

# 강화학습 기반 빌딩의 방별 조명 시스템 조도값 설정 기법

## Reinforcement Learning-Based Illuminance Control Method for Building Lighting System

김 종 민\*, 김 선 용\*\*★

Jongmin Kim\*, Sunyong Kim\*\*★

### Abstract

Various efforts have been made worldwide to respond to environmental problems such as climate change. Research on artificial intelligence (AI)-based energy management has been widely conducted as the most effective way to alleviate the climate change problem. In particular, buildings that account for more than 20% of the total energy delivered worldwide have been focused as a target for energy management using the building energy management system (BEMS). In this paper, we propose a multi-armed bandit (MAB)-based energy management algorithm that can efficiently decide the energy consumption level of the lighting system in each room of the building, while minimizing the discomfort levels of occupants of each room.

### 요 약

전 세계적으로 에너지 사용량이 증가함에 따라 지구온난화와 같은 환경문제가 초래되었으며, 이에 각국은 협정·협약을 통한 에너지 산업의 탈탄소화와 함께 화석 에너지를 신재생에너지로 빠르게 전환 중이다. 발전량이 급변하는 신재생에너지 보급 확대에 따라 효율적인 에너지 관리의 필요성이 대두되는 한편, AI 기술이 발전함에 따라 에너지 관리 분야와 결합한 AI 기반 빌딩 에너지 관리 시스템(Building Energy Management System, BEMS)의 연구 및 개발이 활발히 이루어지고 있다. 본 논문에서는 강화학습 기법 중 Multi-Armed Bandit(MAB) 알고리즘을 활용하여 빌딩 각 방의 조명시스템 전력사용량을 효율적으로 관리함과 동시에 사용자들의 불쾌지수를 최소화할 수 있는 알고리즘을 제안하고, 시뮬레이션을 통해 성능을 검증한다.

*Key words : Reinforcement learning, multi-armed bandit, building energy management, lighting system, illuminance control, user's discomfort*

### 1. 서론

전 세계적으로 에너지 사용량이 증가함에 따라 화석연료를 사용하는 발전소가 증가하였으며, 이는 다양한 환경문제를 초래하였다. 대표적인 환경문제로 온실가스로 인한 지구온난화가 있다. 연료가 연소함에 따라 대기 중에 이산화탄소(CO<sub>2</sub>)가 배출되는데 이는 온실효과가 뛰어나 지구온난화의 주요한 원인으로 지목된다. 이렇듯 점

경문제를 초래하였다. 대표적인 환경문제로 온실가스로 인한 지구온난화가 있다. 연료가 연소함에 따라 대기 중에 이산화탄소(CO<sub>2</sub>)가 배출되는데 이는 온실효과가 뛰어나 지구온난화의 주요한 원인으로 지목된다. 이렇듯 점

\* Undergraduate Student, Dept. of Software, Dongseo University

\*\* Assistant Professor, Dept. of Software, Dongseo University

★ Corresponding author

E-mail : sykim@dongseo.ac.kr, Tel : +82-51-320-4224

※ Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022R1G1A1011513)

Manuscript received Dec. 27, 2021; revised Mar. 21, 2022; accepted Mar. 21, 2022.

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.

점 심각해지는 환경문제에 대해 각국은 에너지 산업의 탈탄소화를 위해 유엔기후변화협약, 교토의정서, 파리기후협정 등의 협정·협약을 통해 노력 중이며, 화석 에너지를 신재생에너지로 빠르게 전환 중이다. 또한, 발전량이 급변하는 신재생에너지 보급 확대에 따라 에너지 관리의 필요성이 급격히 증가하고 있다. 에너지의 효율적 관리는 ‘제5의 연료(The Fifth Fuel)’로 불릴 만큼 상당히 높은 경제성을 가지며[1], 2016년 미래창조과학부 보고에 따르면 효율적인 에너지 관리가 온실가스 감축 효과가 가장 높은 것으로 보고되고 있다. 이에 따라, 전 세계적으로 에너지 사용량의 약 20%를 차지하는 빌딩 에너지 사용량 절감에 대한 필요성이 증대되고 있으며[2], 주거용 빌딩, 상업용 빌딩 등 다양한 종류의 빌딩 에너지 관리 최적화를 통한 빌딩 에너지 관리시스템(BEMS: Building Energy Management System) 개발이 활발히 이루어지고 있다[3].

한편, 인공지능(AI: Artificial Intelligence) 기술이 발전함에 따라 AI가 여러 분야와 접목되고 있으며, AI 기술이 접목된 BEMS 기술 분야가 크게 주목받고 있다. 대표적인 예로 IT 기업 구글은 딥마인드(DeepMind) 신경망 AI를 활용하여 데이터 센터를 냉각하는 데 사용되는 전력을 약 40%가량 줄이는 데 성공하였다[4]. 또한, 국내 기업 KT는 AI기반 중대형 빌딩 BEMS인 기가 에너지 매니저 빌딩(GIGA Energy Manager Building) 기술을 개발하여, 에너지를 평균 10% 절감하였다[5].

AI기반 빌딩 에너지 관리를 위한 다양한 연구도 활발히 진행되고 있다. 관련 연구로써, 논문 [6]에서는 빌딩에 효율적인 에너지 관리를 위해 기계학습(Machine Learning) 기반의 열수요 단기(Short-term) 예측 기법을 제안하며, 네 가지 종류의 빌딩을 대상으로 하는 사례 연구를 통해 예측성능을 검증하였다. 또한, 논문[7]에서는 HVAC(Heating, Ventilation and Air Conditioning) 시스템의 냉방기 전력사용량 모델을 제안하였다. 인공지능 경망 모델을 활용한 전력사용량 예측수행 결과, 99.1%의 정확도를 제공하였다.

본 논문에서는 기계학습 기법의 하나인 강화학습을 이용하여 빌딩의 방별 조명시스템 조도값 설정 알고리즘을 제안한다. 제안한 알고리즘을 통해 빌딩 조명시스템의 전력사용량을 효율적으로 관리함과 동시에, 조명제어에 따른 사용자들의 불편지수를 최소화할 수 있는 전략을 도출하고, 시뮬레이션을 통해 성능을 검증한다.

## II. 빌딩의 방별 조도값 설정 기법 알고리즘

### 1. 시스템 환경

본 논문에서는 기계학습 기법을 활용한 중앙제어시스템을 통해 건물 각 방의 전구의 조도 설정값을 조절하여 빌딩 전체의 전력량을 효과적으로 줄이는 동시에 각 방 사용자들의 불편함(불쾌지수)을 최소화하는 값을 찾는 방법을 기재한다.

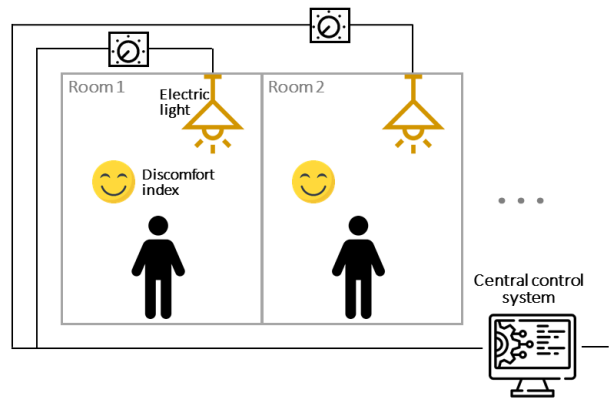


Fig. 1. Building lighting control system model.

그림 1. 빌딩 조명제어 시스템 모델

그림 1은 본 논문에서 고려하는 빌딩의 조명제어 시스템을 보여준다. 빌딩의 각 방에는 조도의 개별 설정이 가능한 조명장치가 설치되어 있으며, 이 때 조명장치의 전력사용량이 빌딩 전체 전력사용량의 높은 비중을 차지함을 가정한다.

방  $n \in 1, \dots, N$ 의 면적을  $S_n [\text{m}^2]$ , 조명장치 설정 조도값을  $I_n [\text{lux}]$ , 각 방의 전구에 발광효율을  $\eta [\text{lm/W}]$ , 전구사용 시간을  $h [\text{hour}]$ 라고 할 때, 방  $n$ 의 전력소비량  $E_n [\text{kWh}]$ 는 다음과 같이 계산된다[8].

$$E_n = \frac{I_n \times S_n \times h}{10^3 \times \eta} \quad [\text{kWh}] \quad (1)$$

예를 들어, 면적이  $100 \text{m}^2$ 인 방에서 발광효율이  $\eta = 15$ 인 백열등을 사용하여  $500 \text{lux}$ 의 조도를 설정한 조명장치를 1시간 동안 사용한다고 가정을 하면

$$E_n = \frac{500 \times 100 \times 1}{1000 \times 15} = 3.33 \quad [\text{kWh}] \quad (2)$$

즉, 1시간당  $3.33 \text{kW}$ 의 전력량을 소비하게 된다.

각 방의 사용자들은 개개인에게 가장 선호하는 조도값  $I_n^* [\text{lux}]$ 을 갖는다. 이 때, 각 방의 사용자들은 본인의 선

호 조도값  $I_n^*$ 와 조명장치로부터 설정된 실제 조도값  $I_n$ 의 차이만큼 비례하게 불편함을 겪을 것이다. 따라서  $\beta$ ,  $\gamma$ 를 임의의 상수라고 할 때, 각 방  $n$ 의 사용자들이 겪는 불쾌지수  $D_n$ 는 다음과 같이 모델링 할 수 있다.

$$D_n = \beta(I_n - I_n^*)^\gamma \quad (3)$$

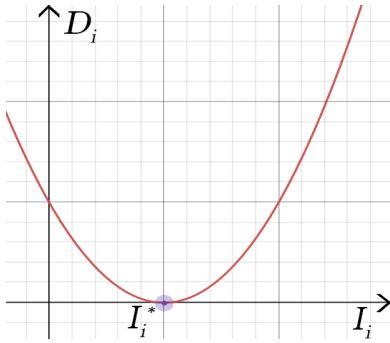


Fig. 2.  $D_n$  with respect to  $I_n$  (an example).  
그림 2.  $I_n$  값에 따른  $D_n$ 의 변화 그래프 (예시)

예를 들어,  $\beta=1$ ,  $\gamma=2$ 인 경우 그림 2와 같은 2차함수 형태의 불쾌지수 모델을 갖는다. 즉,  $I_n$ 가  $I_n^*$ 와 멀어질수록 더 큰 불쾌지수  $D_n$ 를 갖게 된다.

본 논문의 핵심 아이디어는 각 방 조명장치의 조도 설정값  $I_n$ 를 조절하여 빌딩 전체의 전력사용량인  $\sum_{n=1}^N E_n$ 을 최소화 하는 동시에 빌딩 전체의 불쾌지수인  $\sum_{n=1}^N D_n$ 를 최소화 하는 것이므로, 다음 수식으로 정의되는  $R$ (리워드)의 최소값을 찾는 문제로 귀결된다.

$$R = \alpha \sum_{n=1}^N E_n + (1-\alpha) \sum_{n=1}^N D_n \quad (4)$$

이때,  $0 \leq \alpha \leq 1$ 는 전력소비량과 불쾌지수의 가중치를 조절하는 매개변수다. 수식 (4)에서 주목할 점은,  $I_n^*$ 는 각 방 사용자들의 기분이나 그날 날씨, 눈의 피로도 등에 크게 영향을 받으며, 심지어 사용자 스스로도 본인이 선호하는 조도값을 정확히 수치화할 수 없다는 불확실성(Uncertainty)의 특징을 갖는다. 이러한 불확실성이 존재하는 상황 하에서도 학습을 통해 각 방의 최적의 전력량 사용 솔루션을 찾을 수 있도록, 본 논문에서는 강화학습(Reinforcement learning) 기반의 방별 조도값 설정 기법을 제안한다. 특히, 강화학습 기법 중 Multi-Armed Bandit(MAB) 알고리즘을 활용한다.

## 2. Multi-Armed Bandit 기반 방별 조도값 설정 알고리즘 제안

MAB 알고리즘은 리워드(Reward)의 분포를 알 수 없는 다수의 선택지에서 탐색(Exploration)과 활용(Exploitation)을 반복하며 시행착오를 통해 기계가 스스로 학습한 뒤, 어떠한 행동(Action)이 최적의 보상을 제공하는지 찾아낼 수 있는 강화학습 알고리즘이다[9]. 본 논문에서는 각 방 사용자들의 선호 조도값에 대한 불확실성 하에서 각 방의 전력사용량을 최소화 하면서, 동시에 사용자들의 조도에 대한 불편함을 최소화 할 수 있는 방별 최적의 조도값 설정에 대한 전략을 찾기 위하여 MAB 알고리즘을 적용한다.

하나의 액션은 MAB 알고리즘에서의 하나에 arm  $k$ 에 매핑(mapping)되며, 각 arm  $k$ 는 방  $n \in 1, \dots, N$ 의 조명 시스템의 최대 조도값  $I_n^{\max}$  대비 실제 설정 조도값  $I_n^{\text{set}}$ 의 비율인  $a_n$ 의 조합  $\mathbf{a}_k = (a_1, \dots, a_N)$ 에 매핑된다. 따라서,

$$I_n^{\text{set}} = a_n \times I_n^{\max} \quad (5)$$

의 관계가 성립되며, 각  $a_n$ 이  $M$ 개의 discrete 한 레벨을 가질 때, 총 선택 가능한 액션(arm)  $\mathbf{a}_k$ 의 가짓수는  $K=M^N$ 이 된다.

Table 1. Action examples.

표 1. 액션 예시

Action	Room 1	Room 2	Room 3
$\mathbf{a}_1$	100%	100%	100%
$\mathbf{a}_2$	100%	100%	75%
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{a}_k$	50%	25%	0%
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{a}_{125}$	0%	0%	0%

표 1은  $M=5$ ,  $N=3$ 일 때 가질 수 있는 총  $K=125$ 개의 액션을 보여준다.  $\bar{R}(\mathbf{a}_k)$ ,  $c$ ,  $m(\mathbf{a}_k)$ 를 각각 현재까지 액션  $\mathbf{a}_k$ 를 수행했을 때의 평균 리워드( $R$ ), 상수, 현재까지 액션  $\mathbf{a}_k$ 를 수행한 횟수라고 정의하면, MAB 알고리즘에 따라 매 타임스텝  $t$ 에  $K$ 개 중 1개의 최적 액션  $A_t$ 가 다음 수식에 의해 선택된다.

$$A_t \approx \arg \min [\bar{R}(\mathbf{a}_k) - c \sqrt{\frac{\log t}{m(\mathbf{a}_k)}}] \quad (6)$$

결과적으로, 제안하는 빌딩 조명제어 시스템은 매 시간  $t$ 에 각 방  $n$ 의 사용자 선호 조도값  $I_n^*$ 의 불확실성 하에서 수식 (6)에 의하여 현재 상황에서의 최적 액션  $A_t$ 를 선택하여 수행하고, 충분히 많은 수의 시도 끝에 점차적으로 수식 (4)의  $R$  값을 낮추게 된다. 이를 의사코드로 표현하면 다음의 그림 3과 같이 표현할 수 있다.

---

**Algorithm 1** MAB-based Proposed Algorithm
 

---

```

1: // Initialization
2:  $N \leftarrow \#$  of Room,  $M \leftarrow \#$  of Levels
3: for  $k = 1$  to  $K$  do
4:   Play action  $a_k$ 
5:   Update  $\bar{R}(a_k) \leftarrow R$ 
6:    $m(a_k) \leftarrow m(a_k) + 1$ 
7: end for
8: // Main Loop
9: for  $t = 1$  to  $\#$  of Iterations do
10:  Play an action  $a_k$  solving the following problem
11:   $\min_t (\bar{R}(a_k) - c\sqrt{\frac{\log t}{m(a_k)}})$ 
12:   $\bar{R}(a_k) \leftarrow \frac{\bar{R}(a_k) \cdot m(a_k) + R}{m(a_k) + 1}$ 
13:   $m(a_k) \leftarrow m(a_k) + 1$ 
14: end for
  
```

---

Fig. 3. Pseudocode of the proposed MAB-based illuminance control algorithm for each room.

그림 3. 제안하는 MAB 기반 방별 조도값 설정 알고리즘의 의사코드

이때,  $\bar{E}_n$ ,  $\bar{D}_n$ 은 각각 정규화(normalization) 된  $E_n$ ,  $D_n$ 의 표현이다.

### III. 성능평가

#### 1. 시뮬레이션 세팅

제안하는 방별 조도값 설정 알고리즘의 성능평가를 위해 다음과 같이 시뮬레이션을 세팅한다. 빌딩 안에는 총 3개의 방이 있다고 가정하며( $N=3$ ), 모든 방은 100 m<sup>2</sup>의 면적을 가진다고 가정한다( $S_n = 100, \forall n \in \{1, \dots, N\}$ ). 또한, 각 방은 모두 백열등을 사용한다고 가정하여 발광 효율을  $\eta = 15$  lm/W로 세팅한다[10]. 조명 시스템의 최대 조도값은  $I_n^{\max} = 1000, \forall n \in \{1, \dots, N\}$ 으로 설정하며, 이는  $M=20$ 개의 discrete 한 레벨(5%, 10%, ..., 100%)로 나뉜다. 따라서, 총 선택 가능한 액션  $a_k$ 의 가짓수는  $K=20^3$ 가 된다. 이는 곧 MAB 알고리즘의 arm의 개수이다.

각 방 사용자들이 선호하는 조도값에 대한 불확실성을 받

영하기 위해, 선호 조도값  $I_n^*$ 는 가우시안 분포(Gaussian distribution)를 따르도록 다음과 같이 모델링 한다.

$$\begin{aligned} I_1^* &\sim N(500, 10^2), \\ I_2^* &\sim N(700, 10^2), \\ I_3^* &\sim N(900, 10^2) \end{aligned} \quad (7)$$

마지막으로, 수식 (3)의 불쾌지수  $D_n$ 를 모델링하는 상수는  $\beta = 1$ ,  $\gamma = 2$ 로, 수식 (4)의 가중치는  $\alpha = 0.5$ 로, 수식 (6)의 상수는  $c = \sqrt{2}$ 로 설정한다.

#### 2. 성능평가 결과

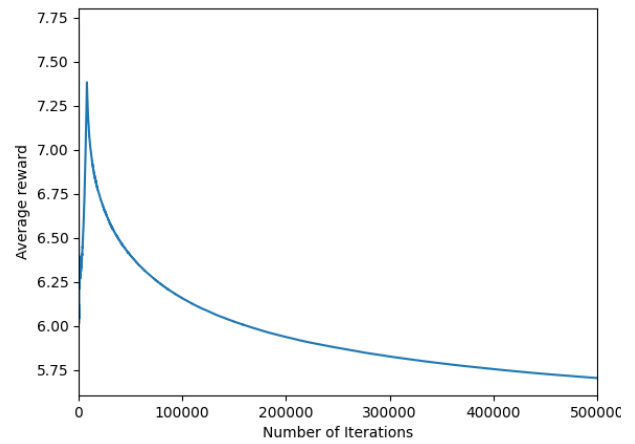


Fig. 4. Average reward w.r.t. number of action attempts.  
그림 4. 액션 시도에 따른 평균 리워드의 변화

그림 4는 액션을 시도하는 횟수에 따른 리워드( $R$ ) 평균값의 변화 그래프를 보여주며, 이는 제안하는 MAB 기반 방별 조도값 설정 알고리즘의 학습 추이를 나타낸다. MAB 기반 알고리즘의 특성상 초기에는 임의로 액션을 시도함에 따라  $R$  값이 매우 높지만, 일정 기간의 탐색 과정 이후에는 활용을 통해 점차적으로 감소하는 경향을 보여준다. 따라서, 학습을 진행함에 따라 전력사용량과 불쾌지수를 동시에 낮출 수 있는 방향으로 개선된 액션을 도출할 수 있음을 알 수 있다.

그림 5는  $5 \times 10^5$ 번의 액션 시도 후 출력된 시뮬레이션 결과 화면을 나타낸다. 강조된 부분에서 볼 수 있듯이,  $\alpha = 0.5$ 인 상황에서 930번째 액션이 7,162회로 가장 많이 선택(시도)되었으며, 이때의 각 방 1, 2, 3의 최적 설정 조도값  $I_n^{set} (n=1, 2, 3)$ 의 설정은 각각 15%(=150 lux), 35%(=350 lux), 55%(=550 lux)로 나타난다.  $\alpha = 0.5$ 로 설정한 결과 전력사용량과 불쾌지수의 가중치가 0.5로 같은 상황이기 때문에, 각 방의 사용자 선호 조도값  $I_n^*$ 의 평균값인 500 lux, 700 lux, 900 lux의 절반

정도로 최적 설정 조도값이 도출된 것으로 예상된다. 그림 6은 각 액션 별 선택된 누적 횟수를 그래프로 보여준다. 그림 5의 결과와 일치하게 930번째 액션이 가장 많이 선택된 모습을 보여준다.

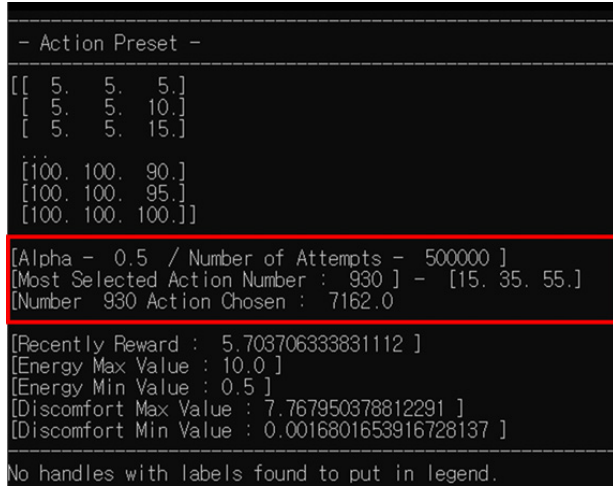


Fig. 5. Simulation results for the proposed MAB-based algorithm.

그림 5. 제안한 MAB 기반 알고리즘의 시뮬레이션 결과

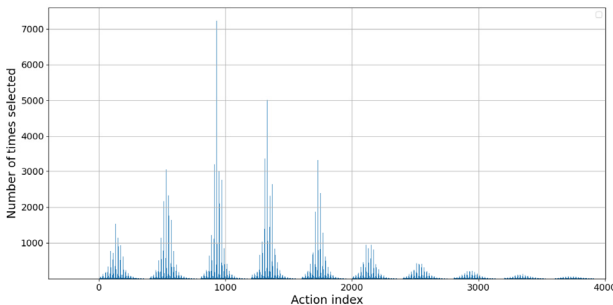


Fig. 6. Number of times each action is selected.

그림 6. 각 액션 별 선택된 누적 횟수

마지막으로, 표 2는 전력소비량과 불쾌지수 사이의 가중치를 조절하는  $\alpha$ 에 따른 최적 설정 조도값  $I_n^{set}$  ( $n=1, 2, 3$ ) 도출 결과를 보여준다. 먼저  $\alpha=0.5$ 인 경우, 그림 5의 결과와 일치하게 최적 설정 조도값  $I_n^{set}$  ( $n=1, 2, 3$ )가 각각 150 lux, 350 lux, 550 lux로 나타난다.  $\alpha=0$ 일 때는 식 (4)에서 불쾌지수 term만 남게 되며, 이는 전력소비량은 고려하지 않은 채 오로지 불쾌지수를 낮추는 것만 고려한다는 의미이기 때문에, 학습결과 각 방의 사용자 선호 조도값의 평균값인 500 lux, 700 lux, 900 lux이 최적 설정 조도값으로 도출됨을 알 수 있다. 이와 반대로  $\alpha=1$ 인 경우, 전력소비량을 낮추는 것만 고려한다는 의미이며, 학습결과 조명시스템의 설정 가능한 조도

값 중 최소값인 50 lux가 각 방의 최적 설정 조도값으로 도출된 결과를 보여준다.

Table 2. The results of learning according to  $\alpha$ .

표 2.  $\alpha$  값에 따른 학습 결과

$\alpha$	$I_1^{set}$ [lux]	$I_2^{set}$ [lux]	$I_3^{set}$ [lux]
0.0	500	700	900
0.1	450	650	850
0.2	400	600	800
0.3	350	550	750
0.4	300	500	700
0.5	150	350	550
0.6	50	200	400
0.7	50	50	100
0.8	50	50	50
0.9	50	50	50
1.0	50	50	50

#### IV. 결론 및 향후 연구

본 논문에서는 강화학습 기반의 빌딩 방별 조명제어 알고리즘을 통해 각 방의 조도값을 조절하여 전력사용량을 낮추는 동시에 이로 인한 각 방 사용자들의 불편함을 최소화 할 수 있는 전력사용량 제어 전략을 도출하였다. 향후 연구로써, 현재 고려하고 있는 조명시스템 외에 냉·난방기, PC 등 전력사용량의 많은 부분을 차지하고 있는 다양한 부하의 전력사용을 제어할 수 있는 알고리즘으로 연구를 확장할 예정이다. 즉, 멀티룸-멀티부하의 전력사용량 제어 알고리즘으로 개선할 계획이다. 또한, 시뮬레이션을 통해 생성된 데이터가 아닌, 실제 측정된 전력사용량 데이터를 활용하여 성능평가를 진행할 계획이다.

#### References

- [1] The Economist Special Report, "Invisible Fuel", 2015. (<https://www.economist.com/special-report/2015/01/15/invisible-fuel>)
- [2] S. Kim and H. Lim, "Reinforcement learning based energy management algorithm for smart energy buildings," *Energies*, vol.11, no.8, 2018. DOI: 10.3390/en11082010
- [3] S. Park et al. "Reinforcement Learning-Based



BEMS Architecture for Energy Usage Optimization,” *Sensors*, vol.20, no.17, 2020.

DOI: 10.3390/s20174918

[4] DeepMind Blog Post, “DeepMind AI Reduces Google Data Centre Cooling Bill by 40%”, 2016. (<https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40>)

[5] Business Korea News, “KT Launches AI-based Energy Management Service for Buildings,” 2019. (<http://www.businesskorea.co.kr/news/articleView.html?idxno=37865>)

[6] A. T. Eseye and M. Lehtonen, “Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models,” *IEEE Transactions on Industrial Informatics*, vol.16, no.12, 2020. DOI: 10.1109/TII.2020.2970165

[7] J. H. Kim, N. C. Seong and W. Choi, “Modeling and optimizing a chiller system using a machine learning algorithm,” *Energies*, vol.12, no.15, 2019. DOI: 10.3390/en12152860

[8] “How to convert lux to watts”, <https://www.rapidtables.com/calc/light/index.html>

[9] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, MIT press, 2018.

[10] T. Murphy, “Maximum Efficiency of White Light.” *Department of physics UC San Diego*, 2011.

#### Sunyong Kim (Member)



2014 : BS degree in Electrical Engineering, Handong Global University.

2020 : PhD degree in Electrical Engineering and Computer Science, GIST.

2020 : Researcher, GIST Institute for Artificial Intelligence.

2020~Present : Assistant Professor, Dept. of Software, Dongseo University.

## BIOGRAPHY

#### Jongmin Kim (Member)



2017~Present : Undergraduate Student, Dept. of Software, Dongseo University.