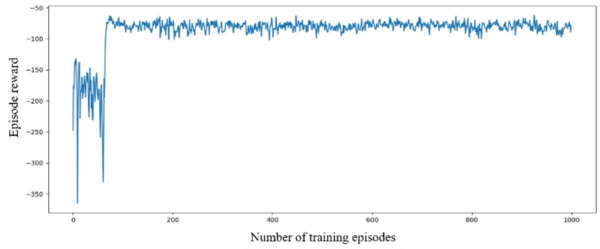
Parameters in the Simulated Environment

RL Environment Data	Value / Distribution
Electricity prices (₩)	KEPCO
Maximum battery capacity E_{max} (kWh)	24
Initial energy ($\times E_{max}$)	$N(0.5, 0.01^2)$
Required distance (km)	{0, 5, 10, 15, 20, 25, 30}
Cost of battery	800
Event probabilities $P_{driving}$, $P_{station}$	0.8, 0.2

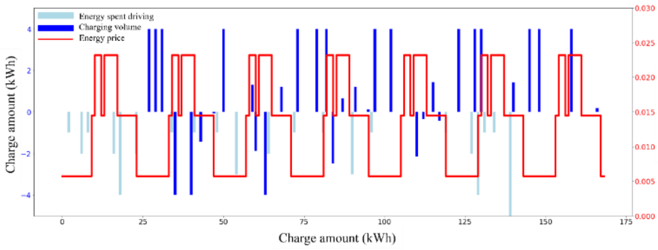
Electricity Price Data

Classification	Time Period		Energy Charge (KRW/kWh)		
			Summer	Spring/Fall	Winter
Low-voltage (≤380V)	Off-peak	23:00~09:00	57.5	58.6	80.6
	Mid-peak	09:00~10:00 12:00~13:00 17:00~23:00	145.2	70.4	128.1
	On-peak	10:00~12:00 13:00~17:00	232.4	75.3	190.7
High-voltage (≥3,300V)	Off-peak	23:00~09:00	52.4	53.4	69.8
	Mid-peak	09:00~10:00 12:00~13:00 17:00~23:00	110.6	64.2	100.9
	On-peak	10:00~12:00 13:00~17:00	163.6	68.1	138.7

Convergence of DDPG



Dynamic Energy Price and Charging/Discharging Volumes for Seven Consecutive Days

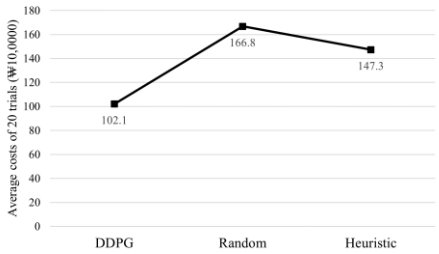


Total Costs per Trial

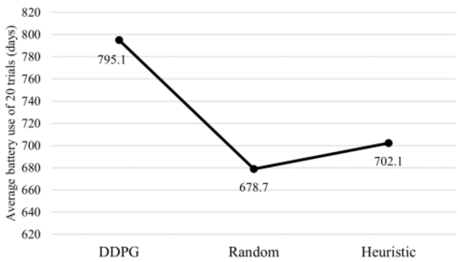
Trial	DDPG	Random	Heuristic
1	101.15	166.43	152.48
2	100.94	158.93	147.26
3	103.41	165.54	146.85
4	98.37	171.42	142.79
5	101.70	167.36	146.85
6	103.09	165.41	149.41
7	100.58	169.99	142.04
8	102.17	164.32	144.92
9	101.80	162.11	147.90
10	99.21	167.66	147.31

(Unit:₩10,000)

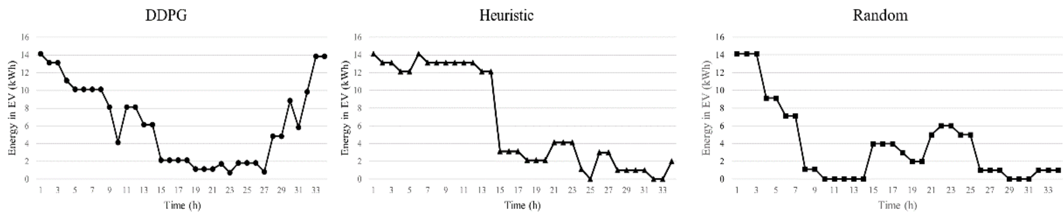
Average Charging Costs



Average Battery Usage Time



Energy Left at Each Corresponding Hours



Experiments

The paper conducts two experiments: comparing average price between the models when there are electricity price fluctuations and there are variations in the traveling distance. The DDPG algorithm successfully reduces average cost of 20 trials in both experiments.

Conclusion

The Deep Deterministic Policy Gradient (DDPG) algorithm is proposed to make decisions about the amount of energy to charge or discharge based on current electricity prices and the amount of energy in the vehicle. The testing results of the algorithm show that the algorithm is capable of making reasonable decisions based on information about the electricity prices and the vehicle's SoC. To further evaluate the performance of the proposed method, random operation and heuristic-based decision making strategy were used for comparison. Experiments conducted in the paper demonstrate that the proposed reinforcement learning algorithm is effective in reducing users' costs while also learning to operate the battery for a longer period of time. At the same time, the algorithm may also impose beneficial effects to the utility, as its strategy has potential to provide more surplus energy to the grid for each vehicle while also allowing the vehicles to charge more energy.