

2024년 전기

Vehicle-to-Grid (V2G)를 위한 알고리즘 및 인공지능 모델 개발



EnerV2Gize

202055577 이선진

202055582 이지은

목차

1. 연구 배경 및 목표

1-1. 연구 선정 배경

1-2. 연구 목표

2. 과제 수행 내용 및 중간 결과

2-1. PST(Power Setpoint Tracking) Problem 평가

2-2. V2G Profit Maximization 평가

2-3. Battery Degradation 관점 평가

2-4. 배터리 효율을 고려한 V2G 충방전 보상/상태함수 알고리즘 설계

2-4-1. 전력망 운영자 관점

2-4-2. 충전 서비스 사업자 관점

2-4-3. 전기차 사용자 관점

2-4-4. 보완할 점

3. 구성원 별 역할 분담

4. 향후 진행 방향

4-1. 일정 계획

4-2. 개발 계획

5. 참고 문헌

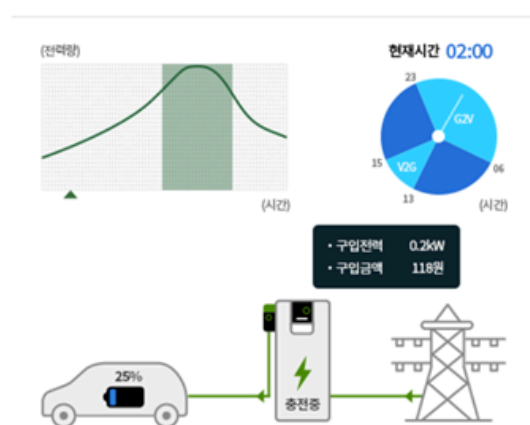
1. 연구 배경 및 목표

1) 연구 선정 배경

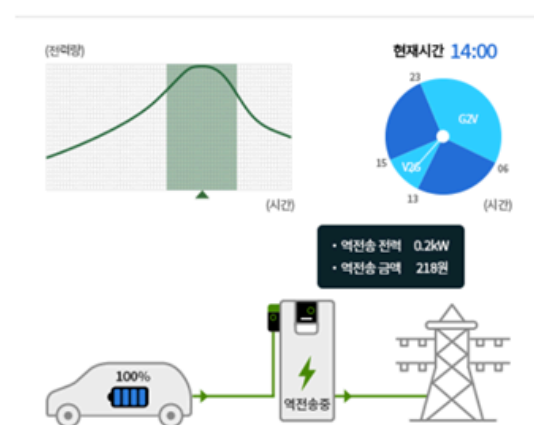
V2G(Vehicle to Grid)는 자동차에서 전력망으로 전기를 이동하는 것을 의미하는데, 전기차에 저장한 배터리를 에너지저장장치(ESS)로 활용하여 전력계통에 연계하는 기술을 뜻한다.

전기 자동차는 엔진 대신 배터리에 있는 전력으로 모터를 구동한다. 전기차 운전자는 전기요금이 저렴한 심야 시간대에 배터리를 충전하고, 출근 이후에는 배터리에 남아있는 전력을 되판다. 이러한 거래 행위를 통해 피크 시간대에는 전력회사에 전력을 팔아 고객은 돈을 벌고, 전력회사는 발전소 가동률을 줄여 수요관리를 하는 양방향 기술이다. 전기차의 전력망으로부터 전력을 공급받아 움직일 수 있는 관점에서 전기차는 에어컨이나 냉장고와 똑같이 전력을 소비하는 장치이나, 배터리에 저장 되어있는 에너지를 다시 전력망으로 되돌린다는 점에서 전기차는 하나의 발전기 역할을 하는 것으로 볼 수 있다.

> 전기요금이 싼 시간에 충전

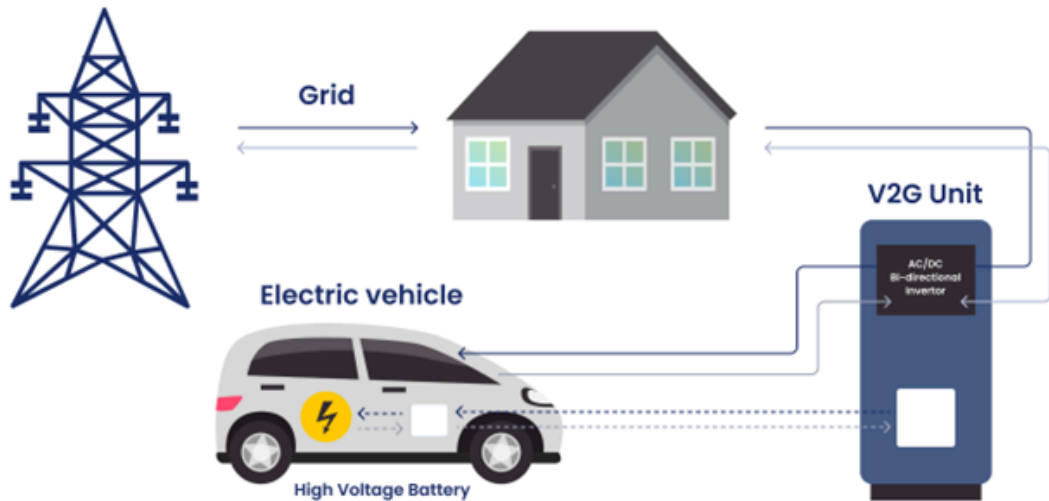


> 피크시간에 역전송



[그림 1] V2G 시스템의 기본 개념

국가적 관점에서 [그림 1]과 같이 전력 소모가 낮은 밤 시간에는 전기차를 충전하여 에너지를 저장해 두었다가 전력 수요가 급증하는 낮 시간에 에너지를 전력망으로 되돌림으로써 V2G 기술을 활용할 수 있다.



[그림 2] V2G 시스템 구성도

현재 V2G 기술은 이산화탄소의 배출을 줄여 친환경적이며, 경제성을 갖추어 지속 가능한 성장을 이어갈 수 있는 산업 모델로 주목받고 있다. [1] 향후 V2G를 활용하여 피크절감 효과 뿐만 아니라 전력계통 주파수 조정, 신재생 에너지의 발판 등 다양한 전력 보조 서비스와 부가가치 창출이 가능할 것으로 기대된다.

다만, 전기차 배터리는 충방전 주기에 따라 성능이 저하되기 때문에, 충전 및 방전 과정에서 배터리의 수명을 최대한 유지하면서도 효율적인 에너지 관리를 수행하는 알고리즘이 필요하다. 또한 전기차 소유자들의 V2G 시스템 참여를 유도를 위해 배터리 사용의 효율성을 극대화하면서 사용자에게 경제적인 혜택을 제공할 수 있는 알고리즘이 필요하다.

2) 연구 목표

논문 및 연구를 통한 V2G 시스템에 대한 이해를 바탕으로 RL(Reinforcement Learning)을 이용한 두 가지 알고리즘 개발을 목표로 한다.

첫째, ‘전기차 배터리 용량 및 현재 상태 등의 파라미터를 이용한 V2G 충/방전 알고리즘 설계’를 통해 전력망을 안정화하고 에너지 효율성을 향상시킨다. 현재 배터리 상태, 전력

요금, 사용자 만족도, 전력망 상태 등을 고려하여 효율적인 충/방전을 가능하게 함으로써 비용 절감 및 재생 에너지 사용을 높일 수 있다.

둘째, ‘전기차 충전 요금, 목표 지역의 계통 에너지 패턴, 전기차 소유주에 대한 혜택 등을 고려한 알고리즘 설계’를 통해 스마트 그리드의 효율성을 극대화한다. 충/방전 알고리즘에서 더 나아가 사용자 운행 패턴, 충/방전 기여에 대한 금전적 혜택을 제공하는 알고리즘을 설계함으로써 전기차의 잠재력을 극대화하고, 스마트 그리드의 효율성을 높인다.

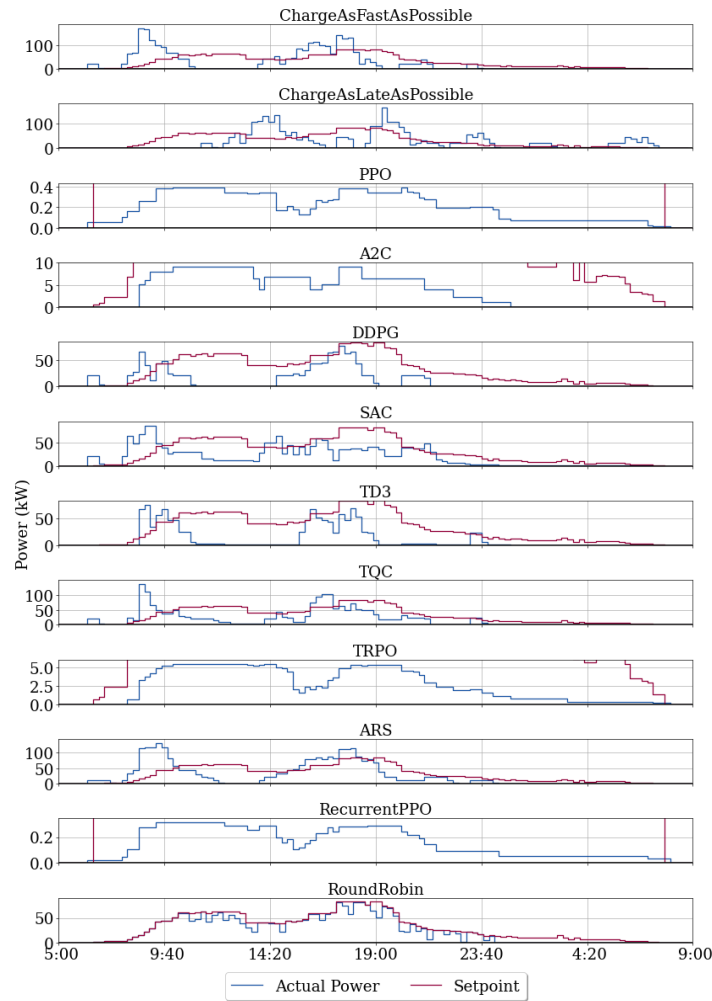
해당 알고리즘을 통해 스마트 그리드의 구현과 전기차의 활용을 극대화하며, 전기차 소유자와 전력망 운영자 모두에게 경제적, 환경적 이점을 제공할 수 있다.

2. 과제 수행 내용 및 중간 결과

본 과제를 수행하기에 앞서 V2G 시뮬레이션 환경인 EV2Gym을 활용하여 PST Problem과 V2G Profit Maximization Problem, 그리고 Battery Degradation 관점에서 성능을 평가해보았다.

1) PST(Power Setpoint Tracking) Problem 평가

실제 전력 소비와 목표 전력 간의 차이를 최소화하는 것을 목표로 시스템이 목표 전력에 도달하도록 유도한다.



[그림 3] PST Problem 평가 그래프

Heuristic 알고리즘 3가지, RL 알고리즘 9가지에 총 12개의 알고리즘에 대해 PST Problem을 평가해 보았다. *Charge As Fast As Possible* (AFAP) 알고리즘은 EV가 연결되자마자 가능한 최대 전력으로 즉시 충전을 시작한다. 변압기의 제약 조건을 고려하지 않고 충전 속도를 최우선으로 한다. *Charge As Late As Possible* (ALAP) 알고리즘은 EV가 목표 SoC에 도달할 수 있도록 가능한 한 늦게 최대 속도로 충전을 시작한다. *Proximal Policy Optimization* (PPO)는 강화 학습 기반 알고리즘으로, EV 충전 시 그리드의 상태를 고려하여 적절한 충전 패턴을 학습하고 주어진 목표에 따라 충전 속도를 유연하게 조절한다. *Advantage Actor-Critic* (A2C)는 강화 학습 기반 알고리즘으로, 충전 패턴을 최적화하기 위해 그리드 상태와 EV의 목표를 동시에 고려하여 점진적으로 충전 속도를 조절한다. *Deep Deterministic Policy Gradient* (DDPG)는 연속적인 action 공간을 다루는 강화 학습 알고리즘으로, EV 충전 시 전력망 상태에 따라 실시간으로 충전 속도를 최적화한다. *Soft Actor-Critic* (SAC)는 그리드 제약 조건을 고려하여 EV 충전 시 전력 소비를 효율적으로 관리한다. *Twin Delayed Deep Deterministic Policy Gradient* (TD3)는 DDPG의 개선

알고리즘으로 그리드 상태와 EV 목표를 기반으로 충전 속도를 조절한다. *Truncated Quantile Critics* (TQC)는 분포 기반 강화 학습 알고리즘으로, 다양한 시나리오에서 안정적인 충전 성능을 보장하기 위해 충전 속도를 조절하고 전력망 제약의 반영하여 최적의 충전 패턴을 학습한다. *Trust Region Policy Optimization* (TRPO)는 강화 학습 알고리즘으로, 정책 변화의 안정성을 유지하면서 최적의 충전 전략을 학습하고, 그리드 상태를 반영하여 EV 충전 속도를 조절한다. *Augmented Random Search* (ARS)는 random 탐색을 통해 최적의 충전 패턴을 찾는 강화학습 알고리즘으로, 그리드 제약 조건을 고려하여 EV 충전 속도를 조절한다. *Recurrent PPO* (RPPO)는 과거의 충전 상태를 기억하여 현재 충전 결정을 최적화하는 알고리즘으로 시간에 따른 전력망 상태 변화를 고려하여 충전한다. *Round Robin* (RR)은 각 EV가 공평하게 에너지 나누도록 순환 방식으로 순차적 EV 최대 전력 setpoint 까지만 충전한다.

Algorithm	Energy Charged (kWh)	ϵ^{usr} (%)	Tracking Error (10^3)	Energy Error (kWh)	Reward (10^3)
A2C	114±38	72±4	140±42	809±158	-140±42
ARS	783±140	100±0	105±46	667±163	-59±23
AFAP	784±140	100±0	171±74	814±176	-95±44
ALAP	784±140	100±0	137±52	715±141	-137±52
DDPG	368±95	83±4	126±42	725±124	-126±42
PPO	6±2	68±4	182±52	930±158	-182±52
RPPO	5±1	68±4	182±52	931±157	-182±52
RR	745±135	98±1	14±5	181±34	-8±4
SAC	551±102	92±3	75±27	572±129	-75±27
TD3	306±93	81±4	121±40	709±134	-121±40
TQC	548±116	91±3	94±36	629±140	-94±36
TRPO	85±20	71±4	156±48	858±154	-156±48

[그림 4] Baseline Algorithms for the PST Problem with 20 Chargers

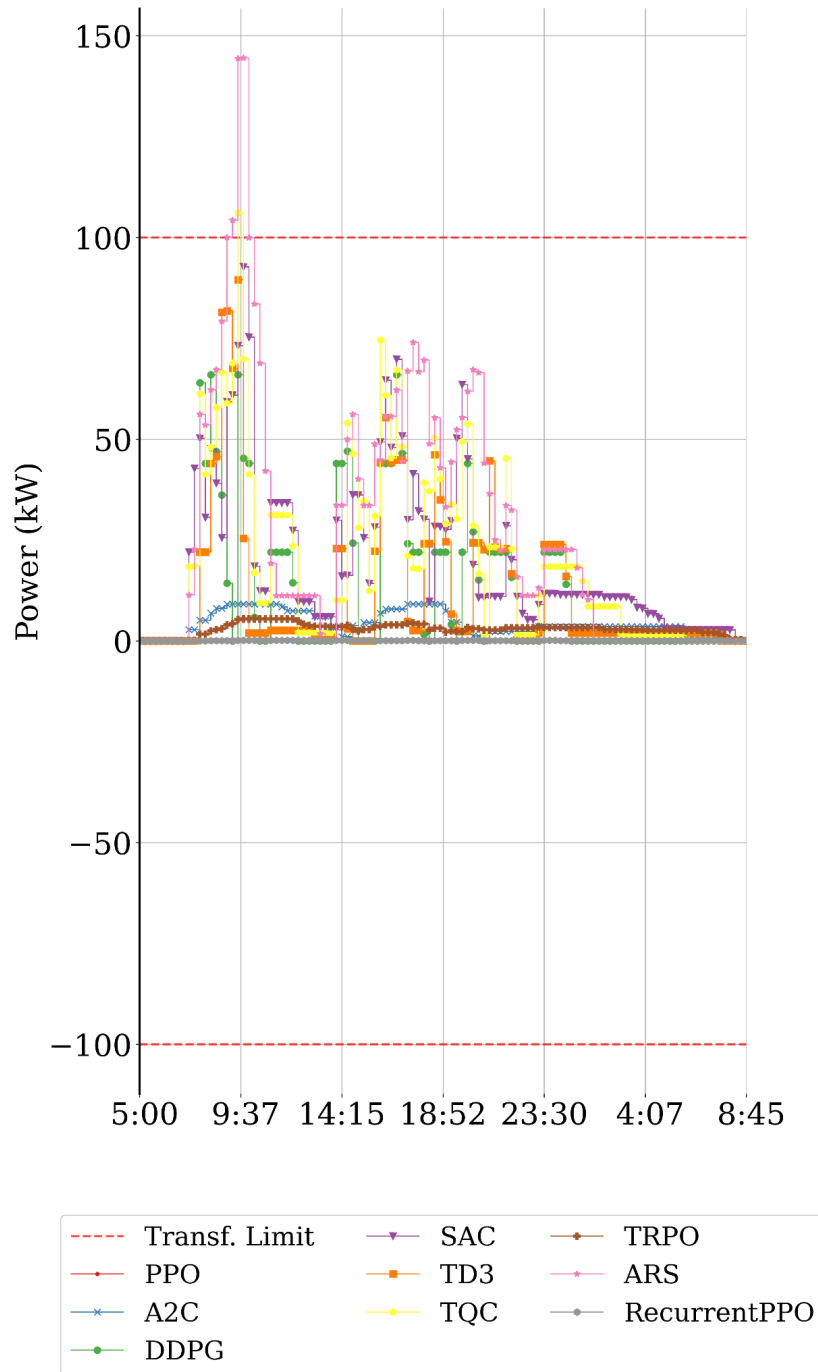
20개의 Charging Station으로 15분 단위 총 85단계 평가하였다.

RL 알고리즘 중에서는 DDPG / SAC / TD3 / TQC / ARS가 상대적으로 목표치와 근사하게 결과를 보임을 알 수 있다.

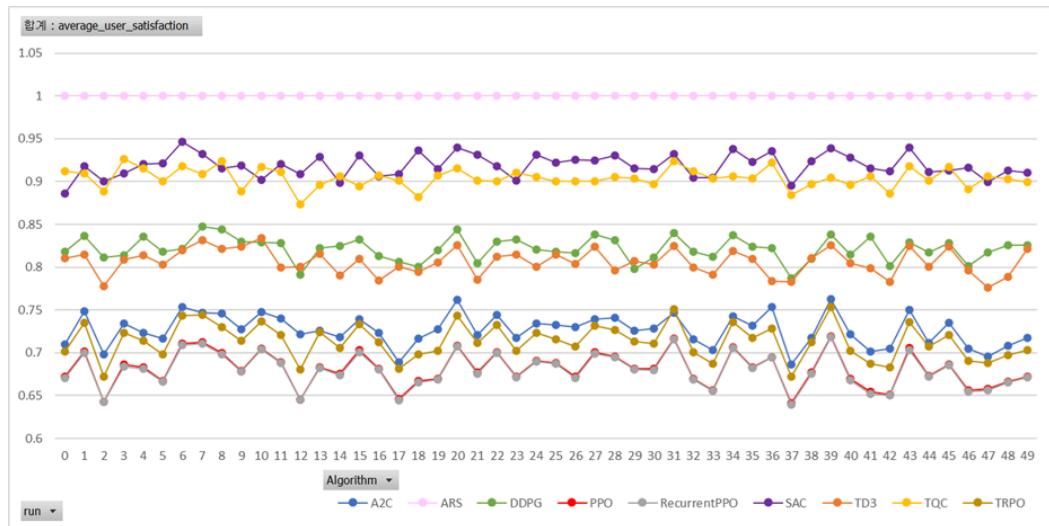
결과 그래프 상 Heuristic의 RR 알고리즘이 가장 우세한 성능을 보이나, RR은 충전소 수의 증가, 제약 조건의 증가에 따라 한계를 보여, RL을 중심으로 알고리즘을 설계 및 평가할 것이다.

2) V2G Profit Maximization 평가

V2G 이익 최대화 문제와 관련하여 RL 알고리즘 별 전력 소비 패턴과 사용자 만족도를 비교해보았다.



[그림 5] 전력 소비 패턴



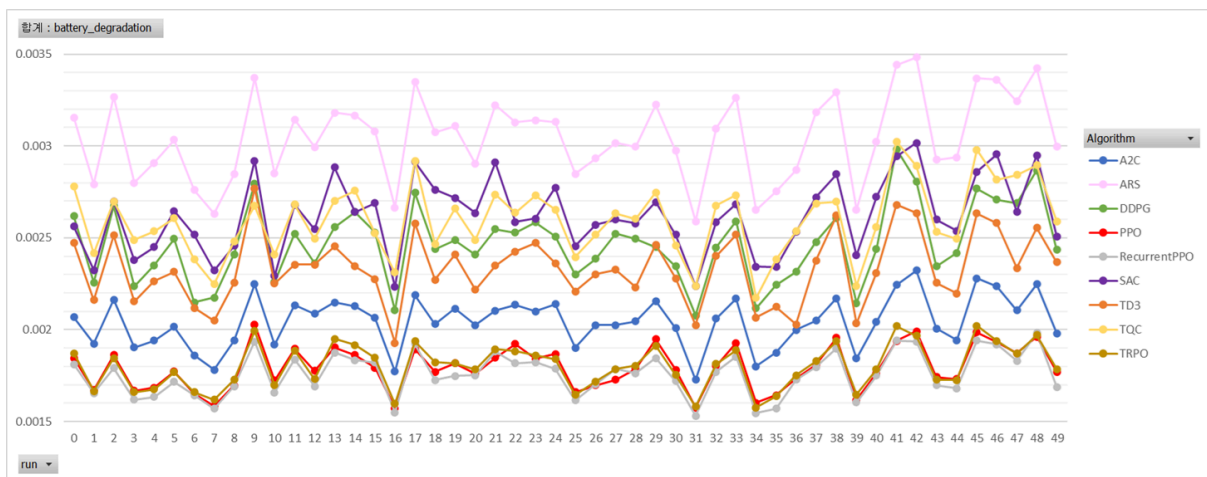
[그림 6] RL 알고리즘 별 사용자 만족도

전력 소비 패턴 그래프에서 **SAC, TQC** RL 알고리즘이 전력 제한 범위 내에서 유연한 형태를 보이며, 사용자 만족도에서 또한 높은 만족도를 보임을 알 수 있다.

반면 전력 제한 범위에서 변동성을 거의 보이지 않은 **A2C, TRPO, PPO, RPPO** 알고리즘의 경우 사용자 만족도에서 낮은 결과를 보인다.

3) Battery Degradation 관점 평가

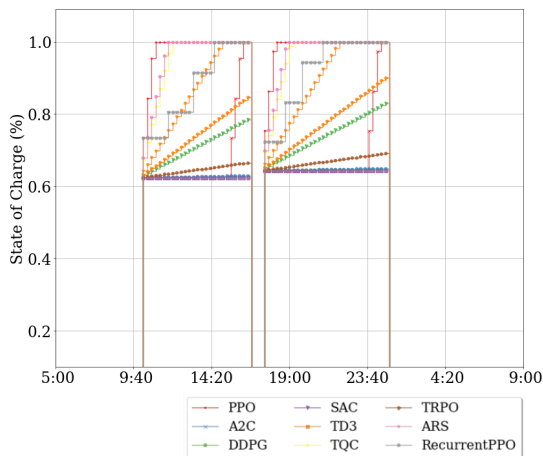
20개의 충전기로 하루 15분 간격 총 85단계, 9가지 RL 알고리즘에 대해 EV 시나리오 각 5,000회 학습 후 평가해보았다.



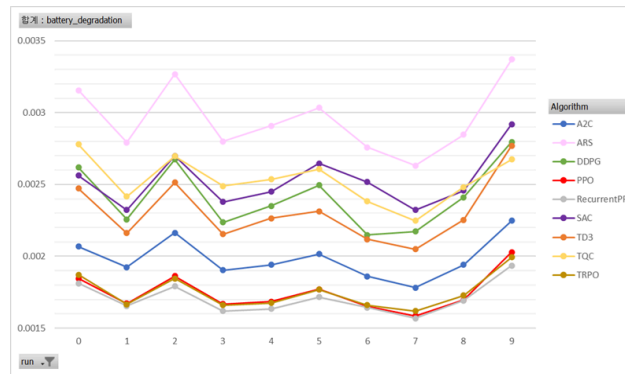
[그림 7] RL 알고리즘 별 배터리 열화 평가 그래프

-> 배터리 열화 계산 방법, 가로축 추가

배터리 열화의 적정 범위가 연간 0.5% ~ 1%, 하루 단위 약 0.0013% ~ 0.0027% 임을 고려하였을 때, 가장 높은 열화를 보이는 ARS 알고리즘이 비효율적이며 하단의 RPPO, PPO, TRPO 알고리즘이 비교적 효율적인 결과를 보임을 알 수 있다.



[그림 8] SoC 그래프 (일부)



[그림 9] 알고리즘 별 배터리 열화 그래프 (일부)

전기차 충전 상태 (SoC)와 배터리 열화 상태를 비교해보았을 때, SoC의 값이 1이 되는 즉, 완전 충전을 진행하는 ARS 알고리즘이 배터리 열화 측면에서도 비효율적인 결과를 보임을 알 수 있다.

위 실험 결과를 토대로 적정 수준의 SoC 범위 내에서 충방전이 이루어지고, 배터리 열화 지표 또한 고려하여 높은 사용자 만족도를 보이는 효율적인 배터리 충방전 알고리즘을 설계할 예정이다.

4) 배터리 효율을 고려한 V2G 충방전 보상/상태함수 알고리즘 설계

전력망 운영자, 전기차 사용자, 충전 서비스 사업자 총 3가지의 이해관계자별로 각자의 이익에 중요한 지표를 생각해보고, 이를 개선하기 위한 보상 함수 및 상태 함수를 설계해 보았다.

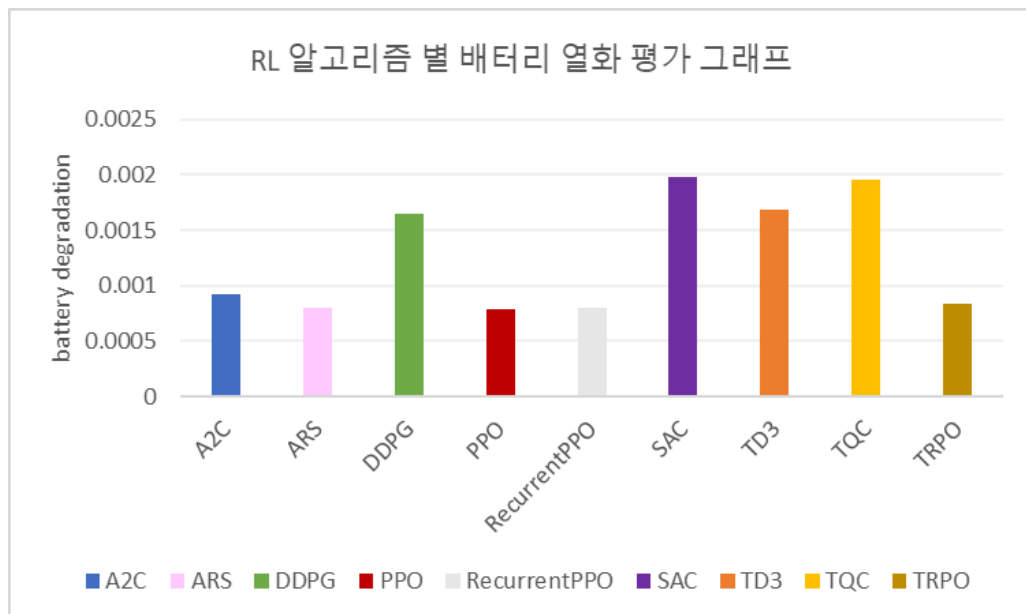
(1) 전력망 운영자

a. 전력망 운영자 - 1

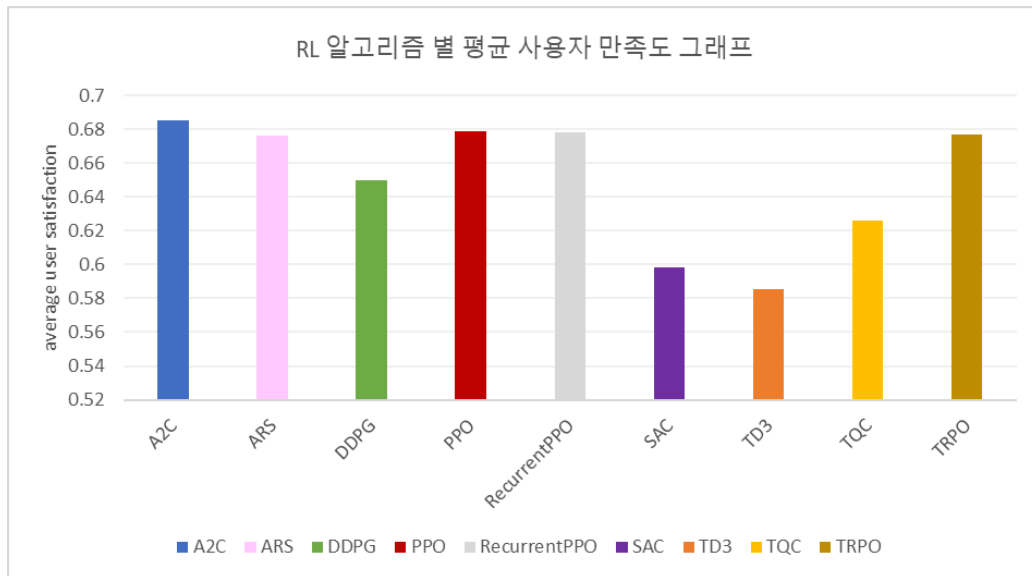
전력망 운영자의 입장에서 전력 사용량, 변압기 과부하, 배터리 SoC, 예상 충전량 초과, 충방전 사이클, 사용자 만족도 6가지 값을 고려하는 보상 함수를 설계했다.

시스템이 목표 전력에 최대한 근접하게 도달하여 운영될 수 있도록 목표 전력과 현재 전력 사용량의 차이를 제공하여 **reward**에 반영하였다. 변압기의 과부하를 최소화하기 위해, 변압기의 과부하 정도에 따라 **penalty**를 부여할 수 있도록 하였다. SoC의 값이 20~80% 사이일 때 배터리의 효율성을 극대화할 수 있으므로, 해당 값을 초과할 경우 **penalty**를 부여하였다. 또한 과도한 충전으로 인한 시스템 부하를 방지하기 위해 예상 충전량 보다 충전 전력을 초과한 경우, EV의 배터리 열화를 최소화하기 위해 충방전 사이클 수가 많을수록 **penalty**를 부과하여 시스템의 효율성과 사용자 만족도를 동시에 높일 수 있도록 고려하였다.

20개의 충전기로 하루 15분 간격 총 85단계, 9가지 RL 알고리즘에 대해 각 5,000회 학습 후 평가해보았다.



[그림 10] RL 알고리즘 별 배터리 열화 평가 그래프



[그림 11] RL 알고리즘 별 평균 사용자 만족도 그래프

해당 그래프의 결과에서 또한 배터리 열화 지표가 낮을수록 사용자 만족도가 상대적으로 높은 수치를 보이는 것을 알 수 있었다.

[그림 7]의 배터리 열화 지표와 비교했을 때 [그림 10]이 0.0015 ~ 0.0035 사이의 결과를 보이던 지표 값이 0.0005 ~ 0.0025 사이로 개선된 결과를 보였음을 확인할 수 있다. 특히

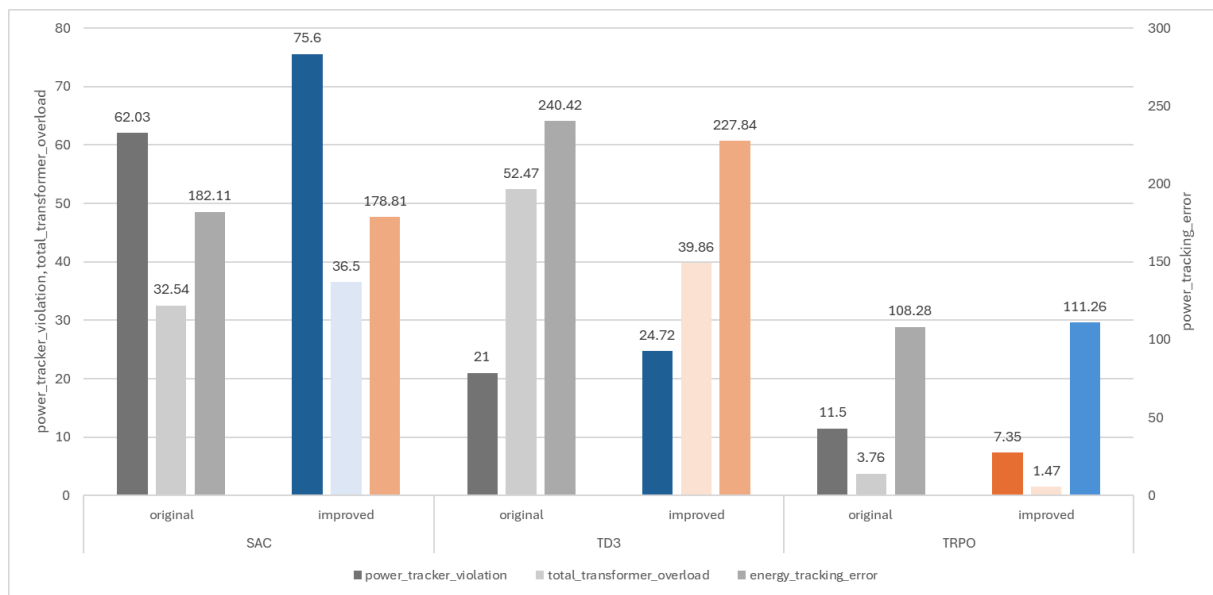
다만, 사용자 만족도의 수치가 약 0.5~0.75%로 낮은 결과를 보이고 있다. 낮은 만족도 수치는 사용자 경험을 저하하고 시스템 성능에 부정적인 요인으로 작용할 수 있어 해당 부분에 있어서는 사용자 만족도를 높일 수 있도록 보완할 것이다.

b. 전력망 운영자 - 2

전력망 운영자의 입장에서는 전력망 안정성, 변압기 과부하 방지, 전력 수요 및 공급 예측 정확도가 중요할 것이라 파악했다. 따라서 `power_tracker_violation`, `total_transformer_overload`, `energy_tracking_error`를 중심으로 개선을 위해 노력했으며, 강력한 제어 및 안정성 확보가 필요한 환경에서 효과적인 알고리즘인 TD3, SAC, TRPO를 이용해 결과를 비교해보았다.

전력망 운영자 관점에서 전력망의 안정성, 변압기 과부하, 사용자 만족도 지표가 중요하다는 판단 하에 최적의 보상이 가능하도록 설계하였다.

총 비용(**total_costs**)를 **reward**의 초기값으로 설정하여 시스템의 운영에서 발생하는 총 비용을 보상의 기준으로 두었다. 변압기가 과부하 정도에 따라 기본 **penalty**를 부여하고, 과부하 상태일 경우 추가로 1000의 값을 더 부과함에 따라 변압기가 최소 전력 제한을 초과하는 것을 방지하도록 하였다. 또한 사용자 만족도도 보상 지표로 두어, 사용자 만족도가 낮은 경우 더 큰 **penalty**를 부과함으로써 총 비용을 절감함과 동시에 높은 사용자 만족도가 도출될 수 있도록 하였다.



[그림 12] 알고리즘별 원본 및 개선 상태에서의 주요 지표 비교: SAC, TD3, TRPO

그래프에서 볼 수 있듯이, 개선한 상태함수, 보상함수를 적용했을 때, **SAC**의 경우 에너지 추적 오차가 감소했고, **TD3**에서는 변압기 과부하 및 에너지 추적 오차를 줄일 수 있었다. **TRPO**의 경우, 전력 추적 위반 및 변압기 과부하의 감소를 확인할 수 있었다.

(2) 전기차 사용자

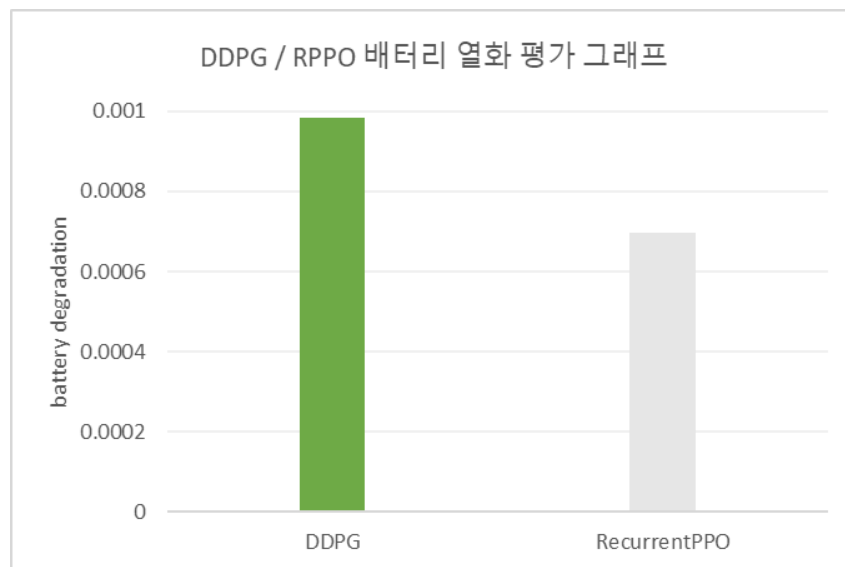
a. 전기차 사용자 - 1

전기차 사용자의 관점에서 배터리 상태, 충전 상태, 사용자 만족도가 가장 중요하다고 판단했다. 이에 따라 **SoC**의 상한(0.8)과 하한(0.2)을 설정하여 해당 수치를 초과하였을 때 **penalty**를 부과하고 추가로 **SoC**가 0.3~0.7 사이에 있을 경우 추가 보상을 제공하였다. 배터리의 과도한 충전방전을 방지하고 배터리 수명과 효율성을 극대화할 수 있는 적정 범위를 유지하도록 유도하였다.

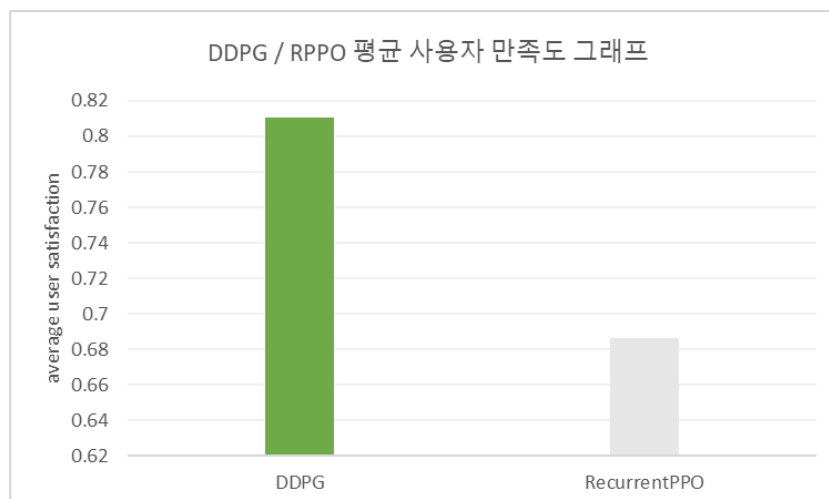
배터리 열화의 경우 **calendar** 열화와 **cycle** 열화 모두 고려하여 열화 값을 **penalty**로 부여함으로써 배터리 상태를 최적화할 수 있도록 설계하였다.

또한 사용자 만족도에서 만족도가 낮을수록 **penalty**를 기하급수적으로 증가시켜 사용자 경험의 질을 개선하도록 유도하였다.

사용자 만족도가 중요한 지표로, 연속적인 행동을 고려하며 경험 재생 메모리(**replay buffer**)를 사용해 학습을 하는 **DDPG**와 안정적이고 사용자 행동의 일관성을 유지하여 성능을 최적화할 수 있는 **RPPO** 알고리즘 두 가지를 중점으로 평가해보았다.



[그림 13] DDPG / RPPO 배터리 열화 평가 그래프



[그림 14] DDPG / RPPO 평균 사용자 만족도 그래프

DDPG 알고리즘의 경우 전체적으로 높은 사용자 만족도를 유지하지만 배터리 열화가 상대적으로 크고 변동값이 크다는 단점을 보이고 있다.

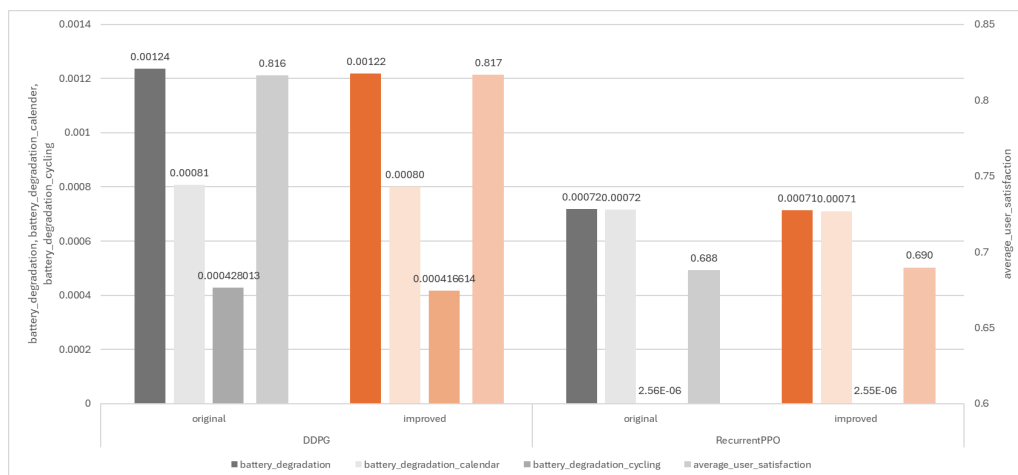
RPPO 알고리즘의 경우 배터리 열화 측면에서는 안정적인 결과를 보이거나 사용자 만족도에서는 다소 낮은 결과를 보이고 있다.

해당 평가를 바탕으로 사용자 만족도 지표가 다소 낮은 결과를 보이고 있어, 사용자 만족도에 대한 보상함수 개선을 통해 사용자 만족도와 배터리 열화 사이에서 균형있는 결과를 만들어 성능을 향상시킬 것이다.

b. 전기차 사용자 - 2

전기차 사용자의 입장에서 충전 비용 절감, 배터리 수명 연장, 충전 편의성이 중요한 지표라고 생각되어 `discharge_price_factor`, `battery_degradation`, `average_user_satisfaction`을 중점으로 살펴보았다. 다만, `discharge_price_factor`는 1.2로 값이 설정 되어 동일하므로 아래 그래프에서는 제외하고 배터리 열화 및 사용자 만족도 수치만 살펴보았다. DDPG는 연속적인 상태 공간을 처리하는데 유리하고, RecurrentPPO는 시계열 데이터에 효과적이므로 DDPG, RecurrentPPO를 중심으로 결과를 살펴보았다.

상태 함수에서는 충전소와 변압기 데이터를 포함해 개인화된 충전 전략을 지원하고, 에너지 효율성을 최적화했고, 보상 함수는 사용자 대기 시간, 배터리 건강, 에너지 교환량, 가격 차이, 및 전력 효율성을 고려해 최적의 충전 경험을 제공하도록 설계해 보았다.



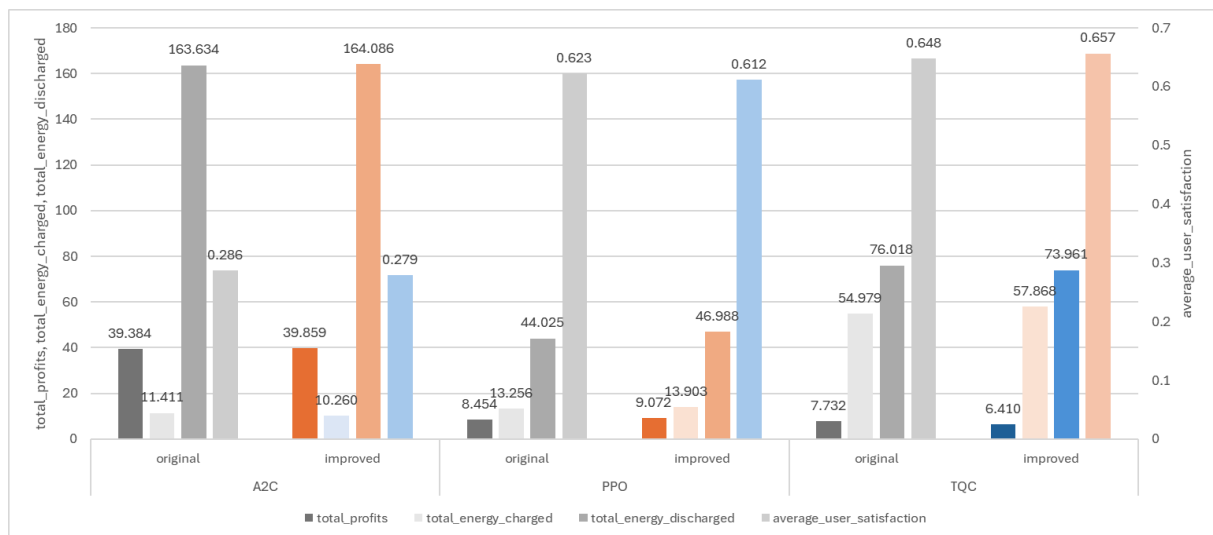
[그림 15] 알고리즘별 원본 및 개선 상태에서의 주요 지표 비교: DDPG, RecurrentPPO

위 그래프를 참고하면, 개선한 상태 함수 및 보상 함수를 적용했을 때, DDPG의 경우 배터리 열화와 사용자 만족도 모든 면에서의 개선을 확인할 수 있다. RecurrentPPO 또한 준수한 결과를 보였으나 battery_degradatoin_cycling 면에서만 소폭 하락한 모습을 보였다.

(3) 충전 서비스 사업자

충전 서비스 사업자에게는 총 수익, 충전소의 운영 효율성, 고객 만족도가 중요할 것이라 생각했다. 따라서, total_profits, total_energy_charged, total_energy_discharged, average_user_satisfaction을 중심으로 살펴보고, 정책 최적화에 강한 PPO와 A2C, 수익 최적화에 적합한 TQC를 이용해 결과를 비교해보았다.

상태 함수는 충전 및 방전 관련 데이터를 포함해 전력망과 충전소의 효율성을 극대화하려 하였고, 보상 함수의 경우 총 비용 절감, 충전 에너지 보상, 에너지 효율성 유지, 그리고 사용자 만족도를 반영하여 전체 운영 성과를 극대화하기 위해 노력하였다.



[그림 16] 알고리즘별 원본 및 개선 상태에서의 주요 지표 비교: A2C, PPO, TQC

그래프에서 볼 수 있듯이, 개선한 상태 함수 및 보상 함수를 적용했을 때, A2C의 경우 총 수익, 방전된 총 에너지가 증가했음을 확인할 수 있다. PPO의 경우, 총 수익, 충전된 총 에너지, 방전된 총 에너지가 증가했으며, TQC에서는 충전된 총 에너지 및 사용자 만족도를 높이는 성과를 보였다.

(4) 보완할 점

V2G의 이해관계자는 크게 전력망 운영자 / 전기차 사용자 / 충전 서비스 사업자 크게 3가지로 나눌 수 있으며, 현재 각각의 입장에서 알고리즘을 설계해 평가를 계속 진행하고 있다. 이를 통해 적절한 접점을 찾아 배터리 효율을 고려한 V2G 충전 알고리즘을 설계할 예정이며, 전기차 충전 요금, 지역의 계통 에너지 패턴, 그리고 전기차 소유주에 대한 혜택 등을 고려한 알고리즘 설계 또한 진행해 볼 예정이다. 각 알고리즘이 가지는 장단점을 보완하여 전력망 운영자, 충전 서비스 사업자, 전기차 사용자 모두의 입장을 균형있게 고려한 최선의 알고리즘을 도출해 낼 계획이다.

3. 구성원 별 역할 분담

이름	역할
이선진	<ul style="list-style-type: none">- V2G 관련 논문 및 이론 학습- EV2Gym 논문 및 학습- 배터리 효율 보상/상태 함수 설계- 스마트 그리드 알고리즘 설계 및 평가- 모델 평가 및 분석- 중간 보고서 작성- 최종 포스터 제작
이지은	<ul style="list-style-type: none">- V2G 관련 논문 및 이론 학습- EV2Gym 논문 및 학습- 배터리 효율 보상/상태 함수 설계- 스마트 그리드 알고리즘 설계 및 평가- 모델 평가 및 분석- 중간 보고서 작성- 최종 보고서 작성

4. 향후 진행 방향

1) 일정 계획

개요	세부항목	8				9				10			
		1	2	3	4	1	2	3	4	1	2	3	4
설계	배터리 충방전 알고리즘												
	스마트 그리드 알고리즘												
보고서	중간보고서												
	최종보고서												

2) 개발 계획

1. 배터리 효율을 고려한 V2G 충방전 알고리즘 설계 - 상태/보상 함수 통합

이해관계자별로 고려했던 상태 및 보상 함수를 기반으로, 이들 요소를 적절히 조합하여 모든 이해관계자에게 유리한 최적의 알고리즘을 개발할 예정이다.

2. 전기차 충전 요금, 목표 지역의 계통 에너지 패턴, 그리고 전기차 소유주에 대한 혜택 등을 고려한 알고리즘 설계

알고리즘의 유효성을 평가하기 위해 다양한 날짜를 설정하여 시뮬레이션을 진행할 계획이다. 특정 날짜 및 시간대에 따른 데이터의 변동성을 분석하고, 이에 따른 전기차 소유주에 대한 혜택을 극대화할 수 있는 알고리즘을 설계하고자 한다.

5. 참고 문헌

[1]

https://tb.kibo.or.kr/ktbs/board/tech-trend/tech_trend.do?mode=download&articleNo=569&attachNo=1447

[그림 2]

<https://erbis.com/blog/the-top-5-v2g-startups-will-you-join-the-energy-industry-revolution/>

<https://ev.or.kr/nportal/evcarInfo/initEvcarSupplyPurposeAction.do>

https://tb.kibo.or.kr/ktbs/board/tech-trend/tech_trend.do?mode=view&articleNo=569&article.ofset=0&articleLimit=10&srSearchVal=v2g&srSearchKey=article_title

<https://github.com/StavrosOrf/EV2Gym/tree/main>

<https://arxiv.org/abs/2404.01849>