

# 강화학습 기반 배터리-슈퍼커패시터 하이브리드 에너지 관리 시스템

안 태 옥<sup>1)</sup> · 임 완 수<sup>\*2)</sup>

금오공과대학교 컴퓨터소프트웨어공학과<sup>1)</sup> · 성균관대학교 전자전기공학부<sup>2)</sup>

## Battery-Supercapacitor Hybrid Energy Management System Based on Reinforcement Learning

Taewook Ahn<sup>1)</sup> · Wansu Lim<sup>\*2)</sup>

<sup>1)</sup>Department of Computer Software Engineering, Kumoh National Institute of Technology, Gyeongbuk 39177, Korea

<sup>2)</sup>School of Electronic and Electrical Engineering, Sungkyunkwan University, Gyeonggi 16419, Korea

(Received 16 May 2024 / Revised 9 August 2024 / Accepted 2 September 2024)

**Abstract** : Batteries and supercapacitors are gaining attention as energy sources for electric vehicles due to their complementary characteristics in energy density and power density. To effectively utilize the advantages of each energy source, a hybrid energy management system that can distribute the required power between batteries and supercapacitors is essential. This paper proposed a reinforcement learning-based hybrid energy management system and a reward function that enhances a battery's state of health(SOH). Simulation results demonstrate that the suggested energy management system (EMS) maintains higher battery SOH than existing EMSs.

**Key words** : Battery(배터리), Electric vehicle(전기자동차), Hybrid energy management system(하이브리드 에너지 관리 시스템), Reinforcement learning(강화학습), Supercapacitor(슈퍼커패시터)

### 1. 서론

전기자동차(EV, Electric Vehicle)는 주 에너지원으로 배터리를 사용하며 최근에는 슈퍼커패시터, 연료 전지와 같은 다양한 에너지원을 함께 사용해 하이브리드 에너지 팩을 구성하고 에너지 효율 및 배터리 수명을 늘리는 연구가 활발히 진행되고 있다.<sup>1-4)</sup>

리튬-이온 배터리는 다른 종류의 배터리에 비해 전력 밀도와 에너지 밀도가 높아서 전기자동차에서 많이 활용되고 있지만, 슈퍼커패시터에 비해서는 전력 밀도가 낮으며 상대적으로 수명이 짧다는 문제가 있다. 반대로 슈퍼커패시터는 전력 밀도가 높고 충전 시간이 짧지만 배터리에 비해 에너지 밀도가 낮다는 문제가 있다. 하지만 배터리와 달리 전기 화학적 반응이 일어나지 않아 수명이 길고, 반응 속도가 짧아 가속에 필요한 급격한 전력 공급에 적합하다. 이처럼 배터리와 슈퍼커패시터는

상호 보완적인 특성을 지니고 있으므로, 두 에너지원으로 하이브리드 에너지 팩을 구성하면 전기자동차에서 유용하게 사용할 수 있다. 하지만 하이브리드 에너지 팩을 구성하는 각 에너지원의 장점을 모두 활용하려면 주행 시 각 에너지원 별 부하를 효율적으로 분배할 수 있는 하이브리드 에너지 관리 시스템(HEMS, Hybrid Energy Management System)이 필요하다. 에너지 관리 시스템의 부하 분배 전략은 크게 세 종류로 나눌 수 있는데,<sup>5)</sup> 규칙 기반 전략(Rule-based strategy), 최적화 기반 전략(Optimization-based strategy), 학습 기반 전략(Learning-based strategy)이 있다.

규칙 기반 전략은 “if-then” 프로시저의 집합으로 정의된 휴리스틱 제어 기법이다. 이들은 내부적으로 결정론 기반 전략(Deterministic-based strategy)과 퍼지 논리 기반 전략(Fuzzy logic-based strategy)으로 세분화 할 수 있다.

\*A part of this paper was presented at the KSAE 2023 Fall Conference and Exhibition

\*Corresponding author, E-mail: [wansu.lim@skku.edu](mailto:wansu.lim@skku.edu)

<sup>†</sup>This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

결정론 기반 전략은 항상 고정된 규칙을 기반으로 동작하는 기법이다. 퍼지 논리 기반 전략은 이진 논리와는 달리 진실의 정도를 다루는 기법으로, 특정 원소가 특정 집합에 속하는 정도를 기반으로 동작한다. 규칙 기반 전략들은 최적의 결과를 보장하지는 않지만, 연산량이 적고 구현이 간단하므로 상용 차량에 많이 사용되고 있다.

최적화 기반 전략은 전역 최적해를 달성할 수 있는 제어 기법이다. 대표적인 예시로는 동적 프로그래밍(DP, Dynamic Programming)이 있다. 동적 프로그래밍은 주어진 문제를 여러 개의 작은 문제로 나누어 먼저 해결하고, 그 결과를 이용해 전체 문제에 대한 최적해를 찾아낸다. 하지만 최적화 기반 전략은 전체 주행 사이클과 전력 소비량 Profile과 같은 미래의 정보들이 필요하고 연산량이 많으므로 실시간 제어 방법으로는 사용할 수 없다.

학습 기반 전략은 인공지능 모델을 사용하는 기법이다. 대표적인 예시로는 머신 러닝 학습 방법 중 하나인 강화학습이 있다. 강화학습은 Agent가 Environment와 상호작용하며 특정 State에서 선택할 수 있는 최적의 Action을 학습하는 것이다. Environment는 Agent의 Action이 얼마나 적합했는지 평가할 수 있는 지표인 보상을 피드백으로써 제공한다. Agent는 매번 입력되는 State에 따라 보상을 최대화할 수 있는 Action을 학습하여 최적에 가까운 부하 분배 전략을 학습할 수 있다.

본 논문은 학습 기반 전략 중 하나인 강화학습 기반 전략(Reinforcement learning-based strategy)을 사용한 배터리-슈퍼커패시터 하이브리드 에너지 관리 시스템을 제안한다. 슈퍼커패시터를 효율적으로 사용하기 위한 보상 함수를 설계하고, Agent는 그 보상 값을 최대화할 수 있는 부하 분배 전략을 학습하여 배터리 수명을 늘릴 수 있다.

## 2. 제안한 하이브리드 에너지 관리 시스템

Fig. 1은 제안한 강화학습 기반 하이브리드 에너지 관리 시스템의 전체 개요도이다. EMS 내부에 위치한 SAC(Soft Actor Critic) Agent는 State로 현재 에너지 팩의 상태와 주행에 필요한 요구 전력을 입력받는다. SAC Agent 내부의 Actor 신경망은 입력 받은 State에 따라, 슈퍼커패시터가 요구전력에서 담당할 비율인 Action을 출력하고, Critic 신경망은 State와 Action을 입력받아 Actor 신경망이 출력한 Action이 얼마나 적합했는지 평가할 수 있는 Q-value를 근사한다. Action에 따라 각 에너지원이 담당할 전력을 계산한 이후에는 각 에너지원에 연결된 양방향 DC-DC 컨버터를 제어하여 모터로 전력을 공급하거나 회생제동으로 전력을 회수한다. 이와 같이 각 에

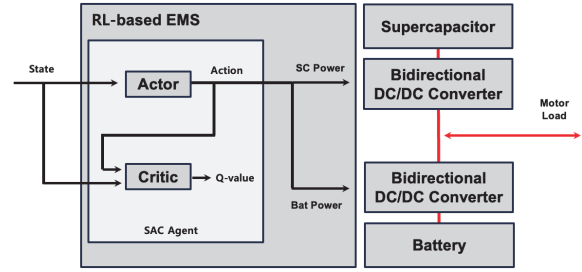


Fig. 1 Proposed RL-based EMS overview

너지원이 양방향 DC-DC 컨버터에 연결되어 있기 때문에 가속 시 발생하는 요구 전력이나 회생제동을 통해 얻는 전력을 배터리와 슈퍼커패시터로 분배할 수 있다.

### 2.1 하이브리드 에너지팩 모델

#### 2.1.1 배터리 모델

배터리 모델은 Simulink Simscape 라이브러리의 Battery 모델을 사용한다. 배터리의 상태를 나타내기 위한 수식들은 다음과 같다.

$$SOH_B(t) = \frac{1}{1 + 0.0005 \times N_B(t)} \quad (1)$$

식 (1)은 배터리 방전 사이클 수에 따른 배터리 State of Health(SOH)를 계산하는 수식이다.  $N_B$ 는 배터리의 방전 사이클 횟수로,  $N_B$ 가 커질수록 SOH는 1에서부터 점차 감소하게 된다. 리튬-이온 배터리는 용량이 초기 용량 대비 20% 감소하면 EOL(End of Life)에 도달했다고 간주하는데,<sup>7)</sup> 식 (1)을 따르면 500회의 방전 주기를 거쳤을 때 80%의 SOH를 갖도록 모델링된다.

$$N_B(t) = \int_0^t \frac{I_B(t) \times H(I_B(t))}{C_B \times SOH_B(t)} dt \quad (2)$$

$$H(x) = \begin{cases} 0, & (\text{if } x < 0) \\ 1, & (\text{if } x \geq 0) \end{cases} \quad (3)$$

식 (2)는 전류 적산법을 통해 배터리의 방전 사이클 수를 계산하는 수식이다. 방전되는 전류만 측정하기 위해 식 (3)의 헤비사이드 함수를 사용하고, 배터리 열화로 인한 용량 감소를 반영하기 위해 초기 용량  $C_B$ 에 SOH를 곱한 값을 나눈다.

$$SOC_B(t) = SOC_B(0) - \int_0^t \frac{I_B(t)}{C_B \times SOH_B(t)} dt \quad (4)$$

식 (4)는 배터리 State of Charge(SOC)를 계산하기 위한 수식이다. 해당 수식에서는 충전 및 방전되는 전류를 모두 포함하기 위해 헤비사이드 함수를 적용하지 않았고, 식 (3)과 동일하게 초기 용량에 SOH를 곱한 값을 나누어 열화로 인한 배터리 용량 감소를 반영한다.

Table 1은 본 논문에서 시뮬레이션에 사용한 배터리의 스펙이다. LG INR18650HG2 리튬이온 배터리 셀의 스펙을 Simulink 모델에 적용하고 직렬과 병렬로 각각 84개와 50개 셀을 연결하여 에너지 팩을 구성한다.

### 2.1.2 슈퍼커패시터 모델

슈퍼커패시터는 Simulink Simscape 라이브러리의 Capacitor 모델<sup>8)</sup>을 사용한다. 커패시터의 상태를 계산하기 위한 수식은 다음과 같다.

$$SOC_{SC}(t) = SOC_{SC}(0) - \int_0^t \frac{I_{SC}(t)}{C_{SC}} dt \quad (5)$$

식 (5)는 슈퍼커패시터의 SOC를 계산하는 수식이다. 슈퍼커패시터는 배터리에 비해 상대적으로 수명이 길기 때문에 본 논문은 슈퍼커패시터 용량 감소를 고려하지 않는다.

Table 2는 본 논문에서 시뮬레이션에 사용한 슈퍼커패시터의 스펙이다. Maxwell BCAP3000 슈퍼커패시터 스펙을 Simulink 모델에 적용하고 222개를 직렬 연결하여 에너지 팩을 구성한다. 슈퍼커패시터 팩은 최소 SOC 제약조건이 존재하는데, 이는 배터리와 슈퍼커패시터의 충방전 특성이 다르기 때문이다. 이들의 충방전 전압 곡선<sup>9)</sup>을 비교해보면, 배터리는 SOC가 감소해도 일정 수준의 전압을 유지할 수 있지만 슈퍼커패시터는 SOC가 감소하면 전압도 선형적으로 감소한다. 따라서 전압 낮은 상황에서 시스템이 요구하는 같은 레벨의 전력을 제공하기 위해 슈퍼커패시터 전류 크기가 높아진다. 이는 시스템의 최대 허용 전류를 초과할 수 있으므로 슈퍼커패시터 팩의 전압이 지나치게 낮아졌을 때 슈퍼커패시터를 사용하지 않게 하도록 최소 SOC 제약조건을 부여하였다.

### 2.1.3 중방향 차량 동역학 모델

주행 사이클에 따른 각 에너지원의 전력 소비를 시뮬레이션하기 위해서는 차량 동역학 모델이 필요하다. 본 논문에서는 Simulink Simscape Driveline 라이브러리의 Longitudinal Vehicle 모델<sup>10)</sup>을 사용하여 전기자동차의 주행을 시뮬레이션 한다. Table 3은 중방향 차량 동역학 모델에서 사용한 파라미터이다.

Table 1 Battery specification

Battery cell voltage	3.6 V
Battery cell capacity	3 Ah
N serial	84
N parallel	50
Battery pack voltage	302 V
Battery pack capacity	150 Ah
Battery pack energy	45.36 kWh

Table 2 Super capacitor specification

SC cell voltage	2.7 V
SC cell capacitance	3,000 F
N serial	222
N parallel	1
SC pack voltage	599 V
SC pack capacitance	13.51 F
SC pack energy	674 Wh
SOC min	0.5

Table 3 Longitudinal vehicle parameter

Vehicle mass	2,400 kg
Tire radius	0.34 m
Tire rolling coefficient	0.0136
Air drag coefficient	0.31
Gravitational acceleration	9.81 m/s <sup>2</sup>
Road grade	0
Motor max power	100,000 W
Motor efficiency	95 %

## 2.2 강화학습 모델

### 2.2.1 Soft Actor Critic Agent

SAC agent는 연속 State 및 Action에서 정책을 학습하는 데 중점을 둔 강화학습 알고리즘이다. SAC는 Actor와 Critic 신경망 기반 강화학습 알고리즘으로, Actor 신경망은 State가 주어졌을 때 Agent의 Action을 근사하고 Critic 신경망은 State와 Action이 주어졌을 때 Q-value를 근사한다. 이와 같이 Actor와 Critic이 상호작용하며 Agent가 Environment로부터 높은 보상을 획득할 수 있는 Action을 파악할 수 있다. SAC의 가장 큰 특징은 학습 시 엔트로피를 추가하여 Agent가 더 많은 Action을 탐색할 수 있도록 유도하는 것이다. SAC는 이를 통해 국소 최적 결과에 수렴하는 문제를 방지하면서 높은 보상을 획득할 수 있도록 학습할 수 있다.

2.2.2 Agent, Environment, State, Action 정의

본 연구에서 Agent는 에너지 관리 시스템 내부의 제어기에 해당하고 Environment는 하이브리드 에너지 팩에 해당한다. Agent가 Environment로부터 얻는 State와 이를 기반으로 출력하는 Action은 다음과 같다.

$$S = \{ P_D, SOC_{SC}, SOC_B \} \tag{6}$$

$$A = \{ R_{SC} \} \tag{7}$$

식 (6)은 State로, 요구 전력(Power Demand)과 배터리와 슈퍼커패시터의 SOC 총 3가지 요소로 구성한다.  $P_D$ 는 요구 전력을 모터의 최대 전력으로 나누어 -1~1 값으로 표현하며 SOC는 0~1 값으로 표현한다. 이를 통해 State 각 요소들의 범위가 달라 신경망 훈련이 제대로 되지 않는 문제를 예방할 수 있다.

식 (7)은 Action으로,  $R_{SC}$ 는 전체 요구 전력에서 슈퍼커패시터가 담당할 비율이며 값은 0에서 1 사이이다. 전체 요구 전력에서 배터리가 담당할 비율은  $1-R_{SC}$ 로 계산한다.

$$P_{SC} = power\ demand \times R_{SC} \tag{8}$$

$$P_B = power\ demand \times (1 - R_{SC}) \tag{9}$$

배터리와 울트라 커패시터가 담당할 비율이 결정된 후에는 식 (8), 식 (9)와 같이 각 에너지원이 담당할 전력을 계산한다. 요구 전력량이 양수라면  $P_B$ 와  $P_{SC}$ 는 각 에너지원이 공급할 전력양이고, 반대로 음수라면  $P_B$ 와  $P_{SC}$ 는 각 에너지원이 회수할 전력양이다.

2.2.3 보상 함수

Fig. 2는 Agent 학습에 사용한 보상 함수이다. 해당 보상 함수는 총 5개 보상 항목을 사용한다. 1~3번 라인은 슈퍼커패시터 방전에 대한 보상으로 슈퍼커패시터의 SOC가 가용 범위 내에 위치하고  $P_D$ 의 미분 값이 0보다 클 때 즉, 요구 전력량이 증가하고 있을 때 보상을 부여한다. 여기서 주어지는 보상은  $P_D$ 와  $R_{SC}$ 를 곱한 값으로, 요구 전력이 높을 때 슈퍼커패시터가 더 많은 양을 보조할수록 높은 보상을 얻을 수 있다. Fig. 3은 3번 라인을 통해 계산되는 보상을 시각화한 그래프로,  $P_D$ 와  $R_{SC}$ 가 모두 1에 가까울수록 높은 보상 값을 획득하는 것을 확인할 수 있다.  $R_{SC}$ 가 동일하더라도  $P_D$ 가 높은 경우에 더 많은 보상을 획득하기 때문에 Agent는  $P_D$ 가 높을수록 더

1. // Discharge Reward
2. if ( $SOC_{SC} > SOC_{SC, Min}$ ) & ( $P_D > 0$ ) & ( $P_{D\_Derivative} > 0$ )
3.     Reward +=  $P_D \times R_{SC}$
- 4.
5. // Charge Reward
6. if ( $SOC_{SC} < 1$ ) & ( $P_D < 0$ )
7.     Reward +=  $R_{SC}$
- 8.
9. // Overdischarge Penalty
10. if ( $SOC_{SC} < SOC_{SC, Min}$ ) & ( $P_D > 0$ )
11.     Reward -=  $R_{SC}$
- 12.
13. // Overcharge Penalty
14. if ( $SOC_{SC} \geq 1$ ) & ( $P_D < 0$ )
15.     Reward -=  $R_{SC}$
- 16.
17. // Basic Reward
18. if (Agent didn't get a penalty)
19.     Reward +=  $SC\_SOC$

Fig. 2 Reward function

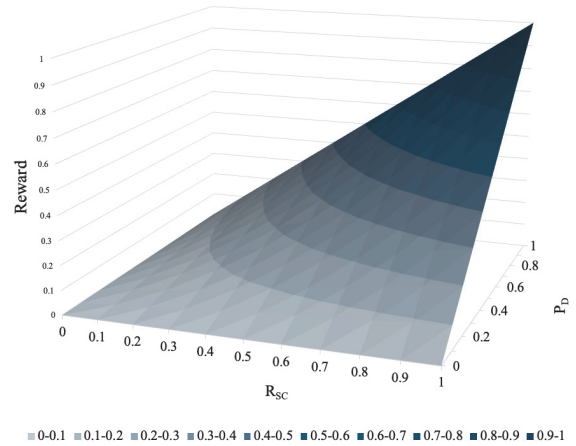


Fig. 3 Charge reward visualization

적극적으로 배터리를 보조하는 전략을 학습하게 된다. 5~7번 라인은 슈퍼커패시터 충전에 대한 보상으로  $P_D$ 가 음수일 때 즉, 모터로부터 회생제동을 통해 에너지 팩을 충전할 때 보상을 부여한다. 이때는  $R_{SC}$ 를 보상 값으로 사용하여 슈퍼커패시터 충전량이 많을수록 높은 보상을 부여한다. 9~11번 라인과 13~15번 라인은 각각 슈퍼커패시터 과방전과 과충전에 대한 패널티로, 보상을  $R_{SC}$ 만큼 감소시킨다. Agent는 이러한 패널티를 최소화하도록 학습하여 슈퍼커패시터의 SOC를 가용 범위 내로 유지하면서 배터리를 보조할 수 있다. 마지막으로

17 ~ 19번 라인은 기본 보상으로 슈퍼커패시터의 SOC가 보상 값으로 주어진다. 이는 Agent가 언제든지 배터리를 보조할 수 있도록 슈퍼커패시터의 SOC를 충분히 유지하도록 하는 역할을 한다. 하지만 Agent가 기본 보상과 과충진/과방전 패널티를 동시에 받으면 패널티를 상쇄하여 Agent가 잘못된 전략을 학습할 확률이 높다. 이를 방지하기 위해 기본 보상은 패널티가 부여되지 않았을 때만 주어진다. Agent는 위와 같이 다양한 조건에 따라 받은 보상과 패널티를 통해 슈퍼커패시터를 어떻게 사용해야 높은 보상을 얻을 수 있는지 학습한다.

### 3. Agent 학습

#### 3.1 주행 사이클 프로파일

본 논문은 Agent의 학습 및 성능 분석에 FTP72, WLTP Class 1, 2, 3 등 총 4개 표준 주행 사이클을 사용하였다.

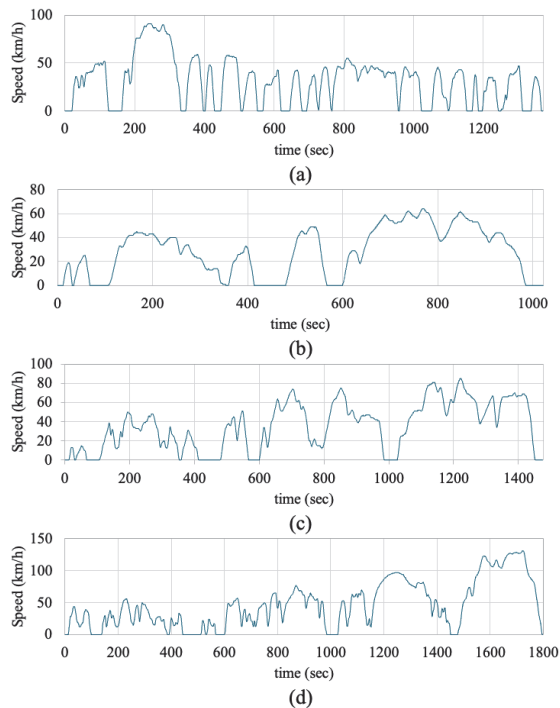


Fig. 4 Driving cycle profiles (a) FTP72 (b) WLTP Class 1 (c) WLTP Class 2 (d) WLTP Class 3

Table 4 Characteristics of driving cycles

Driving cycle	Average speed	Maximum speed	Average discharge power	Average charge power
FTP72	31.4 km/h	91.2 km/h	6,867 W	2,147 W
WLTP class 1	28.5 km/h	64.4 km/h	3,943 W	522 W
WLTP class 2	35.7 km/h	85.2 km/h	6,391 W	1,543 W
WLTP class 3	46.5 km/h	131.3 km/h	10,860 W	2,441 W

Fig. 4는 각 주행 사이클의 속도 프로파일이고 Table 4는 주행 사이클의 출력을 나타낼 수 있는 파라미터로, 평균 속도와 최고 속도 그리고 시뮬레이션 시 측정되는 평균 전력 소비량과 평균 전력 충전량이다. FTP72 사이클은 주행 중 정차와 가속을 빈번하게 반복하고, WLTP Class 1 ~ 3 사이클은 가속과 감속을 반복하며 주행 후반부로 갈수록 최고 속도를 높여 주행한다. FTP72는 평균 31.4 km/h, 최고 91.2 km/h 까지 가속한다. WLTP Class 1은 평균 28.5 km/h, 최대 64 km/h 까지 가속하고 WLTP Class 2는 평균 35.7 km/h, 최대 85 km/h 까지 가속한다. 마지막 WLTP Class 3은 평균 46.5 km/h, 최대 131 km/h 까지 가속한다. Agent는 효과적인 부하 분배 전략을 얻기 위해 다양한 주행 상황을 학습해야 한다. 이를 위해 저속, 중속, 고속 주행이 모두 포함되어 있고 전력 소비와 충전이 가장 활발하게 일어나는 WLTP Class 3 사이클을 Agent 학습에 사용하며, 각 사이클을 모두 시뮬레이션 하여 성능을 분석한다.

#### 3.2 학습 결과

Fig. 5는 에피소드를 150회 학습시키며 측정한 보상을

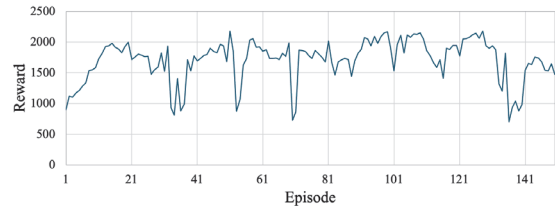


Fig. 5 Rewards for each episode

Table 5 SAC agent parameter

Number of hidden layer	3
Number of hidden units per layer	128
Activation function	ReLU
Learning rate	0.01
Batch size	64
Discount factor	0.99
Target smoothing factor	0.001
Replay buffer size	10,000

나타낸 그래프이다. 학습 초기 약 21회 에피소드를 학습할 때까지는 획득 보상이 급격하게 증가하였고, 이후에는 증가와 감소를 반복하였다. 128번째 에피소드를 학습하였을 때 Agent는 가장 높은 2180 보상을 획득하였고, 해당 시점의 Agent를 사용해 다른 선행연구와 성능을 비교하였다. 학습에 사용한 파라미터들은 Table 5와 같다.

### 3.3 성능 비교

#### 3.3.1 Agent의 분배 전략 분석

Fig. 6은 Agent 학습에 사용했던 WLTP Class 3 주행 사이클을 시뮬레이션 한 결과이다. (a)는 전력 분배 그래프

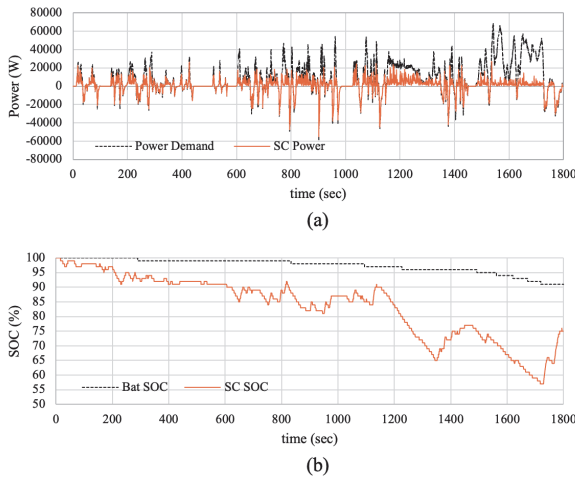


Fig. 6 WLTP Class 3 simulation results with proposed EMS (a) Power Distribution (b) SOC

로 요구 전력과 슈퍼커패시터가 소비한 전력을 나타내고, (b)는 전력 분배에 따른 각 에너지원의 SOC 변화를 나타낸다. Agent는 기본적으로 요구 전력이 증가할 때 슈퍼커패시터의 전력을 소모하여 배터리를 보조하였고, 회생제동으로 얻는 에너지 대부분을 슈퍼커패시터 충전에 사용하였다. 주행 1,500초 이후에는 요구 전력이 가파르게 증가함에도 슈퍼커패시터가 주행 전반부에 비해 적은 전력을 공급하였는데, 이는 슈퍼커패시터 에너지 팩 SOC가 이미 70 %까지 감소한 상태이기 때문이다. Agent는 슈퍼커패시터 에너지 팩의 현재 SOC 상태에 따라 전력 공급량을 제어하거나 회생제동을 통해 충전하여 슈퍼커패시터의 최소 SOC 제약조건을 만족시켰고, 이를 통해 Agent가 보상 함수의 의도를 잘 학습하였음을 알 수 있다.

#### 3.3.2 기존 EMS와 성능 비교

본 논문은 기존 EMS 2개와 제안한 EMS의 성능을 비교하였다. 첫 번째 비교 대상은 기존의 강화학습 기반 EMS<sup>1)</sup>로 본 논문과 같은 SAC Agent를 사용하지만 학습에 사용한 보상 함수에서 차이가 존재한다. 두 번째 비교 대상은 규칙 기반 EMS<sup>11)</sup>로 가속, 정속, 감속 등 총 3가지 주행 상태에 따라 다른 알고리즘을 사용하여 부하를 분배한다.

Fig. 7은 WLTP Class 2 주행 사이클을 각 EMS로 시뮬레이션한 결과로, 전력 요구량과 각 EMS의 슈퍼커패시터 소비전력 그리고 SOC 그래프이다. (a)와 (c)에서 제안한 EMS와 규칙 기반 EMS는 회생제동으로 얻는 에너지 대부분을 슈퍼커패시터 충전에 사용했다. (b)에서 강화

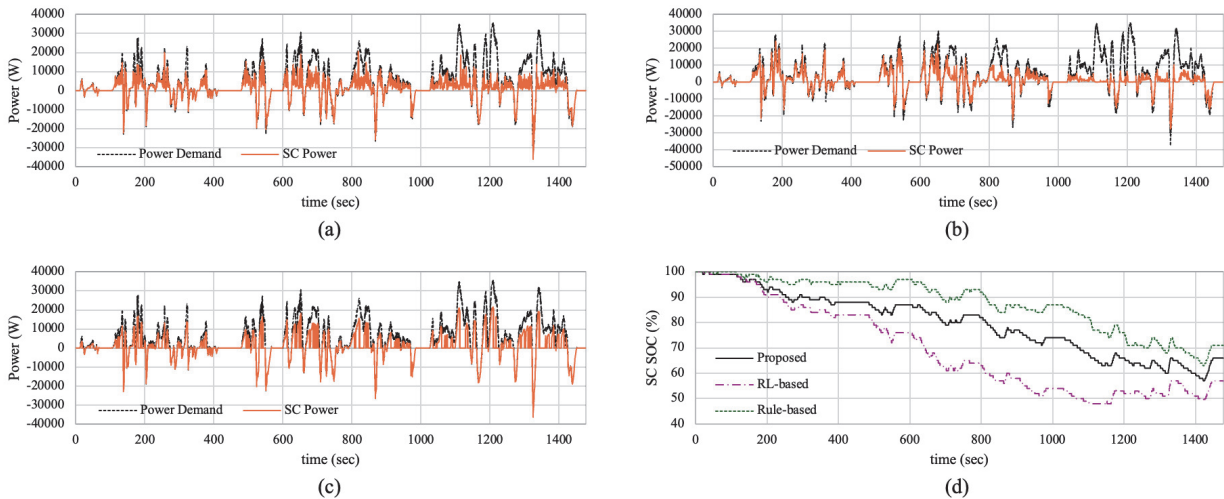


Fig. 7 Performance comparison with other studies (a) Proposed EMS power distribution (b) RL-based EMS power distribution (c) Rule-based EMS power distribution (d) Supercapacitor SOC Comparison

학습 기반 EMS 역시 회생제동 시 슈퍼커패시터를 위주로 충전하지만 (a), (c)에 비해 상대적으로 충전량이 적다. (a)와 (c)에서 제안한 EMS와 규칙 기반 EMS는 요구 전력이 증가할 때 슈퍼커패시터가 모든 전력을 담당하는 것이 아니라 일정량을 나누어 보조하고 있다. 하지만 (b)에서 강화학습 기반 EMS는 슈퍼커패시터에서 전력을 소비할 때 요구 전력 대부분을 슈퍼커패시터로 공급하고 있다. (d)는 각 EMS의 슈퍼커패시터 SOC 그래프이다. 제안하는 EMS와 규칙 기반 EMS는 최소 SOC 제약 조건을 만족시켰으나 강화학습 기반 EMS는 일부 구간에서 SOC가 50 % 이하로 감소하는 모습을 보였다.

Table 6은 FTP72, WLTP Class 1 ~ 3 주행 사이클을 시뮬레이션 한 결과를 표로 나타낸 것이다. 주요 성능 비교 파라미터는 슈퍼커패시터의 평균 전력 소비량과 충전량 그리고 배터리의 평균 전력 소비량과 SOH이다. 이때 배터리 SOH는 주행 사이클 시뮬레이션 한 번으로 비교하기 어렵기 때문에 해당 파라미터는 10,000번의 시뮬레이션을 진행한 결과를 사용하여 비교하였다.

FTP72 주행 사이클에서는 강화학습 기반 EMS가 2,859 W로 가장 많은 슈퍼커패시터 평균 전력 소비량을 기록했다. 그로 인해 배터리의 평균 전력 소비량은 4016 W로 가장 낮았고 배터리 SOH는 83 %로 가장 높았다. 하지만 슈퍼커패시터 평균 전력 충전량은 1624 W로 가장 적었기 때문에 슈퍼커패시터의 최소 SOC가 48.5 %까지 감소하여 최소 SOC 제약 조건을 만족시키지 못하였다.

제안하는 EMS와 규칙 기반 EMS는 각각 배터리 SOH가 82.7 %, 82.8 %까지 감소하여 강화학습 기반 EMS와 유사한 성능을 보이면서 슈퍼커패시터의 최소 SOC가 모두 70 % 이상으로 최소 SOC 제약 조건을 만족하였다.

WLTP Class 1 주행 사이클에서도 강화학습 기반 EMS가 2,226 W로 가장 많은 슈퍼커패시터 평균 전력 소비량을 기록하였고 배터리 SOH는 94.6 %로 가장 높았다. 하지만 슈퍼커패시터 최소 SOC가 47.3 %까지 감소하여 최소 SOC 제약 조건을 만족시키지 못하였다. 제안하는 EMS는 슈퍼커패시터 평균 전력 소비량이 2,105 W이고, 배터리 SOH는 94.2 %까지 감소하여 강화학습 기반 EMS와 유사한 성능을 보였다. 또한 슈퍼커패시터의 최소 SOC는 53.8 %로 최소 SOC 제약 조건 역시 만족하였다. 규칙 기반 EMS는 슈퍼커패시터 평균 전력 소비량이 1474 W로 가장 적었고 그로 인해 배터리 SOH는 92.2 %까지 감소하여 가장 낮은 성능을 보였다.

WLTP Class 2 주행 사이클에서는 모든 EMS의 슈퍼커패시터 평균 전력 소비량 편차가 가장 작았고 배터리 SOH에서도 유사한 성능을 보였다. 하지만 제안하는 EMS와 규칙 기반 EMS는 슈퍼커패시터 최소 SOC를 만족하였지만 강화학습 기반 EMS는 슈퍼커패시터의 최소 SOC가 47.3 %로 제약조건을 만족시키지 못하였다.

마지막 WLTP Class 3 주행 사이클에서는 규칙 기반 EMS가 3116 W로 가장 많은 슈퍼커패시터 평균 전력 소비량을 기록하였고 배터리 SOH는 56.8 %로 가장 높았

Table 6 Performance comparison with other studies

		Supercapacitor				Battery		
		SOC <sub>Final</sub>	SOC <sub>Min</sub>	Average charge power	Average discharge power	SOC <sub>Final</sub>	Average discharge power	SOH
FTP72	Proposed	75.2 %	71.4 %	1,978 W	2,768 W	96.7 %	4,099 W	82.7 %
	RL-based <sup>1)</sup>	56.5 %	48.5 %	1,624 W	2,859 W	97.0 %	4,016 W	83.0 %
	Rule-based <sup>11)</sup>	80.4 %	76.6 %	2,163 W	2,815 W	96.5 %	4,055 W	82.8 %
WLTP class 1	Proposed	54.8 %	53.8 %	438 W	2,105 W	98.9 %	1,837 W	94.2 %
	RL-based <sup>1)</sup>	48.4 %	47.0 %	402 W	2,226 W	98.9 %	1,718 W	94.6 %
	Rule-based <sup>11)</sup>	77.5 %	76.2 %	523 W	1,474 W	98.4 %	2,467 W	92.2 %
WLTP class 2	Proposed	65.2 %	57.1 %	1,396 W	2,356 W	96.4 %	4,040 W	81.7 %
	RL-based <sup>1)</sup>	56.4 %	47.3 %	1,154 W	2,292 W	96.6 %	4,107 W	81.3 %
	Rule-based <sup>11)</sup>	71.8 %	63.6 %	1,543 W	2,357 W	96.3 %	4,042 W	81.7 %
WLTP class 3	Proposed	75.4 %	56.8 %	2,256 W	2,864 W	91.3 %	8,023 W	55.5 %
	RL-based <sup>1)</sup>	70.3 %	49.1 %	1,837 W	2,550 W	91.4 %	8,343 W	53.8 %
	Rule-based <sup>11)</sup>	72.5 %	50.0 %	2,432 W	3,116 W	91.4 %	7,782 W	56.8 %
Average	Proposed	67.6 %	59.7 %	1,517 W	<b>2,523 W</b>	95.8 %	<b>4,499 W</b>	<b>78.5 %</b>
	RL-based <sup>1)</sup>	57.9 %	47.9 %	1,254 W	2,481 W	95.9 %	4,546 W	78.1 %
	Rule-based <sup>11)</sup>	75.5 %	66.8 %	1,665 W	2,440 W	95.6 %	4,586 W	78.3 %

다. 그 다음으로는 제안하는 EMS가 슈퍼커패시터 평균 전력 소비량 2,864 W, 배터리 SOH 55.5 %를 기록하였다. 강화학습 기반 EMS는 1837 W로 가장 적은 슈퍼커패시터 평균 전력 충전량을 기록하였고 배터리 SOH는 56.8 %까지 감소하여 가장 낮은 성능을 보였다.

네 가지 주행 사이클 시뮬레이션 결과의 평균을 살펴보면 제안하는 EMS, 강화학습 기반 EMS 그리고 규칙 기반 EMS의 배터리 SOH는 각각 78.5 %, 78.1 %, 78.3 %로 유사한 성능을 보였다. 하지만 강화학습 기반 EMS는 슈퍼커패시터 평균 전력 충전량이 상대적으로 적고 최소 SOC 제약 조건을 만족시키지 못하는 문제가 있었고 규칙 기반 EMS는 저속 주행 사이클인 WLTP Class 1에서 슈퍼커패시터 평균 전력 소비량이 적었다는 문제가 있다. 반면 제안하는 EMS는 모든 주행 사이클에서 기존 EMS와 유사한 성능을 보였고 평균값에서는 슈퍼커패시터 평균 전력 소비량이 2,523 W로 가장 많았고 배터리 SOH는 78.5 %로 가장 높았다.

#### 4. 결론

본 논문은 강화학습 기반의 배터리-슈퍼커패시터 하이브리드 에너지 관리 시스템을 제안하였다. 우선 슈퍼커패시터의 충방전에 따른 보상과 패널티를 설계하여 강화학습 Agent가 배터리 수명을 증가시킬 수 있는 부하 분배 전략을 학습하였다. 학습 후에는 여러 주행 사이클을 이용하여 시뮬레이션을 진행해 기존 EMS와 제안한 EMS의 성능을 비교하였다. 제안한 EMS는 회생제동을 통해 얻는 전력을 슈퍼커패시터 충전에 사용하여 슈퍼커패시터의 최소 SOC를 유지하였고 가속 시 배터리를 보조하여 배터리의 평균 전력 소비량을 낮췄다. 제안한 EMS는 모든 주행 사이클에서 기존 EMS들과 유사한 성능을 보이면서 슈퍼커패시터 최소 SOC 제약조건을 만족시켰다. 또한 평균값에서는 슈퍼커패시터의 평균 전력 소비량이 가장 높았고 배터리 SOH도 가장 높게 유지하였다.

#### 후 기

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2021R111A3056900, RS-2024-00349885).

#### References

1) D. Kim, S. Hong, S. Cui and I. Joe, "Deep Reinforcement Learning-Based Real-Time Joint Optimal Power Split for Battery-Ultracapacitor

-Fuel Cell Hybrid Electric Vehicles," *Electronics*, Vol.11, No.12, 2022

2) R. Abdelhedi, A. Lahyani, A. Ammari, A. Sari and P. Venet, "Reinforcement Learning-Based Power Sharing between Batteries and Supercapacitors in Electric Vehicles," *IEEE International Conference on Industrial Technology*, pp.2072-2077, 2018.

3) C. U. Udeogu and W. Lim, "Improved Deep Learning-Based Energy Management Strategy for Battery-Supercapacitor Hybrid Electric Vehicle With Adaptive Velocity Prediction," *IEEE Access*, Vol.10, pp.133789-133802, 2022.

4) T. Ahn and W. Lim, "Combining Reinforcement Learning and Rule-Based Strategies for Electric Vehicle Battery-Ultracapacitor Hybrid Energy Management System," *KSAE Annual Conference Proceedings*, pp.1645-1650, 2023.

5) A. Gautam, M. Tariq, J. P. Pandey, K. S. Verma and S. Urooj, "Hybrid Sources Powered Electric Vehicle Configuration and Integrated Optimal Power Management Strategy," *IEEE Access*, Vol.10, pp.121684-121711, 2022.

6) Mathworks, Behavioral Battery Model – MATLAB, <https://www.mathworks.com/help/sps/ref/battery.html>, 2024.

7) J. Kim, J. Park, S. Kwon, S. Sin, B. Kim and J. Kim, "Remaining Useful Life Prediction of Lithium-ion Batteries through Empirical Model Design of Discrete Wavelet Transform Based on Particle Filter Algorithm," *Transactions of KSAE*, Vol.30, No.3, pp.199-206, 2022

8) Mathworks, Linear Capacitor in Electrical Systems – MATLAB, <https://www.mathworks.com/help/sps/ref/capacitor.html>, 2024.

9) W. Jang, S. C. Hong, J. P. Hong, T. Hwang, J. S. Oh, S. Ko, G. Lee, K. Ahn, H. Kim, J. Suhr and J. D. Nam, "Performance Characteristics of Li-ion Battery and Supercapacitor Hybrid Cell for High Power / Low Temperature Discharge," *Transactions of KSAE*, Vol.21, No.6, pp.49-57, 2013.

10) Mathworks, Longitudinal Motion Abstract Vehicle with Ideal Brakes and Tires – MATLAB, <https://www.mathworks.com/help/sdl/ref/longitudinalvehicle.html>, 2024.

11) J. Armenta, C. Nunez, N. Visairo and I. Lazaro, "An Advanced Energy Management System for Controlling the Ultracapacitor Discharge and Improving the Electric Vehicle Range," *Journal of Power Sources*, Vol.284, pp.452-258, 2015.