

신재생 에너지 최적 관리를 위한 강화학습 기반 EMS 시스템 설계

김행복(*), 반승훈, 우근혁, 이민욱, 임수민, 조 민, 권기현

(*) 강원대학교 정보통신공학과, kimhb0117@kangwon.ac.kr

Design of RL EMS System for Optimal Management of Renewable Energy

Kim Haeng-Bok, Ban Seung-Hun, Woo Geun-Hyeok, Lee Min-Uk, Lim Soo-Min, Cho Min, Kwon Ki-Hyeon

Kangwon National University, Department of Electronic Information Communication Engineering

요약

수소 기반 전력 관리 시스템인 H₂EMS를 통해 그리드 전력 비용을 최소화하는 방법을 제안한다. H₂EMS를 통해 그리드 전력의 변동 단가를 고려하여 에너지원 간 밸런스를 조절하기 위하여 강화학습 머신러닝 기법을 사용하였다. 잉여 전력은 수소로 변환하여 저장되고, 그리드 전력의 단가를 고려하여 공급되도록 설계하였다. 실험 결과 강화학습 기반의 H₂EMS 시스템이 그리드 전력 사용량을 절감할 수 있음을 보여준다.

1. 서론

태양광과 수소 등 분산 에너지 자원을 사용하는 소규모 전력망에서는 효율적인 에너지 관리 제어가 필수적이다. 이 연구는 전력망과 분산 에너지 자원의 계획적 사용을 통해 그리드 전력 사용량을 줄이고 전력비용을 낮추는 것을 목표로 한다. 이를 위해 강화학습 기반 수소 에너지 관리 시스템(H₂EMS)을 제안한다. 가상 EMS 환경을 구현하고 Q-learning[1]을 사용하여 최적의 운용 계획을 도출한다. H₂EMS를 통해 신재생 에너지의 변동성을 효과적으로 대응하고, 마이크로 그리드를 제어하여 자체 에너지 활용 극대화하여 계통 안정성을 확보하고 전력 비용을 절감할 수 있다.

2. 연구 내용

2.1 데이터

데이터는 K 대학교 S 캠퍼스 A 건물의 100kW 태양광 발전 데이터를 사용하였으며, A 건물은 '교육용(을) 고압 A 선택 1' 전기요금을 적용받는다[2]. 부하 데이터는 A 건물 소비전력을 토대로 구성했다. 각 데이터는 5분 간격이며, 단위는 kWh이다.

2.2 강화학습 그리드 전력 절감

본 연구에서는 강화학습 방법 중에서 Q-learning을 사용한 비용 최소화 EMS를 제안한다. 실험에서 사용된 Q-learning 모델은 상태(State), 행동(Action), 보상(Reward)으로 구성된다. 상태는 수소 탱크의 충전 상태를 나타내며, 저장 용량은 최대 100kg으로 설정하였고, 초기 충전량은 50%로 시작한다. 상태의 가용 범위는 20%에서 90% 사이로 제한했다. 행동은 표 1에 나열된 옵션을 포함하며, 보상은 단계별 그리드 단가와 그리드 사용량을 기반으로 계산되었다(식 1).

Table 1. Action

표 1. 액션

Action Number	PV	Hydrogen	Grid
1	O	X	X
2	O	O	X
3	O	X	O
4	O	O	O
5	X	O	X
6	X	X	O
7	X	O	O
8	X	X	X

학습 과정에서 에이전트는 ϵ -greedy 정책을 통해 확률 ϵ 로 무작위 행동을 선택하거나 Q-테이블에 따라 행동을 결

정한다(식 2). 이 접근 방식은 그리드 비용 최적화를 위한 전략 결정에 효과적인 학습 경로를 제공한다.

$$\begin{aligned} tP_{grid} &: \text{Use grid power by step} \\ tCost &: \text{Grid prices by step} \\ \text{Reward} &= tP_{grid} \cdot tCost \end{aligned} \quad (1)$$

$$\begin{aligned} \alpha &: \text{Learning rate} \\ \gamma &: \text{Discount rate} \\ \max(Q_{next}) &: \text{The maximum } Q\text{-Value among all possible} \\ &\quad \text{actions from the next state} \\ Q_{new} &= (1-\alpha) \cdot Q_{old} + \alpha (R + \gamma \cdot \max(Q_{next})) \end{aligned} \quad (2)$$

파라미터별 이동 평균 그래프의 기울기에 절댓값을 취한 값을 구한 뒤 최종 비용으로 나눈 값에 AUC를 더해 계산된 값(식 3)을 히트맵으로 나타냈다. 그림 1은 그리드 탐색 히트맵을 나타냈다.

$$\begin{aligned} \text{Slope} &: \text{Norm}(\text{moving average value slope}) \\ \text{Price} &: \text{Norm}(\text{moving avg final grid price}) \\ \text{AUC} &: \text{Norm}(\text{Area under curve}) \\ \text{Value} &= \frac{|\text{Slope}|}{\text{price}} + \text{AUC} \end{aligned} \quad (3)$$

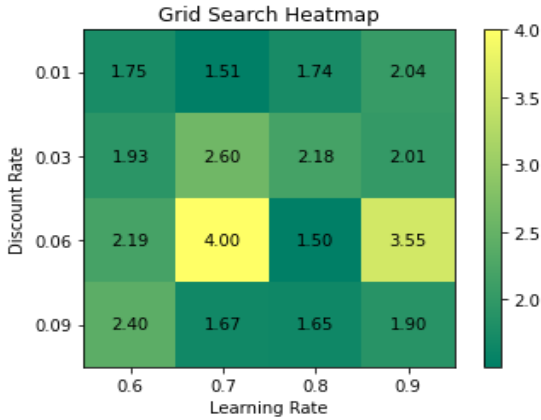


Figure 1. Grid Search HeatMap

그림 1. 그리드 탐색 히트맵

그리드 탐색 히트맵을 통해 성능이 4.0으로 높은 Discount rate 0.06, Learning rate 0.7 파라미터를 사용하여 300,000 에피소드의 강화학습을 진행했다. 그림 2는 휴리스틱과 강화학습을 이용한 그리드 전력, 태양광, Hydrogen(수소) 저장, 수송가의 전력 사용량을 나타낸 것이다.

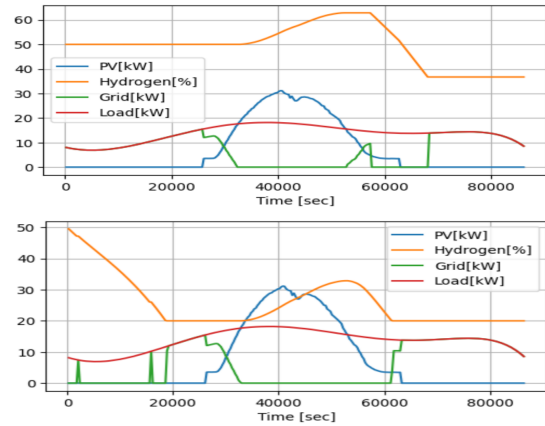


Figure 2. Results: Heuristic(top) RL(bottom)

그림 2. 결과:휴리스틱(상) 강화학습(하)

3. 기대효과 및 활용

본 연구를 통해 에너지 사용 최적화로 전기요금 절감 효과를 기대할 수 있고, 실시간 에너지 관리를 통해 에너지 사용량에 대한 모니터링과 전력 공급의 안정성 및 신뢰성 제 공할 수 있을 것으로 기대된다.

4. 결론

본 연구에서는 태양광 데이터와 부하 데이터를 토대로 계절에 따른 태양광 발전량의 변동을 고려하여 수소연료전 지와 강화학습을 통해 1일 단위의 에너지 관리를 최적화하 였다. 전통적인 방법인 휴리스틱과 전력 요금을 비교했을 때 1일 기준 휴리스틱은 15081.46원, 강화학습은 14972.61원으로 약 0.72% 절감 효과를 확인했다. 이러한 연구 결과는 안정적인 EMS 시스템의 실현 가능성을 보여 주며 범용적으로 적용 가능할 것이며, 실시간 에너지 관리 를 통해 사용량에 대한 파악과 전력 공급의 안정성 및 신뢰 성을 제공할 것으로 예상된다.

감사의 글

본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력 기반 지역 혁신 사업의 결과입니다. (2022RIS-005)

참고 문헌

- [1] E. Samadi, A. Badri and R. Ebrahimpour, "Q-Learning-Oriented Distributed Energy Management of Grid-Connected Microgrid," 2021 29th Iranian Conference on Electrical Engineering (ICEE), Tehran, Iran, Islamic Republic of, 2021, pp. 318-322, doi: 10.1109/ICEE52715.2021.9544152.
- [2] Korea Electric Power Corporation. (n.d.). Electricity Rate Table in Korean, for Educational Use (Type A/Type B). KEPCO ON. <https://online.kepco.co.kr/PRM004D00>