

강화학습 기반 V2G 전력거래 이익 최대화

노란전력

CONTENTS

연구 배경 및 목적

연구 조건 및 제약 사항

시뮬레이션 설계

연구 내용

01 연구 배경 및 목적

[“V2G”]

1

환경문제로 인한 저탄소 정책
도입 가속화

2

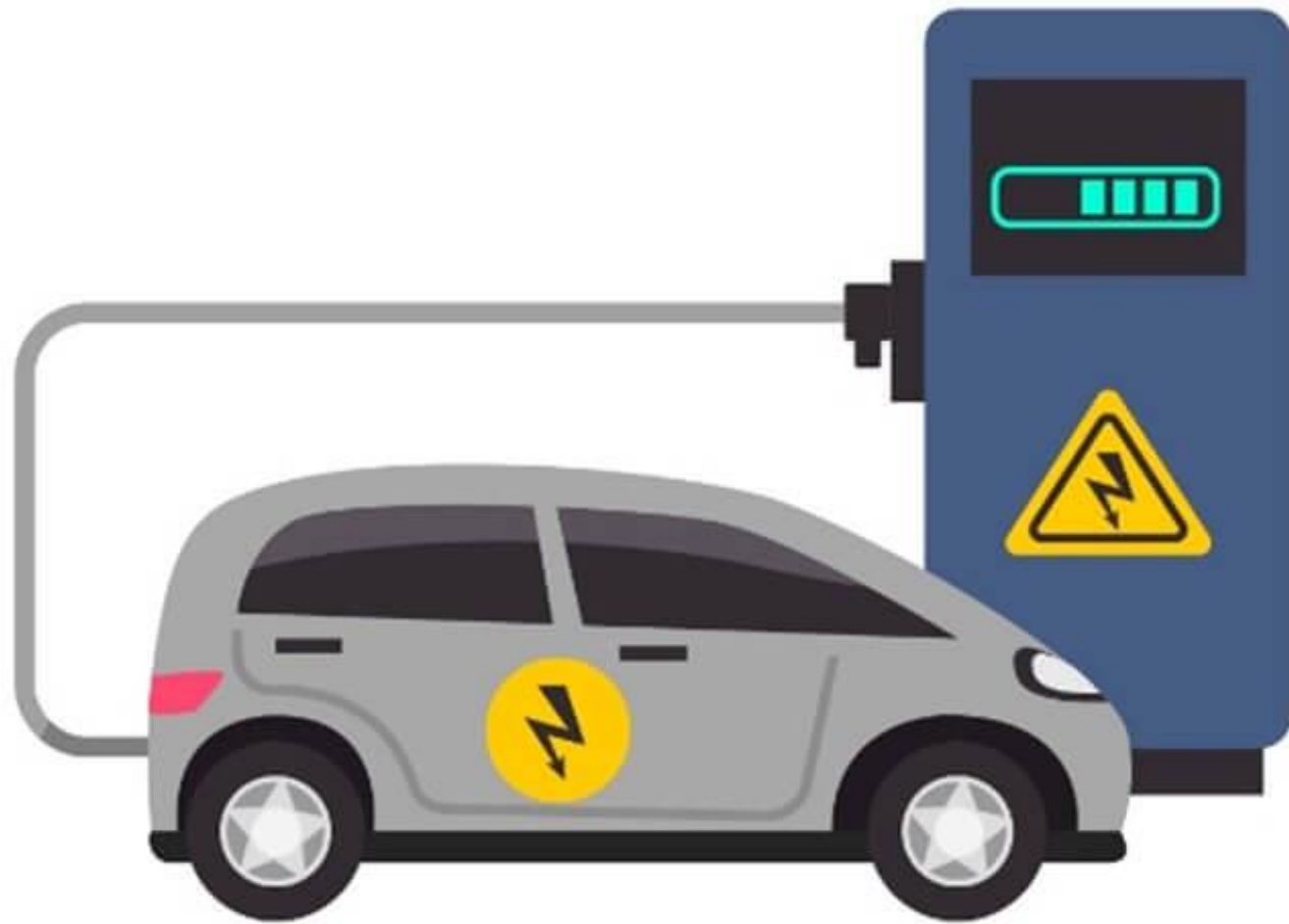
국내 신재생 에너지 보급 확대

3

전세계적으로 전기차 보급 증가

01 연구 배경 및 목적

V2G



전력망의 안정적 유지

에너지관리 최적화

사회 전반적 전력비용 절감

**전기차 소유자가 V2G를 통해 얻는 금전적 혜택이
배터리 수명 단축으로 인한 비용을 상쇄할 만큼 충분하지 않은 경우**



01 연구 배경 및 목적



02 연구 조건 및 제약 사항

개발한 알고리즘의 평가를 위하여 실제 전기차 충전소와 전기차를 활용하는 것은 현실적으로 어려움이 있기에 시뮬레이션 환경을 활용하여 이익을 최대화할 방안을 모색 할 필요가 있다

TABLE I
OVERVIEW OF EXISTING EV SIMULATORS FOCUSING ON SMART CHARGING STRATEGIES.

Simulator Name	V2G	Power Network Impact	EV Models	EV Behavior	Charging Stations	Available Baseline Algorithms	RL Ready	Programming Language	Comments
V2G-Sim [12]	Yes	Partial	Diverse	Real Charging Transaction Probability Distributions	Uniform	- Heuristics, Mathematical Programming	No	Not Open-source	- Customizable V2G simulations.
EVLlibSim [13]	Yes	No	Diverse	Randomized	Uniform	- Heuristics, Mathematical Programming	No	Java	- Easily customizable simulations with a visual interface.
EV-EcoSim [14]	Yes	Complete	Uniform	Randomized	Uniform	- Mathematical Programming	No	Python	- Grid-Impact analysis
evsim [15]	No	Partial	Diverse	Real Charging Transaction Probability Distributions	Uniform	- Heuristics, Mathematical Programming	No	R	- Simulate and analyze the charging behavior of EV users.
OPEN [16]	Yes	Complete	Uniform	Randomized	Uniform	- Heuristics, Mathematical Programming	No	Python	- Modelling, control & simulations for smart local energy systems.
ACN-Sim [17]	No	Complete	Uniform	Real Charging Transactions	Uniform	- Heuristics, MPC, RL, Mathematical Programming	Yes	Python	- Designing a complete simulator framework.
SustainGym [18]	No	Complete	Uniform	Real Charging Transactions	Uniform	- RL	Yes	Python	- Providing a benchmark for sustainable RL applications.
Chargym [19]	Yes	Very Limited	Uniform	Randomized	Uniform	- Heuristics, RL	Yes	Python	- Comparing RL algorithms for smart charging.
EV2Gym (Ours)	Yes	Partial	Diverse	Real Charging Transaction Probability Distributions	Diverse	- Heuristics, MPC, RL, Mathematical Programming	Yes	Python	- Comprehensive simulator for any control algorithm.

EMS - Energy Management System

03 시뮬레이션 설계

V2G-Sim

- EV, 모델과 동작과 같은 풍부한 기능을 갖췄지만 오픈소스가 아니며 강화학습 개발을 위한 환경을 제공하지 않는다.

EV-EcoSim

- 자세한 배터리 사용률 및 성능 저하 모델이 포함되어 있으나, 현실적인 EV 사양 및 동작이 아닌 배터리에 중점을 두고 있다.

ACN-Sim

- 표준화된 Gym, EV 환경이 포함되어 있으며 주차장 특성에 따라 개발되었다가 장. EV 확립된 시뮬레이터 플랫폼 중 하나이며 다른 오픈소스 그리드 시뮬레이터를 사용하여 전력 네트워크 계산을 지원한다 하지만 V2G 지원을 염두에 두지 않아 연구에는 부적합하다

EVLibSim

- 다양한 EV, Java 모델을 제공하고 있지만 로 작성되었으며 그리드에 미치는 영향을 시뮬레이션하지 않는다.

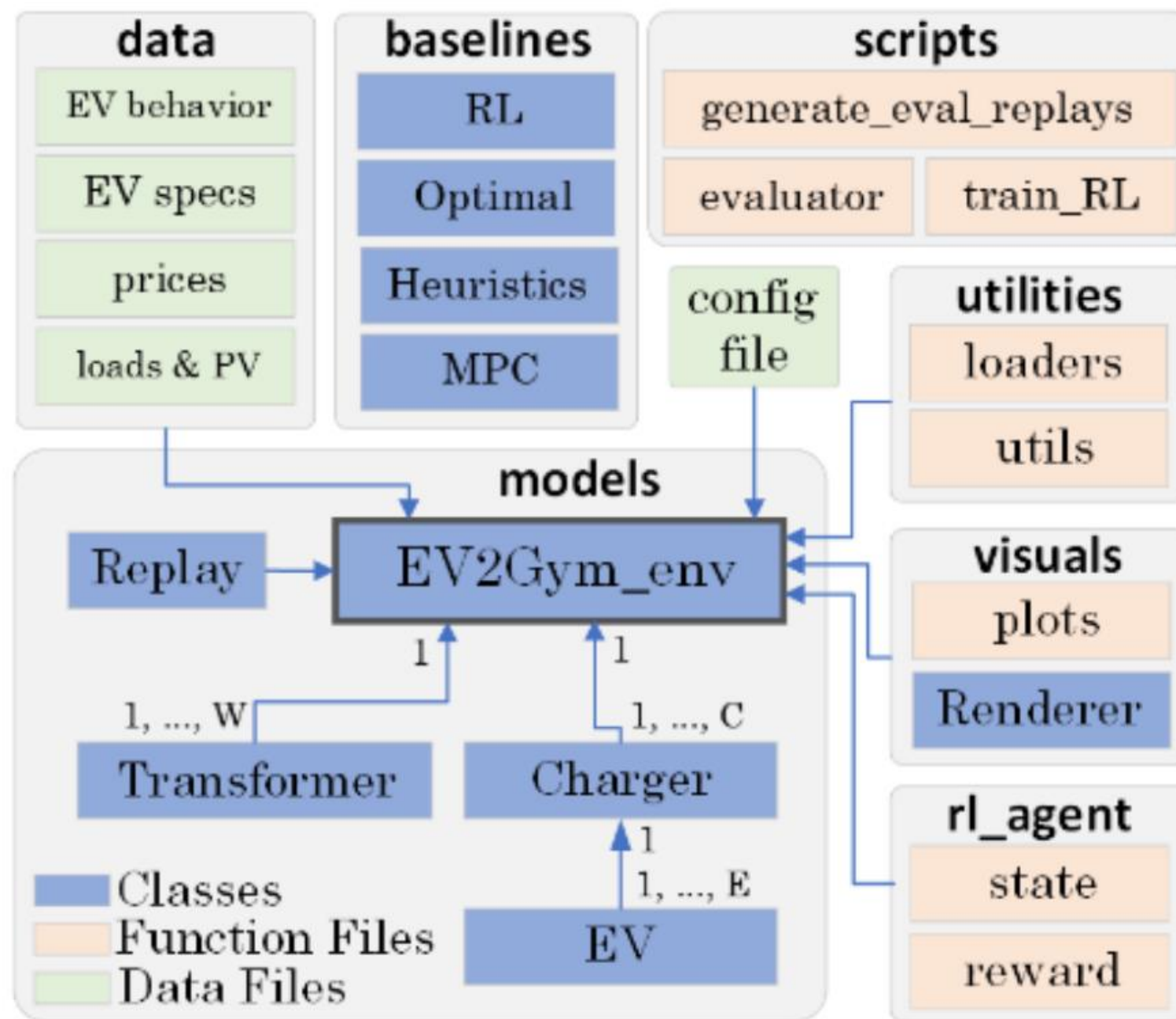
EVsim

- EV 사용자의 행동을 평가하고 분석하는데에 중점을 두었으며 사실적인 EV, V2G 행동 데이터로 구축되어 있으나의 영향을 조사할 수 있는 옵션이 포함되어 있지 않다

Chargym

- 비용 및 페널티 설계에 중점을 둔 강화학습 알고리즘을 개발하는데 사용한다. EV EV 모든 와 충전기가 동일한 사양을 가지고 동작이 실제 데이터를 기반으로 하지 않는 등 기본 모델이 매우 단순하여 강화학습 연구에 부적합하다

03 시뮬레이션 설계



EV2GYM 패키지 폴더 및 파일 구조

장점

강화학습 환경을 위한 파이썬 패키지 Gymnasium을 기반으로 구성되어 간소화된 강화학습 알고리즘 평가가 가능

EV2GYM 은 오픈소스 검 모듈식으로 필요에 따라 다양한 추가 기능을 추가 및 수정이 원활하게 가능하다

[V2G Profit Maximization with Loads]

주어진 정보

충전기에 연결된 EV의 출발 및 SOC 정보

충전 에너지 가격, 부하 및 PV 발전

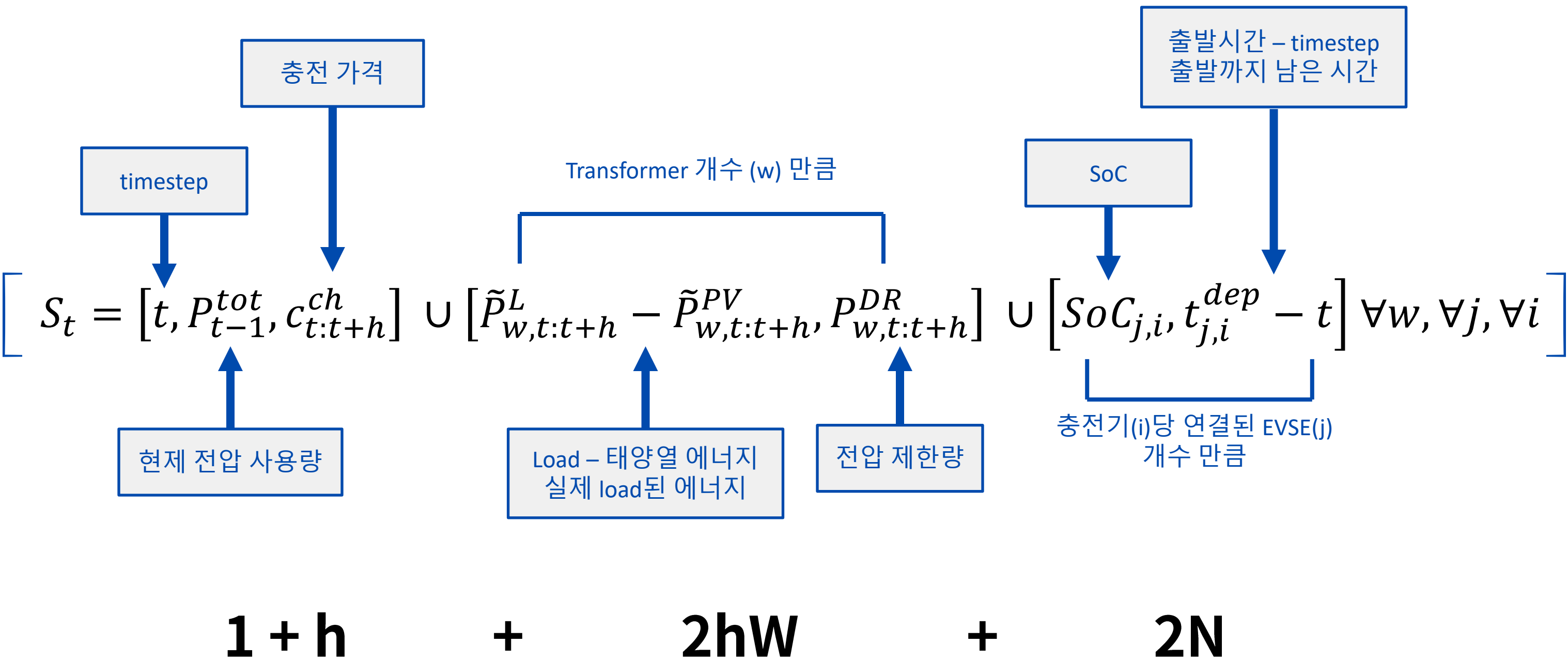
목표

EV 사용자의 요구 충족
+
CPO의 이익 극대화
+
Transformer 과부하 방지

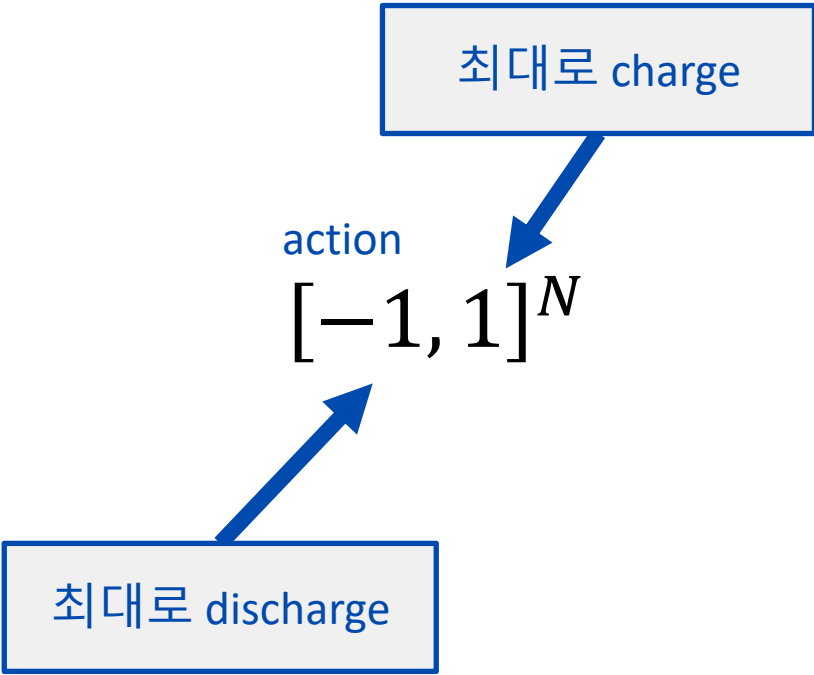
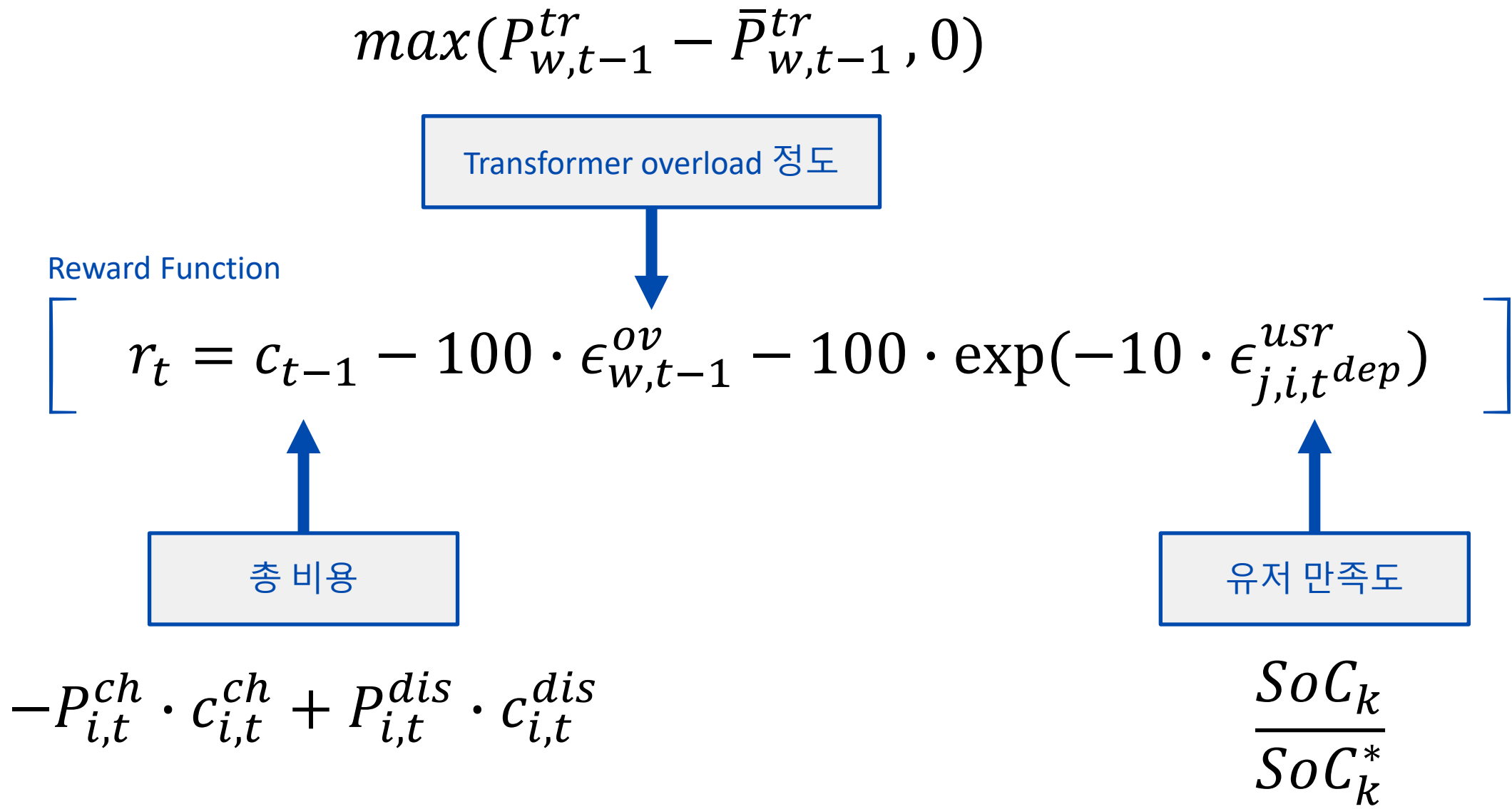
행동

충전기별 충 · 방전 비율

04 강화학습 구조 - state



04 강화학습 구조 - reward, objective function, action



objective Function

$$\left[\max_{I_{j,i,t}^{ch}, I_{j,i,t}^{dis}} \sum_{t \in T} \sum_{i \in C} r_t \cdot \Delta t \right]$$

05 강화학습 알고리즘

TABLE VII
BASELINES COMPARISON FOR THE V2G PROFIT MAXIMIZATION PROBLEM WITH 10 EV CHARGERS, TRANSFORMER LOADS, PV, AND DR EVENTS.

Algorithm	Profits/ Costs (€)	ϵ^{usr} (%)	Energy Ch. (kWh)	Energy Disch. (kWh)	Tr. Ov. (kWh)	Total Q^{lost} ($\times 10^{-3}$)	$\sum d^{cal}$ ($\times 10^{-3}$)	$\sum d^{cyc}$ ($\times 10^{-3}$)	Execution Time (s)	Reward ($\times 10^3$)
AFAP	-23.4 \pm 6.4	100 \pm 0	109 \pm 29	0 \pm 0	138 \pm 86	0.40 \pm 0.09	0.15 \pm 0.04	0.24 \pm 0.07	0.02 \pm 0.00	-13.8 \pm 8.6
ALAP	-25.2 \pm 7.1	100 \pm 0	109 \pm 29	0 \pm 0	106 \pm 78	0.38 \pm 0.09	0.14 \pm 0.04	0.24 \pm 0.07	0.02 \pm 0.00	-10.6 \pm 7.8
RR	-22.6 \pm 5.6	99 \pm 1	107 \pm 26	0 \pm 0	11 \pm 19	0.45 \pm 0.11	0.17 \pm 0.04	0.29 \pm 0.08	0.03 \pm 0.00	-1.2 \pm 1.9
ARS	17.7 \pm 9.9	53 \pm 11	45 \pm 19	107 \pm 35	36 \pm 67	0.57 \pm 0.14	0.12 \pm 0.03	0.45 \pm 0.11	0.09 \pm 0.01	-3.6 \pm 6.7
DDPG	25.3 \pm 9.5	44 \pm 9	33 \pm 16	125 \pm 34	43 \pm 88	0.59 \pm 0.13	0.12 \pm 0.03	0.47 \pm 0.11	0.12 \pm 0.02	-4.4 \pm 8.8
PPO	23.7 \pm 10.2	46 \pm 9	36 \pm 16	123 \pm 38	37 \pm 89	0.58 \pm 0.14	0.12 \pm 0.03	0.46 \pm 0.11	0.13 \pm 0.27	-3.7 \pm 8.9
TD3	31.1 \pm 10.6	38 \pm 8	25 \pm 13	142 \pm 41	74 \pm 140	0.61 \pm 0.13	0.12 \pm 0.03	0.50 \pm 0.11	0.12 \pm 0.03	-7.5 \pm 14.0
TOC	1.9 \pm 21.6	72 \pm 24	66 \pm 39	62 \pm 58	35 \pm 56	0.53 \pm 0.13	0.14 \pm 0.04	0.40 \pm 0.11	0.16 \pm 0.02	-3.5 \pm 5.6
TRPO	13.5 \pm 8.5	57 \pm 10	41 \pm 17	88 \pm 30	7 \pm 13	0.52 \pm 0.12	0.14 \pm 0.03	0.40 \pm 0.09	0.13 \pm 0.01	-0.7 \pm 1.3
MPC	17.8 \pm 9.3	76 \pm 9	480 \pm 107	462 \pm 109	183 \pm 167	1.68 \pm 0.42	0.12 \pm 0.03	1.55 \pm 0.39	108.18 \pm 19.44	-18.3 \pm 16.7
Optimal	3.7 \pm 6.1	100 \pm 0	578 \pm 135	469 \pm 117	4 \pm 12	1.45 \pm 0.35	0.15 \pm 0.04	1.31 \pm 0.32	33.30 \pm 28.05	-0.4 \pm 1.2

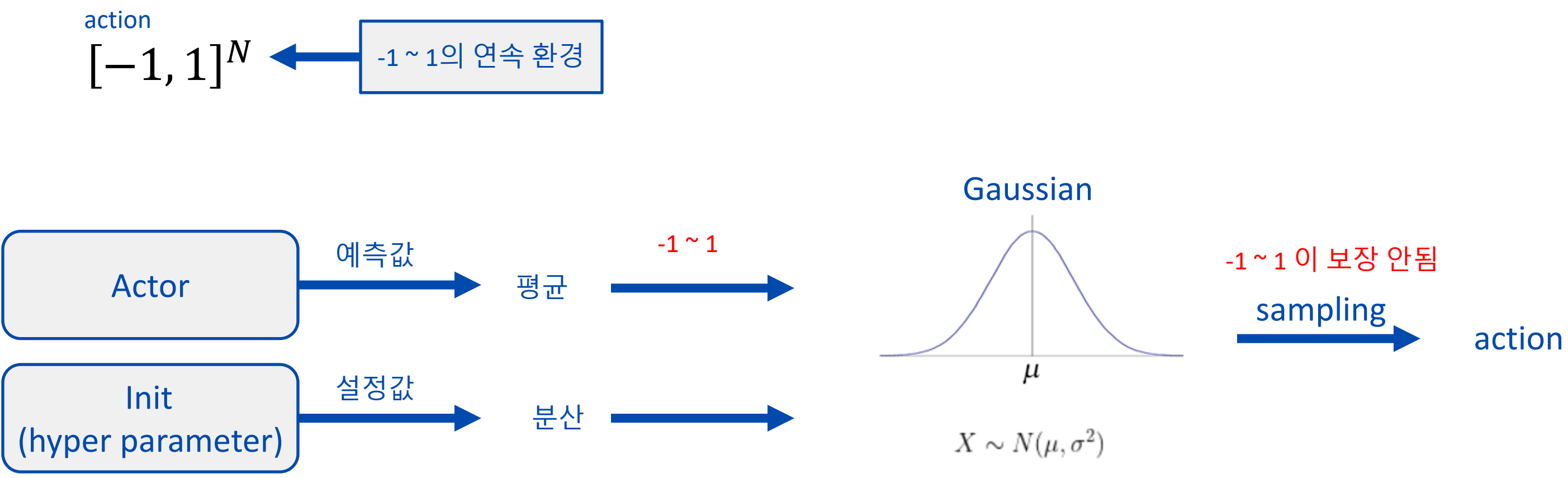
TRPO

- 최대한 큰 step으로 학습하면서도 과장된 향상을 막는 것이 목적
- 정책의 향상이 보장되는 알고리즘을 바탕으로 근사하여 수렴
- Step-size를 KL divergence로 제한하여 과장된 향상을 막음

PPO

- TRPO와 같은 목적
- TRPO와 다르게 KL divergence 대신 Clipping을 사용
- 2차 미분없이 1차 미분으로 계산할 수 있어 구현이 간단

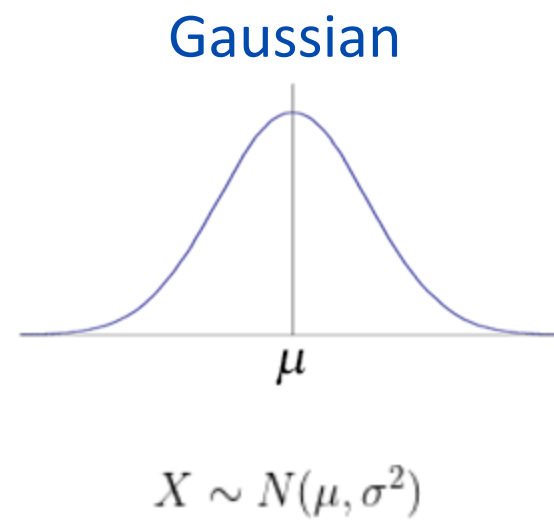
05 강화학습 알고리즘



11% 정도의 오류값

706350 6000640 0.1177124440059727

05 강화학습 알고리즘



-1 ~ 1 이 보장 안됨

sampling

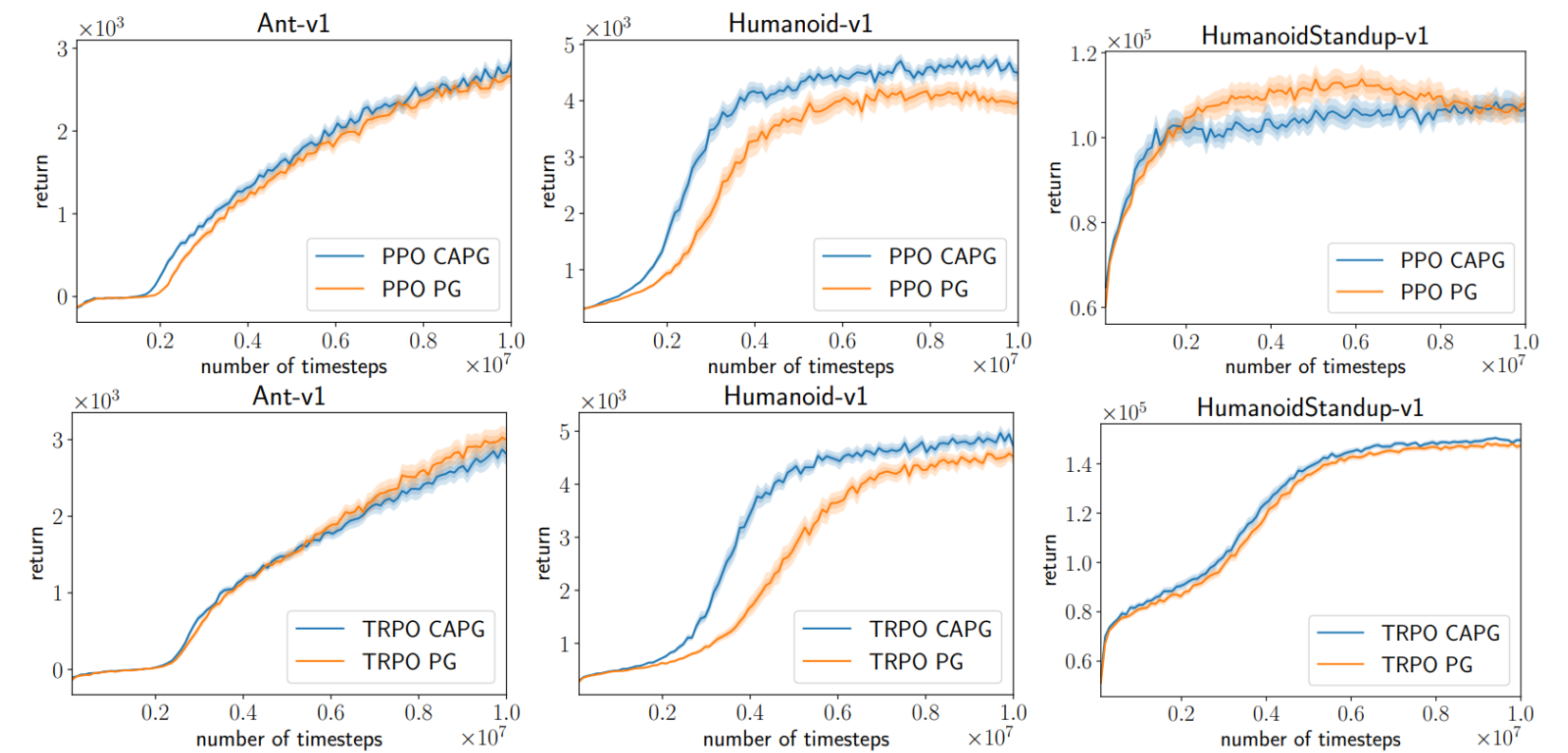
clipping

action

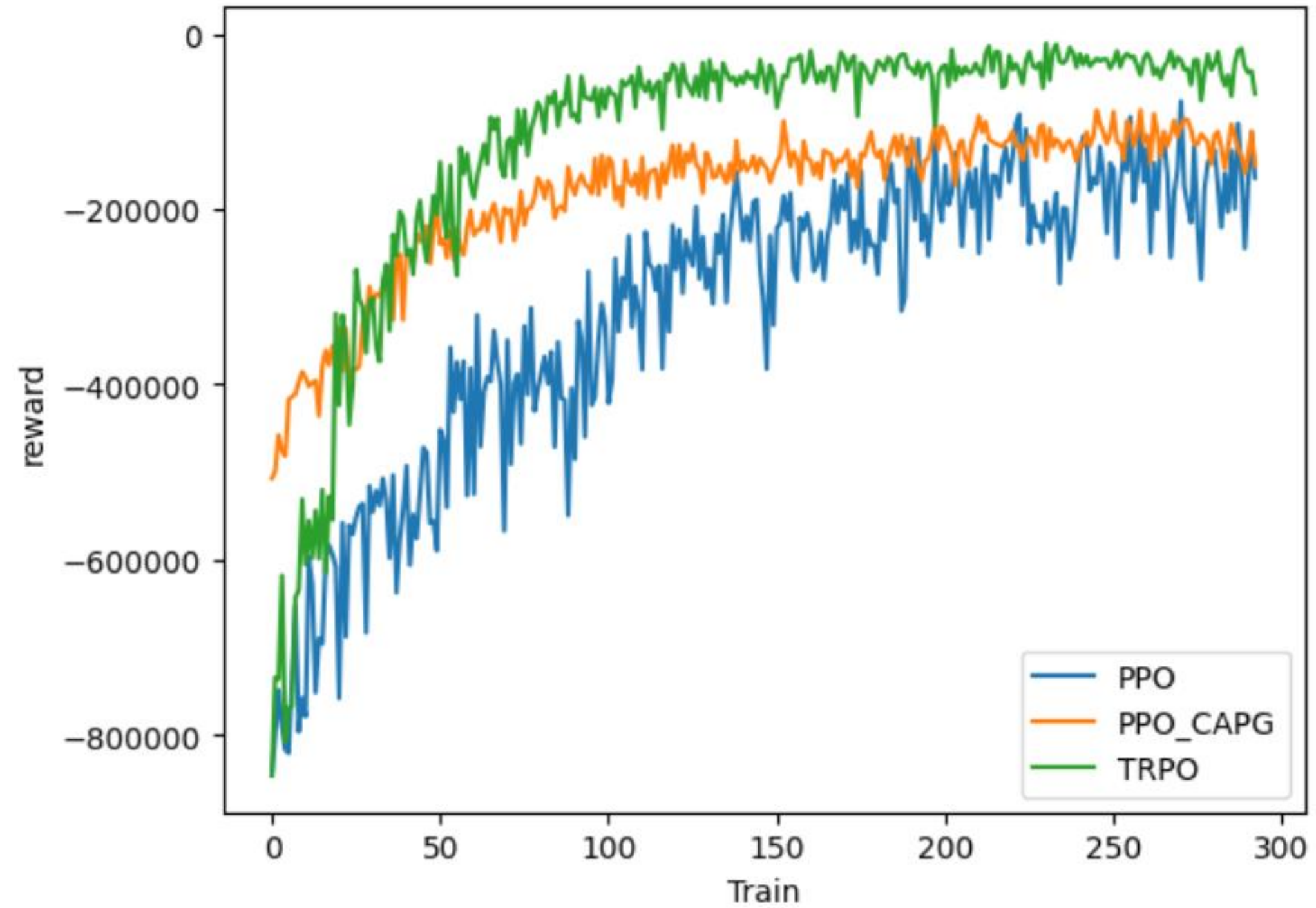
-1 ~ 1 범위에 맞춤

Clipped Action Policy Gradient

[출처] : <https://arxiv.org/abs/1802.07564>



06 시뮬레이션 결과



06 시뮬레이션 결과

Env	algorithm	Profits/costs \$	ϵ^{usr} (%)	Energy Ch. (kWh)	Energy Disch. (kWh)	Tr. Ov. (kWh)	Total Q^{lost} ($\times 10^{-3}$)	$\sum d^{cal}$ ($\times 10^{-3}$)	$\sum d^{cyc}$ ($\times 10^{-3}$)	Execution Time (s)	Reward ($\times 10^{-3}$)
env1	TRPO	35.1	41	53	183	5	0.50	0.10	0.40	0.02	-0.53
	PPO	47.4	31	52	219	80	0.54	0.08	0.46	0.02	-8.09
	PPO CAPG	24.3	54	63	151	9	0.44	0.12	0.32	0.04	-0.97
env2	TRPO	36.6	41	68	199	4	0.53	0.10	0.43	0.02	-0.52
	PPO	47.6	31	50	216	27	0.53	0.08	0.45	0.02	-2.80
	PPO CAPG	24.1	54	57	114	15	0.42	0.12	0.30	0.04	-1.59
env3	TRPO	35.7	42	71	199	21	0.53	0.10	0.44	0.02	-2.18
	PPO	47.6	31	52	218	120	0.54	0.08	0.46	0.02	-12.07
	PPO CAPG	24.1	54	56	143	2	0.42	0.12	0.30	0.04	-0.31
env4	TRPO	36.8	41	82	213	4	0.59	0.10	0.49	0.02	-0.46
	PPO	47.6	31	48	214	11	0.52	0.08	0.45	0.02	-1.22
	PPO CAPG	24.2	54	59	146	12	0.43	0.12	0.31	0.04	-1.25
env5	TRPO	39.7	37	56	201	1	0.52	0.10	0.42	0.02	-0.15
	PPO	47.4	31	45	212	45	0.52	0.08	0.44	0.02	-4.60
	PPO CAPG	24.2	54	58	145	6	0.42	0.12	0.31	0.04	-0.66