

# 강화학습 기반 V2G Station 연계형 스마트 에너지 빌딩 전력 제어 기법

최석민\* · 김선용\*\*

## Reinforcement Learning Based Energy Control Method for Smart Energy Buildings Integrated with V2G Station

Seok-Min Choi\* · Sun-Yong Kim\*\*

### 요 약

전 세계적으로 전력 소비량이 꾸준히 증가하고 있으며, 특히 빌딩의 전력 소비 비율은 세계 전력 소비 비율의 20% 이상을 차지할 만큼 그 비중이 크다. 이에 따라 빌딩에서의 전력 소비를 효율적으로 관리하는 빌딩 에너지 관리 시스템(BEMS, Building Energy Management System)의 연구 및 개발이 활발히 진행되고 있으며, 특히 최근에는 인공지능 기술의 발달로 인해 Smart BEMS 연구가 주목받고 있다. 본 논문에서는 강화학습 기반 V2G(Vehicle-to-Grid) Station 연계형 스마트 에너지 빌딩 전력 제어 기법을 제안한다. 실제 빌딩의 전력량 데이터 기반 성능평가 결과, 학습이 진행됨에 따라 빌딩에서의 전력 요금이 감축하는 것을 확인하였다.

### ABSTRACT

Energy consumption is steadily increasing, and buildings in particular account for more than 20% of the total energy consumption around the world. As an effort to cost-effectively manage the energy consumption of buildings, many research groups have recently focused on Smart Building Energy Management Systems (BEMS), which are deepening the research depth by applying artificial intelligence(AI). In this paper, we propose a reinforcement learning-based energy control method for smart energy buildings integrated with V2G station, which aims to reduce the total energy cost of the building. The results of performance evaluation based on the energy consumption data measured in the real-world building shows that the proposed method can gradually reduce the total energy costs of the building as the learning process progresses.

### 키워드

Reinforcement Learning, Q-learning, Building Energy Management System, V2G Station, Energy Storage System  
강화학습, Q 러닝, 빌딩 에너지 관리 시스템, V2G 스테이션, 에너지 저장 장치

## 1. 서 론

전 세계적으로 전력 소비량이 증가하고 있으며 이는 1980년대부터 현재까지 꾸준한 상승세를 보인다

[1]. 특히 빌딩은 세계 전력 소비 비율 중 약 20% 이상을 차지할 정도로 그 비중이 높는데, 최근에는 코로나19로 인해 실내 활동이 많아짐에 따라 전체 전력 소비는 감소했으나 빌딩에서의 전력 소비는 약 30%

\* 동서대학교 석사과정 (seokmchoi@gmail.com)

\*\* 교신저자 : 동서대학교 소프트웨어학과

• 접수 일 : 2024. 04. 23

• 수정완료일 : 2024. 05. 18

• 게재확정일 : 2024. 06. 12

• Received : Apr. 23, 2024, Revised : May. 18, 2024, Accepted : Jun. 12, 2024

• Corresponding Author : Sun-Yong Kim

Dept. of Software, Dongseo University

Email : sykim@dongseo.ac.kr

증가했다[2-3]. 이렇듯 꾸준히 증가하는 빌딩의 전력 소비에 비해 빌딩 대부분은 아직 비효율적으로 전력을 소비하고 있다[4]. 이에 따라 빌딩에서의 전력 소비를 효율적으로 관리하는 BEMS가 적용된 스마트 빌딩의 연구·개발이 주목받고 있다[5-6]. 특히 최근에는 인공지능 기술의 발달로 인해 빌딩에서의 전력 소비를 더욱 효율적으로 관리하고 최적화하는 Smart BEMS 연구가 주목받고 있다[7-8]. Smart BEMS 연구의 예시로 구글 딥마인드에서 신경망 알고리즘을 BEMS에 적용하여 데이터 센터에서의 전력을 약 40% 정도 줄이는 데 성공하였고, 구글에서 인수한 NEST사는 온도 조절에 강화학습을 적용하여 사용자의 출/퇴근, 기상/수면시간을 자동으로 파악하여 상황에 맞는 적절한 온도를 설정하는 시스템을 개발했다[9-10]. 또한 Utility, 태양광 발전(PV, Photovoltaics), ESS, V2G 시스템 그리고 V2G Station 등이 연계된 빌딩의 BEMS에 강화학습을 적용하여 효율적인 전력 공급을 통해 빌딩의 전력 요금을 줄이는 연구도 수행되었다[2], [11-12].

본 논문에서는 Smart BEMS 연구의 일환으로 Utility, V2G Station 그리고 ESS와 연계된 빌딩의 전력을 효율적으로 관리하는 강화학습 기반 V2G Station 연계형 스마트 에너지 빌딩 전력 제어 기법을 제안한다. 그리고 실제 빌딩의 전력량 데이터를 기반으로 제안하는 기법의 성능을 평가한다.

## II. 강화학습 기반 빌딩 전력 제어 기법

### 2.1 시스템 환경

그림 1은 본 논문에서 제안하는 기법의 환경을 보여준다. 빌딩은 Utility, V2G Station 그리고 ESS가 연계되어 있다. Utility와 V2G Station은 각각 빌딩에 전력을 공급할 때 타임 스텝  $t$ 에 따라 일정한 전력 요금을 가진다. 이때 타임 스텝  $t$ 는 현재 시간을 의미하고,  $\tau$ 는 각 타임 스텝의 길이를 의미한다. 하루를 주기로 반복되는 타임 스텝  $t$ 를 표현하기 위해  $\hat{t}$ 을 수식 (1)로 정의한다.

$$\hat{t} = t \bmod \frac{h}{\tau} \quad \dots (1)$$

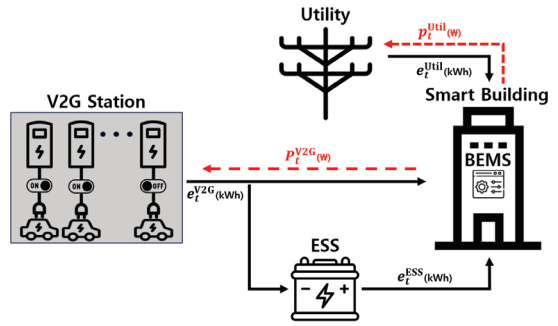


그림 1. Smart BEMS 시스템 환경  
Fig. 1. Environment of Smart BEMS system

여기서  $h$ 는 하루의 시간 수(i.e., 24)를 나타낸다.

Utility의 시간대는 하루를 기준으로 3가지로 구분되며 수식 (2)로 정의한다. 23시부터 08시까지는 전력 요금이 가장 낮은 시간대로 low-load period(경부하), 09시부터 10시, 12시부터 13시, 17시부터 23시는 전력 요금이 중간인 middle-load period(중간부하), 10시부터 12시, 13시부터 17시는 전력 요금이 가장 높은 high-load period(최대부하)로 구분된다.

$$T(\hat{t}) = \begin{cases} L, & \text{if low-load period,} \\ M, & \text{if middle-load period,} \\ H, & \text{if peak-load period,} \end{cases} \quad \dots (2)$$

타임 스텝  $t$ 일 때 빌딩이 소비하는 전력량은  $e^{Bldg}$  (kWh)로 정의한다. 타임 스텝  $t$ 일 때 Utility의 전력 요금 단가는  $p_{T(\hat{t})}^{Util} (\$/kWh)$ 로 정의한다. 빌딩이 Utility로부터 공급받는 전력량은  $e^{Util} (kWh)$ 이며, 이 때 발생하는 전력 요금은  $p_t^{Util} (\$)$ 로 정의한다. 따라서 Utility의 전력 요금 계산식은  $p_t^{Util} = e_t^{Util} \times p_{T(\hat{t})}^{Util}$ 이다.

한편 타임 스텝  $t$ 일 때 V2G Station으로부터 공급받는 전력량은  $e_t^{V2G} (kWh)$ 이며, 이 때 발생하는 전력 요금은  $p_t^{V2G} (\$)$ 로 정의한다. V2G Station은 전력 요금 단가가 항상 일정하다고 가정하며, V2G Station의 소켓을 On/Off 함으로써 빌딩에 공급할 전력량 및 전력 요금이 결정된다. 이때 V2G Station의 'On 상태의 소켓 개수'는  $N$ 으로 정의한다. V2G Station의 소켓 하나로부터 공급받는 전력량의 단위는  $\Delta e^{V2G} (kWh)$

이며, 이에 대한 전력 요금의 단가는  $p^{V2G}$ (₩/kWh)이다. 따라서 V2G Station으로부터 공급받는 전력량 계산식은  $e_t^{V2G} = N \times \Delta e^{V2G}$ 이며, 전력 요금 계산식은  $p_t^{V2G} = e_t^{V2G} \times p^{V2G}$ 이다.

타임 스텝  $t$ 일 때 ESS로부터 공급받는 전력량은  $e_t^{E2B}$ (kWh)이며, ESS에 저장된 전력량은  $e_t^{ESS}$ (kWh)로 정의한다. ESS는 V2G Station이 빌딩에 전력을 공급하고 남은 전력을 저장하는 곳이며, ESS의 최대 저장 가능한 용량(i.e, capacity)은  $E^{ESS}$ 로 정의한다.

## 2.2 제안하는 강화학습 기반 V2G Station 연계형 스마트 에너지 빌딩 전력 제어 기법

본 논문에서는 빌딩의 전력 공급을 제어하는 BEMS에 강화학습 기법의 핵심 알고리즘인 Q-Learning을 적용한다. 제안하는 기법에서 BEMS는 시간대별로 V2G Station의 소켓을 On/Off하고, ESS의 사용 여부를 결정하여 효율적인 전력 공급 제어를 목표로 한다. Q-Learning은 학습 주체(Agent)가 현재 상태(State)를 고려하여 미래 어느 시점까지의 가장 좋은 보상(Reward)을 제공해 줄 것으로 기대되는 행동(Action)을 찾을 수 있는 정책(Policy)을 학습하는 기법이다. 이때, State-Action 조합별 Reward의 기댓값(Q-Value)을 출력해 주는 Q-function을 활용한다.

본 논문에서 제안하는 Q-Learning 알고리즘의 State는 수식 (3)을 따른다.

$$s_t = \{T(\hat{t}), T(t+\tau), e_t^{Bldg}, e_t^{ESS}\} \quad \dots (3)$$

수식 (3)의  $T(\hat{t})$ ,  $T(t+\tau)$ 는 각각 현재 시간대, 다음 시간대를 의미한다.  $s_t$ 의 요소인 시간대를  $T(\hat{t})$ ,  $T(t+\tau)$  두 가지로 구분 짓는 이유는 다음과 같다. 시간대는 경부하, 중간부하, 최대부하 3가지로 구분되는데  $T(\hat{t})$ ,  $T(t+\tau)$ 의 조합은 각 시간대의 경우의 수에 따라 총 9가지의 경우의 수가 존재한다. 예를 들어  $T(\hat{t})$ 은 경부하일 때  $T(t+\tau)$ 는 경부하, 중간부하, 최대부하가 될 수 있으며, 이는 모두 다른 경우의 수를 의미한다. 따라서 시간대를 현재 시간대와 다음 시간대로 구분하였다.  $e_t^{Bldg}$ ,  $e_t^{ESS}$ 는 각각 타임

스텝  $t$ 일 때 빌딩에서 소비하는 전력량, 타임 스텝  $t$ 일 때 ESS에 저장된 전력량을 의미한다.  $e_t^{Bldg}$ ,  $e_t^{ESS}$ 는  $T(\hat{t})$ 일 때의 값만 알 수 있고, 미래인  $T(t+\tau)$ 일 때에는 어떤 값을 가질지 알 수 없는 불확실성을 가지는 요소이다.

Q-Learning 알고리즘의 Action은 수식 (4)를 따른다.

$$a_t = \{N, M, B\} \quad \dots (4)$$

수식 (4)의  $M$ 은 'On 상태의 소켓 개수'  $N$  중 '즉시 빌딩에 전력을 공급하는 소켓 개수'를 의미하고, 이를 통해 빌딩에 즉시 공급할 전력량을 알 수 있다. 예를 들어  $N=4$ ,  $M=3$ 일 때,  $M$ 개의 소켓에 대한  $\Delta e^{V2G} \times 3$ (kWh)만큼의 전력량은 즉시 빌딩에 공급하고,  $N-M$ 개의 소켓에 대한 전력량인  $\Delta e^{V2G} \times (4-3)$ (kWh)는 ESS로 충전된다.

$B$ 는 ESS에 저장된 전력의 사용 여부를 나타내며, 이는 0과 1로 구분된다.  $B=0$ 일 때 빌딩은 ESS에 저장된 전력은 공급받지 않고, V2G Station 혹은 Utility로부터만 전력을 공급받는다. 반면  $B=1$ 일 때 빌딩은 V2G Station으로부터는 전력을 공급받지 않고, ESS 혹은 Utility로부터만 전력을 공급받는다. ESS는 빌딩에 전력을 공급할 때 항상 ESS에 저장된 전력량 중 공급할 수 있는 최대치만큼의 전력량을 공급한다. 예를 들어 빌딩에서 소비하는 전력량이 250(kWh)이고 ESS에 저장된 전력량이 300(kWh)라면, 빌딩은 ESS로부터 250(kWh)를 공급받고 ESS에는 50(kWh)만큼의 전력량이 남는다. 반면 빌딩에서 소비하는 전력량이 250(kWh)이고 ESS에 저장된 전력량이 100(kWh)라면, 빌딩은 ESS에 저장된 100(kWh)만큼의 전력량을 공급받고 나머지 전력량은 Utility로부터 공급받는다.

Q-Learning 알고리즘의 Reward는 수식 (5)를 따른다.

$$r_t = \begin{cases} \max(e_t^{Bldg} - (\Delta e^{V2G} \times M), 0) \\ \times p_{T(\hat{t})}^{Util} + e_t^{V2G} \times p^{V2G}, & \text{if } B=0, \\ \max(e_t^{Bldg} - e_t^{E2B}, 0) \times p_{T(\hat{t})}^{Util}, & \text{if } B=1, \end{cases} \quad \dots (5)$$

수식 (5)의  $r_t$ 는 빌딩이 소비한 전력량에 대한 전력 요금을 나타내며,  $B$ 에 따라 전력 요금이 다르게 계산된다. 빌딩은 소비하는 전력량을 V2G Station 혹은 ESS로부터 우선적으로 공급받아 충당하고, 충당하지 못한 전력량은 Utility로부터 공급받다고 가정한다.

수식 (3), (4), (5)를 통해 도출된 Q-Learning의 최종 업데이트 식은 수식(6)을 따른다.

$$Q(s_t, a_t) = (1 - \beta) Q(s_t, a_t) + \beta[r + \gamma \min_a Q(s_{t+\tau}, a)] \quad \dots (6)$$

---

**Algorithm 1** Proposed Method Using Q-learning
 

---

```

1: Initialize all parameters
2: Initialize learning parameters  $\gamma$ ,  $\beta$  and  $\epsilon$ 
3: for each time step  $t$  do
4:   Decaying  $\beta$  and  $\epsilon$  linearly for 5000 times
5:   Select an action  $a$  by  $\epsilon$  and execute it
6:   Receive immediate reward  $r_t$ 
7:   Observe the future Q-value  $\min_a Q(s_{t+\tau}, a)$ 
8:   Update the table  $Q(s_t, a_t)$  as follows:
9:    $Q(s_t, a_t) \leftarrow (1 - \beta)Q(s_t, a_t) + \beta[r + \gamma \min_a Q(s_{t+\tau}, a)]$ 
10:   $s_t \leftarrow s_{t+\tau}$ 
11: end for
  
```

---

그림 2. 제안하는 기법의 의사코드  
Fig 2. Pseudocode of proposed Method

그림 2는 Q-Learning 알고리즘을 제안하는 기법에 적용했을 때의 의사코드를 표현한 것이다. 우선, 첫 번째 줄에서 시스템 파라미터들의 초깃값을 설정한다. 두 번째 줄에서는 Q-Learning의 학습 파라미터인  $\gamma$ ,  $\beta$  그리고  $\epsilon$ 의 초깃값을 설정한다. 3-11번째 줄은 타임 스텝  $t$ 에 따른 학습 과정을 나타낸다.

### III. 성능평가

#### 3.1 시뮬레이션

본 논문에서 제안하는 강화학습 기반 V2G Station 연계형 스마트 에너지 빌딩 전력 제어 기법의 성능을 평가하기 위해 다음과 같은 환경을 고려하여 시뮬레

이션을 진행하였다. 빌딩에서 소비하는 전력량( $e^{\text{Bldg}}$ ) 데이터는 광주과학기술원(GIST)의 실제 빌딩 데이터를 기반으로 가우스 분포(Gaussian Distribution)를 따라 증강하였다<sup>1)</sup>.

표 1은 시뮬레이션에서 사용된 파라미터의 값을 보여준다.  $Q(s_0, a_0), \forall (s_0, a_0)$  pairs는 타임 스텝  $t$ 가 0일 때 Q-Table의  $Q(s_0, a_0)$ 에 대한 모든  $s_0, a_0$  쌍의 값을 뜻한다.  $Q(s_0, a_0), \forall (s_0, a_0)$  pairs는 충분히 큰 값인  $2 \times 10^6$ 로 설정하였다.  $p_{T(i)}^{\text{Util}}$ 은 한국전력공사의 산업용 전력 요금 단가를 고려하여 설정하였다.  $\Delta e^{\text{V2G}}$ 은 전기자동차의 평균 배터리 용량인 약 70(kWh)를 고려하여 50(kWh)로 설정하였다<sup>2)</sup>. 이때 발생하는  $p^{\text{V2G}}$ (₩/kWh)는

표 1. 시뮬레이션 파라미터  
Table 1. Simulation input parameters

|   |   |
|---|---|
| $Q(s_0, a_0), \forall (s_0, a_0)$ pairs | $2 \times 10^6$   |
| $p_{T(i)}^{\text{Util}}$ (₩/kWh)        | 80, if $T(\hat{t}) = L$ ,<br>130, if $T(\hat{t}) = M$ ,<br>200, if $T(\hat{t}) = H$ , |
| $\Delta e^{\text{V2G}}$ (kWh)           | 50  |
| $p^{\text{V2G}}$ (₩/kWh)                | 100   |
| $E^{\text{ESS}}$ (kWh)                  | 500   |
| $N^{\text{cap}}$                        | 6   |
| $\beta$                                 | 1.0 to 0.1 (decaying)   |
| $\epsilon$                              | 1.0 to 0.1 (decaying)   |
| $\gamma$                                | 0.9   |

Utility의 정부하와 중간부하 사이의 단가인 100 (₩)으로 설정하였다. 위와 같이 설정한 이유는 전기자동차의 소유자가 정부하일 때 전기자동차의 배터리를 충전한다고 가정하면 전기자동차 소유자 및 빌딩 모두 이득을 얻을 수 있는 가격이기 때문이다.  $E^{\text{ESS}}$ 는 500(kWh)로 설정하였으며, V2G Station의 총 소켓 개수인  $N^{\text{cap}}$ 은 6으로 설정하였다. 상대적으로 큰 Q-Table의 크기(270 x 29)를 고려하여, 학습 초반에는 높은 빈도의 탐색을 수행하기 위해  $\epsilon$ 가 decaying(1.0 to 0.1) 하도록 설정하였다. 학습률은 학

1) <https://www.gist.ac.kr/kr/main.html>

2) <https://ev-database.org/cheatsheet/useable-battery-capacity-electric-car>

습 초반에 높은  $\epsilon$  값을 통해 탐색하는 동안, Q-table 업데이트 시 기존의 Q-value 보다 새로운 Q-value의 영향을 더 많이 받도록 decaying(1.0 to 0.1) 하도록 설정하였다.

### 3.2 성능평가 결과

그림 3은 제안하는 강화학습 기반 V2G Station 연계형 스마트 에너지 빌딩 전력 제어 기법의 성능평가 그래프이다. 본 논문에서는 제안하는 기법의 성능평가를 위해 Q-Learning의  $a_t$ 를 랜덤하게 결정하는 Random 기법, Q-Learning에서 Discount factor인  $\gamma$  값을 0으로 설정한 Greedy 기법을 비교군으로 추가했다. 또한, 빌딩이 Utility로부터만 전력을 공급받았을

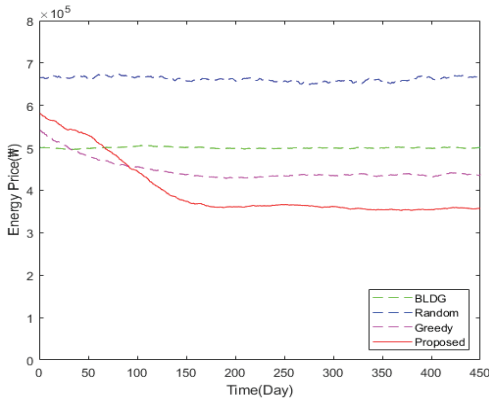


그림 3. 학습에 따른 빌딩의 전력 요금 변화  
Fig 3. Building's energy price via learning process

때의 전력 요금인 BLDG를 추가하여 전력 요금을 비교하였다.

그림 3에서 Random 기법의 평균 전력 요금은 약 660,000(₩)으로, 이는 BLDG보다 높은 수치를 알 수 있다. Greedy 기법은 전력 요금이 제안하는 기법보다 빠르게 감소하지만, 수렴 후 평균 전력 요금은 상대적으로 높은 430,000(₩)이다. 반면 제안하는 기법에서는 전력 요금이 Greedy 기법에 비해 늦게 수렴하지만, 약 180일 이후부터는 약 350,000(₩)으로 수렴하여, BLDG 대비 약 30%, Greedy 기법 대비 약 18%의 전력 요금 감축 효과를 보여준다.

그림 4는 V2G Station의 총 소켓 개수( $N^{cap}$ )에 따른 Random 기법, Greedy 기법 그리고 제안하는 기법의 평균 전력 요금을 비교한 그래프이다. Random 기

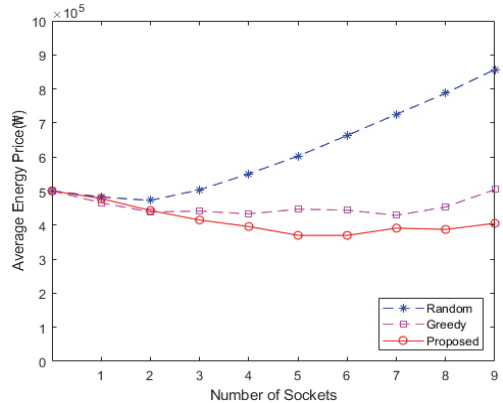


그림 4. V2G Station의 소켓 개수에 따른 평균 전력 요금의 비교

Fig 4. Comparison of average energy price w.r.t. the number of sockets in the V2G station

법에서는 소켓 개수가 2개일 때까지 평균 전력 요금이 감소하지만, 3개 이후로는 꾸준히 상승하는 경향이 보인다. 한편 Greedy 기법에서는 평균 전력 요금이 불규칙한 것을 알 수 있다. 반면 제안하는 기법에서는 소켓 개수가 증가할수록 전력 요금이 꾸준히 감소하다가, 소켓 개수가 6개일 때 평균 전력 요금이 가장 낮아지는 것을 알 수 있다. 그림 4를 통해 제안하는 기법은 본 논문의 환경에서  $N^{cap}$ 이 6일 때 전력 요금이 가장 낮은 것을 알 수 있다. 그러나 빌딩의 규모, ESS의 용량 등 고려하는 환경에 따라 최적의  $N^{cap}$ 은 달라질 것으로 예상된다.

## IV. 결론 및 향후 연구

본 논문에서는 강화학습 기반 V2G Station 연계형 스마트 에너지 빌딩 전력 제어 기법을 제안하였고, 실제 빌딩의 전력량 데이터를 기반으로 성능평가 했을 때 전력 요금이 감축하는 것을 확인하였다. 향후 연구로써, 시간대별 전력 요금이 실시간으로 변하는 환경을 고려하고, PV 등 추가적인 요소가 연계된 빌딩의 환경을 고려하여 심층 강화학습 기법을 적용하는 연구를 수행할 계획이다[13-15].

### 감사의 글

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1G1A1011513)

### References

- [1] Q. Abu Al-Haija, "A Stochastic Estimation Framework for Yearly Evolution of Worldwide Electricity Consumption," *Forecasting*, vol. 3, no. 2, 2021, pp. 256-266.
- [2] S. Kim and H. Lim, "Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings," *Energies*, vol. 11, no. 8, 2018, pp. 2010.
- [3] K. Moncef and A. Mohammad, "Review analysis of COVID-19 impact on electricity demand for residential buildings," *Renewable and Sustainable Energy Reviews*, vol. 143, 2021, pp. 110888.
- [4] A. Allouhi, Y. El Fouih, T. Kousksou, A. Jamil, Y. Zeraouli, and Y. Mourad, "Energy consumption and efficiency in buildings: current status and future trends," *J. of Cleaner production*, vol. 109, 2015, pp. 118-130.
- [5] H. Kang, J. Park, J. Park, and C. Kim, "A Study on Energy Management System for Energy Saving in Building," *J. of Institute of Korean Electrical and Electronics Engineers*, vol. 10, no. 7, 2015, pp. 787-792.
- [6] L. H. José, S. Roberto, C. Álvaro, P. Ricardo, and L. Isabel, "A Fuzzy-Based Building Energy Management System for Energy Efficiency," *Buildings*, vol. 8, no. 2, 2018, pp. 14-23.
- [7] K. Park, Y. Kim, S. Kim, K. Kim, W. Lee, and H. Park, "Building Energy Management System based on Smart Grid," In *Proceedings of the 2011 IEEE 33rd International Telecommunications Energy Conference*, Amsterdam, Netherlands, 2011, pp. 1-4.
- [8] T. Mazhar, M. A. Malik, I. Haq, I. Rozeela, I. Ullah, M. A. Khan, D. Adhikari, M. T. Ben Othman, and H. Hamam, "The Role of ML, AI and 5G Technology in Smart Energy and Smart Building Management," *Electronics*, vol. 11, no. 23, 2022, pp. 3960.
- [9] J. Kim and S. Kim, "Reinforcement Learning-Based Illuminance Control method for Building Lighting System," *J. of Institute of Korean Electrical and Electronics Engineers*, vol. 26, no. 1, 2022, pp. 56-61.
- [10] C. White and J. N. Gilmore, "Imagining the thoughtful home: Google Nest and logics of domestic recording," *Critical Studies in Media Communication*, vol. 40, no. 1, 2023, pp. 6-19.
- [11] S. Choi, S. Lee, and S. Kim, "Building-Integrated V2G Station Scheduling Algorithm based on Reinforcement Learning," In *Proceedings of Symposium of the Korean Institute of communications and Information Sciences*, Jeju, Korea, 2023, pp. 393-393.
- [12] H. Noh and Y. Lim, "Comparative Analysis of Battery Optimization in Grid Considering Consumption Patterns," *J. of the Korea Institute of Electronic Communication Science*, Vol. 30, no. 2, 2023, pp. 549-552.
- [13] S. Lee, "A routing algorithm based on deep reinforcement learning in SDN," *J. of the Korea Institute of Electronic Communication Science*, vol. 16, no. 6, 2021, pp. 1153-1160.
- [14] Z. Zhang, A. Chong, Y. Pan, C. Zhang, and K. P. Lam, "Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning," *Energy and Buildings*, vol. 199, 2019, pp. 472-490.
- [15] Y. Zhao, S. Cen, S. Seong, J. Hur, and C. Lim, "Path Planning with Obstacle Avoidance Based on Double Deep Q Networks," *J. of the Korea Institute of Electronic Communication Science*, vol. 18, no. 2, 2023, pp. 231-240.

## 저자 소개



### **최석민(Seok-Min Choi)**

2024년 동서대학교 소프트웨어학과 졸업(공학사)  
2024~현재 동서대학교 컴퓨터공학과 재학(석사)  
※ 관심분야 : 스마트그리드, 강화학습



### **김선용(Sun-Yong Kim)**

2014년 한동대학교 전산전자공학부 졸업(공학사)  
2019년 University of Maryland, Baltimore County 방문연구원  
2020년 GIST 전기전자컴퓨터공학부 졸업(공학박사)  
2020년 GIST AI연구소 연구원  
2020년~현재 동서대학교 소프트웨어학과 조교수  
※ 관심분야 : 스마트그리드, 강화학습

