Reinforcement Learning-Based Energy Management of Smart Home with Rooftop Solar Photovoltaic System, Energy Storage System, and Home Appliances

Taeseop Park, Bumsuk Lee, Haejoong Lee

Sogang University, EE

2021.06.10

Networking for Intelligence Communications and Energy

Reinforcement Learning-Based Energy Management of Smart Home with Rooftop Solar Photovoltaic System, Energy Storage System, and Home Appliances

☐ Contents

sensors

MDPI

Article

Reinforcement Learning-Based Energy Management of Smart Home with Rooftop Solar Photovoltaic System, Energy Storage System, and Home Appliances

Sangyoon Lee and Dae-Hyun Choi *

School of Electrical and Electronics Engineering, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 156-756, Korea; sangyoon1207@naver.com

* Correspondence: dhchoi@cau.ac.kr; Tel.: +82-2-820-5101

check for updates

Received: 13 August 2019; Accepted: 9 September 2019; Published: 12 September 2019

Abstract: This paper presents a data-driven approach that leverages reinforcement learning to manage the optimal energy consumption of a smart home with a rooftop solar photovoltaic system, energy storage system, and smart home appliances. Compared to existing model-based optimization methods for home energy management systems, the novelty of the proposed approach is as follows: (1) a model-free Q-learning method is applied to energy consumption scheduling for an individual controllable home appliance (air conditioner or washing machine), as well as the energy storage system charging and discharging, and (2) the prediction of the indoor temperature using an artificial neural network assists the proposed Q-learning algorithm in learning the relationship between the indoor temperature and energy consumption of the air conditioner accurately. The proposed Q-learning home energy management algorithm, integrated with the artificial neural network model, reduces the consumer electricity bill within the preferred comfort level (such as the indoor temperature) and the appliance operation characteristics. The simulations illustrate a single home with a solar photovoltaic system, an air conditioner, a washing machine, and an energy storage system with the time-of-use pricing. The results show that the relative electricity bill reduction of the proposed algorithm over the existing optimization approach is 14%.

Keywords: home energy management system; reinforcement learning; artificial neural network; smart home; consumer comfort; smart grid

Reinforcement Learning-Based Energy Management of Smart Home with Rooftop Solar Photovoltaic System, Energy Storage System, and Home Appliances

☐ Background

- TOU price : time-of-use price (Utility Company)
- Appliance Energy Consumption

- Outdoor Temperature (Weather Station)
- PV systems (solar photovoltaic)

- Comfort setting
- Appliance Parameters

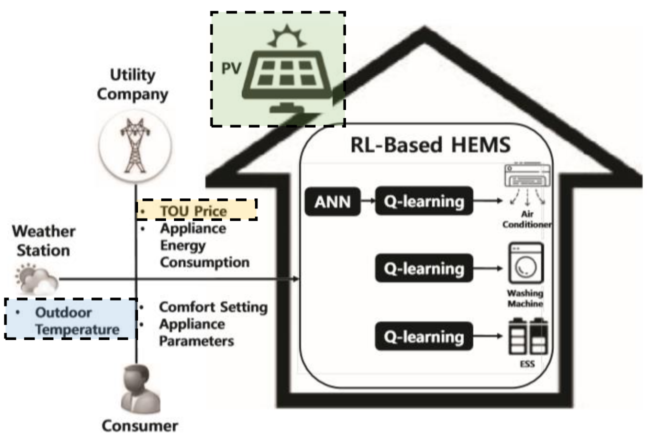
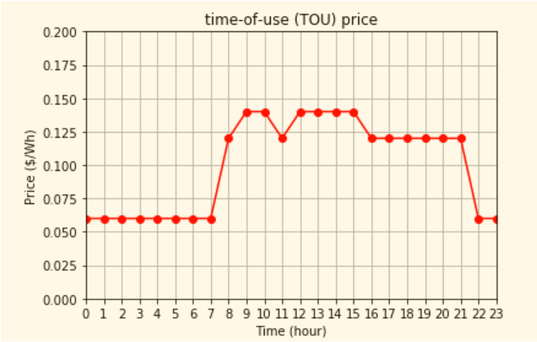
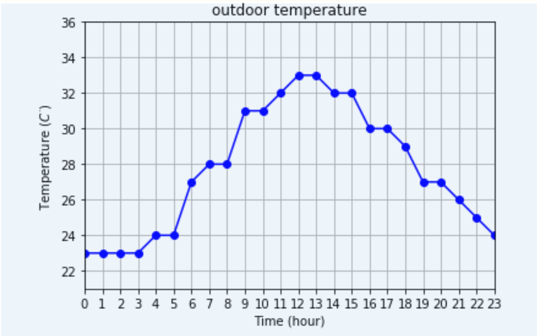
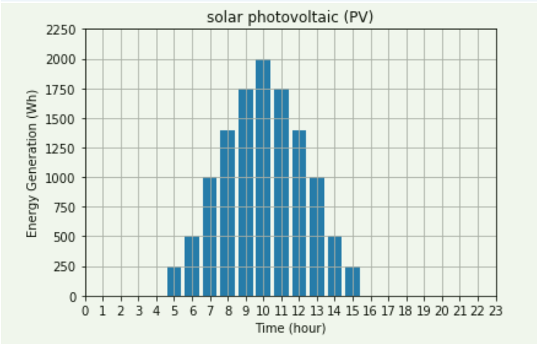


Figure 1. Illustration of the proposed home energy management system (HEMS) framework.







Introduction

☐ **Compared to the existing model-based HEMS optimization approaches, we propose a HEMS algorithm using a model-free reinforcement learning (RL).**

- The Q-learning method is applied to the energy consumption scheduling of different home appliances (**air conditioner**, **washing machine**, and **ESS**), whereby the agent of each appliance determines the optimal policy independently to **reduce its own electric cost within the consumer comfort level** and the appliance operation characteristics.
- An **ANN model** to learn the relationship between the indoor temperature and energy consumption of the air conditioner more accurately, which is integrated into the **Q-learning module to achieve improved performance of the air conditioner agent**.

Simulation showed two results.

1. Using reinforcement learning and ANN prediction techniques, power consumption and user dissatisfaction were reduced.

2. It can save more energy than the existing mixed-integer linear programming algorithm.

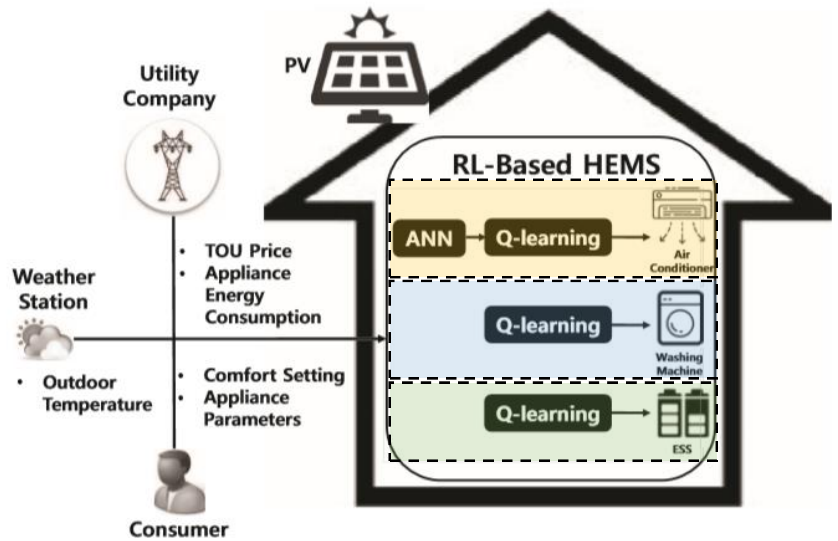


Figure 1. Illustration of the proposed home energy management system (HEMS) framework.

Related Research

☐ model-based HEMS optimization

- the scheduling of different types of home appliances along with electric vehicles using linear programming (LP)
- load scheduling considering the consumer comfort level using mixed integer nonlinear programming (MINLP)
- convex programming based on relaxed MINLP using an L1 regularization technique
- load scheduling for a single consumer or multiple consumers using MILP
- LP-based joint optimization of energy supplies and electric loads through three-stage scheduling (prediction, supply control, and demand control)
- the natural aggregation algorithm (NAA)-based HEMS method consisting of forecasting, day-ahead scheduling, and actual operation
- robust optimization for the scheduling of home appliances to resolve the uncertainty of consumer behavior the outdoor temperature and consumer comfort levels and distributed HEMS architectures consisting of a local and global HEMS.
- using real-time pricing, a HEMS optimization method that considers the operational dependency of various types of home appliances and consumer life style requirements was proposed in

☐ model-free reinforcement learning (RL)

System Model for HEMS

□ Preliminary

$\boldsymbol{A^c}$: A controllable appliance is an appliance of which the operation is scheduled and controlled by the HEMS.

- A_r^c: a **reducible** appliances (e.g. air conditioner)
- A_s^c: a **shiftable** appliances
  - $A_s^{c,NI}$: a **non-interruptible** load (e.g. : washing machine)
  - $A_s^{c,I}$: an **interruptible** load (e.g. : ESS)

$\boldsymbol{A^{uc}}$: An uncontrollable appliance cannot be scheduled and operated by the HEMS. (e.g. : TV, PC or lighting)

Therefore, $\boldsymbol{A^{uc}}$ maintains the fixed energy consumption scheduling.

System Model for HEMS

☐ Conventional HEMS Optimization Formulation

Object Function

$$\min_{E_t^{net}, T_t^{in}} \underbrace{\sum_{t \in T} \pi_t E_t^{net}}_{J_1(E_t^{net})} + \underbrace{\epsilon \sum_{t \in T} \left| T_t^{in} - T_t^{set} \right|}_{J_2(T_t^{in})}.$$

Total electricity cost

Total penalty
(consumer discomfort cost)

- π_t : the TOU price
- E_t^{net} : net energy consumption
- T_t^{set} : the preferred consumer temperature
- T_t^{in} : the indoor temperature
- ϵ : Penalty for consumer thermal discomfort cost.

System Model for HEMS

☐ Conventional HEMS Optimization Formulation

Constraint 1 (Net Power Consumption)

the predicted PV generation output

$$E_t^{\text{net}} = \sum_{a \in \mathcal{A}} E_{a,t} - \widehat{E}_t^{\text{PV}}, \tag{2}$$

$$\sum_{a \in \mathcal{A}} E_{a,t} = \sum_{a \in \mathcal{A}_r^c} E_{a,t} + \sum_{a \in \mathcal{A}_s^{c,NI}} E_{a,t} + \sum_{a \in \mathcal{A}_s^{c,I}} \left(E_{a,t}^{\text{ch}} - E_{a,t}^{\text{dch}} \right) + \sum_{a \in \mathcal{A}^{uc}} E_{a,t}. \tag{3}$$

- A_r^c: a **reducible** appliances (e.g. AC)
- A_s^c: a **shiftable** appliances
- $A_s^{c,NI}$: a **non-interruptible** load (e.g. : WM)
- $A_s^{c,I}$: an **interruptible** load (e.g. : ESS)
- A^{uc} : An **uncontrollable appliance** cannot be scheduled and operated by the HEMS.

- E_t^{net}: Net energy consumption at time slot t
- $E_{a,t}^{ch}$: Charging energy of ESS a at time slot t
- $E_{a,t}^{dch}$: Discharging energy of ESS a at time slot t

System Model for HEMS

☐ Conventional HEMS Optimization Formulation

Constraint 2 (Operating Characteristics for Controllable Appliances)

e.g. air conditioner

Constraint for the temperature dynamics of the reducible appliance

$$T_t^{\text{in}} = T_{t-1}^{\text{in}} + \alpha(\widehat{T}_{t-1}^{\text{out}} - T_{t-1}^{\text{in}}) + \beta E_{a,t}, \quad (4)$$

$$T^{\text{min}} \leq T_t^{\text{in}} \leq T^{\text{max}}, \quad (5)$$

Consumer preferred indoor temperatures range

$$E_a^{\text{min}} \leq E_{a,t} \leq E_a^{\text{max}}. \quad (6)$$

Capacity range

- T_t^{in}: in-door temperature at time t
- T_{t-1}^{in}: in-door temperature at time $t-1$
- $\hat{T}_{t-1}^{out}$: the predicted outdoor temperature at time $t-1$
- $E_{a,t}$: the energy consumption of the reducible appliances
- α , β : the environmental parameters

NICELAB
Networking for Intelligence Communications and Energy

서강대학교
SOGANG UNIVERSITY

System Model for HEMS

☐ Conventional HEMS Optimization Formulation

Constraint 3 **(the desired operation of shiftable appliances with a non-interruptible load)**

e.g. washing machine

$$b_{a,t}^{c,NI} = 0, \qquad t \in [1, \omega_s^{\text{pref}}) \cup (\omega_f^{\text{pref}}, T], \tag{7}$$

$$\sum_{t=\omega_s^{\text{pref}}}^{\omega_f^{\text{pref}}} b_{a,t}^{c,NI} = L_a, \tag{8}$$

$$\sum_{t=p}^{p+L_a-1} b_{a,t}^{c,NI} \geq (b_p^{c,NI} - b_{p-1}^{c,NI}) L_a, \quad \forall p \in (\omega_s^{\text{pref}}, \omega_f^{\text{pref}} - L_a + 1) \tag{9}$$

$$E_{a,t} = b_{a,t}^{c,NI} E_a^{\max}. \tag{10}$$

- $b_{a,t}^{c,NI}$: the binary decision variable , “1” for consumption, “0” otherwise
- L_a: the operation period of L_a hours during a day
- E_a^{max}: Maximum consumption of appliance a

System Model for HEMS

☐ Conventional HEMS Optimization Formulation

Constraint 4 (the operational dynamics of the state of energy (SOE) for the ESS)

e.g. ESS

$$SOE_{a,t} = SOE_{a,t-1} + \eta_a^{\mathrm{ch}} E_{a,t}^{\mathrm{ch}} - \frac{E_{a,t}^{\mathrm{dch}}}{\eta_a^{\mathrm{dch}}}, \tag{11}$$

$$SOE_a^{\min} \leq SOE_{a,t} \leq SOE_a^{\max}, \tag{12}$$

$$E_a^{\mathrm{ch,min}} b_{a,t}^{c,I} \leq E_{a,t}^{\mathrm{ch}} \leq E_a^{\mathrm{ch,max}} b_{a,t}^{c,I}, \tag{13}$$

$$E_a^{\mathrm{dch,min}}(1 - b_{a,t}^{c,I}) \leq E_{a,t}^{\mathrm{dch}} \leq E_a^{\mathrm{dch,max}}(1 - b_{a,t}^{c,I}). \tag{14}$$

- $\eta_a^{ch}, \eta_a^{dch}$: the charging and discharging efficiency
- $E_{a,t}^{ch}, E_{a,t}^{dch}$: the charging and discharging energy
- SOE_a^{min}, SOE_a^{max}: the SOE capacity constraint with SOE_a^{min} and SOE_a^{max}
- $b_{a,t}^{c,I}$: the binary decision variable that determines the ESS on/off status "1" for charging, "0" otherwise
- $E_a^{dch,max}, E_a^{dch,min}$: Maximum (Minimum) discharging energy of ESS a

System Model for HEMS

☐ Conventional HEMS Optimization Formulation

MINLP -> MILP by means of the linearization of the nonlinear o.f.

Non-linear

$$\min_{E_t^{net}, T_t^{in}} \sum_{t \in T} \pi_t E_t^{net} + \epsilon \sum_{t \in T} \left| T_t^{in} - T_t^{set} \right|.$$

$$\Delta T_t = |T_t^{\text{in}} - T^{\text{set}}|, \tag{15}$$

$$\Delta T_t \geq T_t^{\text{in}} - T^{\text{set}}, \tag{16}$$

$$\Delta T_t \geq T^{\text{set}} - T_t^{\text{in}}. \tag{17}$$

Linear

- MINLP (mixed-integer non-linear programming)
- MILP (mixed-integer linear programming)

Formulation of RL- and ANN-Based Home Energy Management

☐ Home Energy Management via Q-Learning

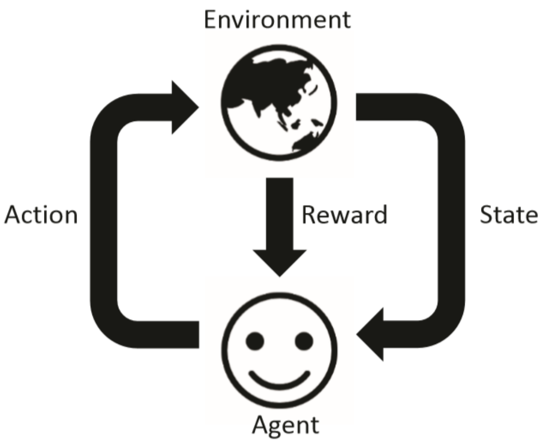


Figure 2. Conceptual architecture of reinforcement learning (RL).

s_t : State at time slot t

a_t : Action at time slot t

$r(s_t, a_t)$: Reward at time slot t

$Q(s_t, a_t)$: Function that tells how good an action is in which s_t

Bellman equation

$$Q^*_{v^*}(s_t, a_t) = r(s_t, a_t) + \gamma \max Q(s_{t+1}, a_{t+1}).$$

$\gamma \in [0, 1]$

discount rate = 0.9

$$\underbrace{Q(s_t, a_t) \leftarrow}_{update} (1-\theta)Q(s_t, a_t) + \theta[r(s_t, a_t) + \gamma \max Q(s_{t+1}, a_{t+1})].$$

$\theta \in [0, 1]$

learning rate = 0.1

$$v^* = \arg\max Q(s_t, a_t).$$

Formulation of RL- and ANN-Based Home Energy Management

☐ Home Energy Management via Q-Learning

State Space

$$\underbrace{\mathcal{S}^{\mathrm{WM}}}_{set} = \{E_t^{\mathrm{WM}}\}, \ \mathcal{S}^{\mathrm{AC}} = \{E_t^{\mathrm{AC}}\}, \ \mathcal{S}^{\mathrm{ESS}} = \{SOE_t^{\mathrm{ESS}}\}, \qquad \forall t = 1, \ldots, 24,$$

Action

$$\mathcal{A}^{\mathrm{WM}} = \{\mathrm{On}, \mathrm{Off}\},$$
$$\mathcal{A}^{\mathrm{AC}} = \{0, \Delta E^{\mathrm{AC}}, 2\Delta E^{\mathrm{AC}}, \ldots, 8\Delta E^{\mathrm{AC}}, 9\Delta E^{\mathrm{AC}}\},$$
$$\mathcal{A}^{\mathrm{ESS}} = \{-4\Delta E^{\mathrm{ESS}}, -3\Delta E^{\mathrm{ESS}}, -2\Delta E^{\mathrm{ESS}}, -1\Delta E^{\mathrm{ESS}}, 0, 1\Delta E^{\mathrm{ESS}}, 2\Delta E^{\mathrm{ESS}}, 3\Delta E^{\mathrm{ESS}}, 4\Delta E^{\mathrm{ESS}}\}.$$

Reward

$$r^{\mathrm{Total}} = r_t^{\mathrm{WM}} + r_t^{\mathrm{AC}} + r_t^{\mathrm{ESS}}.$$

- $\Delta E^{AC} = 40\mathrm{Wh}$, $\Delta E^{ESS} = 150\mathrm{Wh}$

Formulation of RL- and ANN-Based Home Energy Management

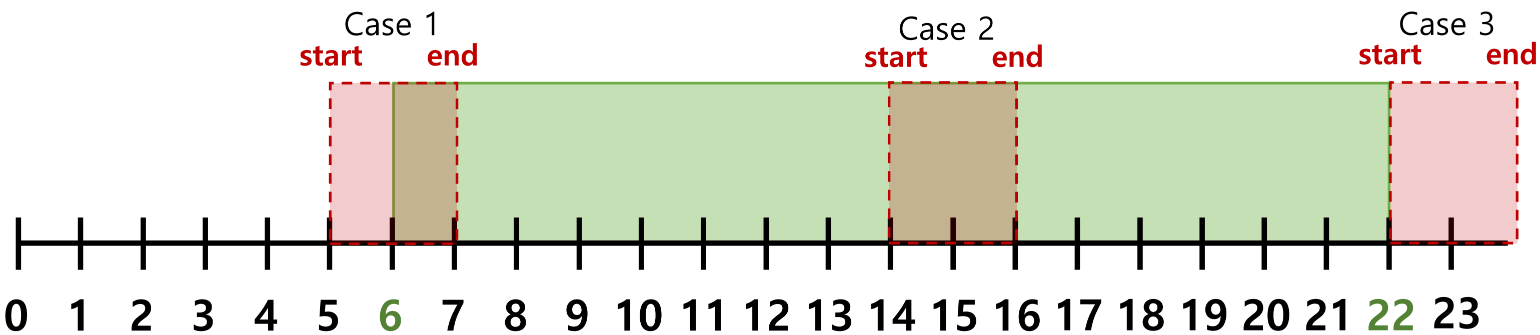
☐ Home Energy Management via Q-Learning

(i) the reward function for the **WM** agent is expressed as

the electric cost and consumer undesired operation of the **WM**

$$r_t^{\mathrm{WM}} = \begin{cases} -[\pi_t E_t^{\mathrm{WM}} + \overline{\delta}(\omega_s^{\mathrm{pref}} - t)], & \text{if } t < \omega_s^{\mathrm{pref}}, \\ -[\pi_t E_t^{\mathrm{WM}} + \underline{\delta}(t - \omega_f^{\mathrm{pref}})], & \text{if } t > \omega_f^{\mathrm{pref}}, \\ -\pi_t E_t^{\mathrm{WM}}, & \text{otherwise,} \end{cases}$$

Case 1
Case 2
Case 3



- w_s^{pref}, w_f^{pref} : the consumer **preferred starting** and **finishing** times of the WM
- $\overline{\delta}, \underline{\delta}$: the **penalties** for **early** and **late** operation, respectively, compared to the consumer preferred operation interval

Formulation of RL- and ANN-Based Home Energy Management

☐ Home Energy Management via Q-Learning

(ii) the reward function for the **AC** agent is expressed as

the electric cost and consumer thermal discomfort of the **AC**

Prediction of Indoor Temperature via ANN

$$r_t^{\mathrm{AC}} = \begin{cases} -\left[\pi_t E_t^{\mathrm{AC}} + \kappa\left(T^{\min} - T_t^{\mathrm{in}}\right)\right], & \text{if } T_t^{\mathrm{in}} < T^{\min}, \\ -\left[\pi_t E_t^{\mathrm{AC}} + \kappa\left(T_t^{\mathrm{in}} - T^{\max}\right)\right], & \text{if } T_t^{\mathrm{in}} > T^{\max}, \\ -\pi_t E_t^{\mathrm{AC}}, & \text{otherwise}, \end{cases}$$

- T_t^{in} : **Prediction of Indoor** Temperature via ANN
- T^{min}, T^{max} : the consumer **preferred temperature**
- k : the **penalty** for the consumer thermal discomfort

NICELAB
Networking for Intelligence Communications and Energy

서강대학교 Sogang University

Formulation of RL- and ANN-Based Home Energy Management

☐ Home Energy Management via Q-Learning

(iii) the reward function for the ESS agent consists of a negative electric cost and negative energy underutilization cost

$$r_t^{\mathrm{ESS}} = \begin{cases} -[\pi_t E_t^{\mathrm{ESS}} + \overline{\tau}(SOE_t - SOE^{\max})], & \text{if } SOE_t > SOE^{\max}, \textit{ overcharging} \\ -[\pi_t E_t^{\mathrm{ESS}} + \underline{\tau}(SOE^{\min} - SOE_t)], & \text{if } SOE_t < SOE^{\min}, \textit{ undercharging} \\ -\pi_t E_t^{\mathrm{ESS}}, & \text{otherwise,} \end{cases}$$

- SOE_t : State of energy of ESS
- SOE^{max}, SOE^{min} : Maximum (Minimum) state of energy of ESS
- $\overline{\tau}, \underline{\tau}$: the **penalties** for the ESS **overcharging** and **undercharging**

Formulation of RL- and ANN-Based Home Energy Management

☐ Prediction of Indoor Temperature via ANN

$$\hat{f}(T^{\text{in}}_{t-1}, T^{\text{min}}, T^{\text{max}}, \hat{T}^{\text{out}}_{t}, E^{\text{AC}}_{t}) = T^{\text{in}}_{t},$$

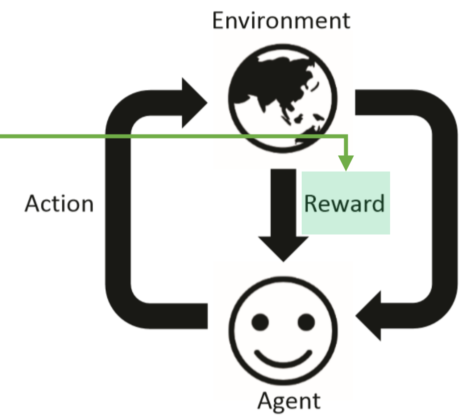
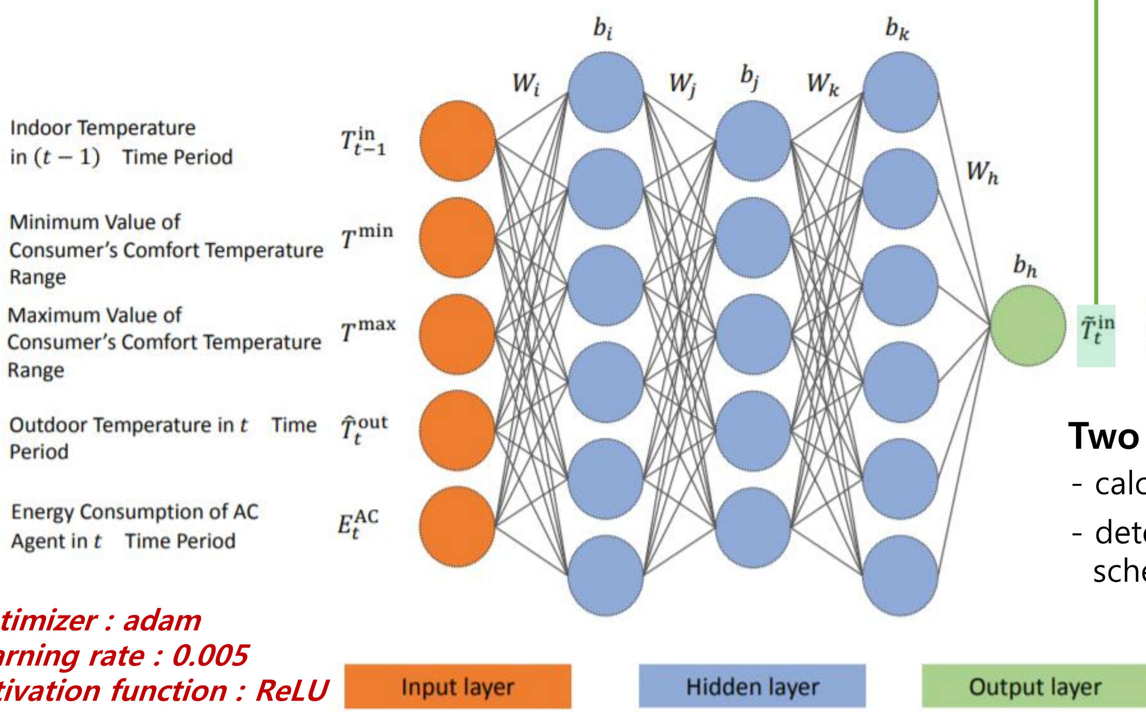


Figure 2. Conceptual architecture of reinforcement learning (RL).



Two advantages

- calculate the dissatisfaction cost more precisely
- determine the optimal energy consumption schedule more efficiently

Optimizer : adam
Learning rate : 0.005
Activation function : ReLU

Figure 3. Architecture of the proposed artificial neural network (ANN) model.

Formulation of RL- and ANN-Based Home Energy Management

☐ Algorithm

Algorithm 1: Q-learning-based energy management of smart home with PV system, ESS, and home appliances.

Initialize each appliance's energy demand, dissatisfaction parameters, and Q-learning parameters
%%Learning with ANN for temperature prediction of AC agent
Indoor temperature at time period $t-1 \rightarrow T^{\text{in}}_{t-1}$
Minimum and maximum value of consumer's comfort temperature range $\rightarrow T^{\text{min}}, T^{\text{max}}$
Predicted outdoor temperature at time $t \rightarrow \hat{T}^{\text{out}}_t$
Energy consumption of AC agent at time $t \rightarrow E^{\text{AC}}_t$
Predicted indoor temperature at time $t \rightarrow T^{\text{in}}_t$
Learning process with ANN and approximate the temperature prediction model $\hat{f}$
$T^{\text{in}}_t = \hat{f}(T^{\text{in}}_{t-1}, T^{\text{max}}, T^{\text{min}}, \hat{T}^{\text{out}}_t, E^{\text{AC}}_t)$
Initialize Q-value of each agent
for *episode* = 1, *MaxEpisode* **do**
  ▷ Initialize state, action and time period
  for *time step* = 1, 24 **do**
    ▷ Select a_t from present s_t using ϵ-greedy policy
    ▷ Take action a_t; observe $r(s_t, a_t)$ and s_{t+1}
    $Q(s_t, a_t) \leftarrow (1-\theta)Q(s_t, a_t) + \theta\,[r(s_t, a_t) + \gamma \max Q(s_{t+1}, a_{t+1})]$
  end
end
Find optimal policy with largest Q-value

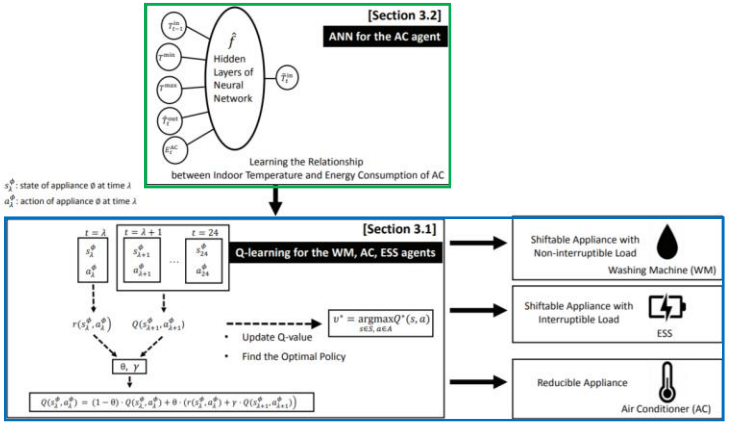


Figure 4. Overall architecture of the proposed HEMS algorithm using Q-learning and ANN.

NICELAB Networking for Intelligence Communications and Energy

서강대학교 SOGANG UNIVERSITY

Numerical Examples

☐ Simulation Setup

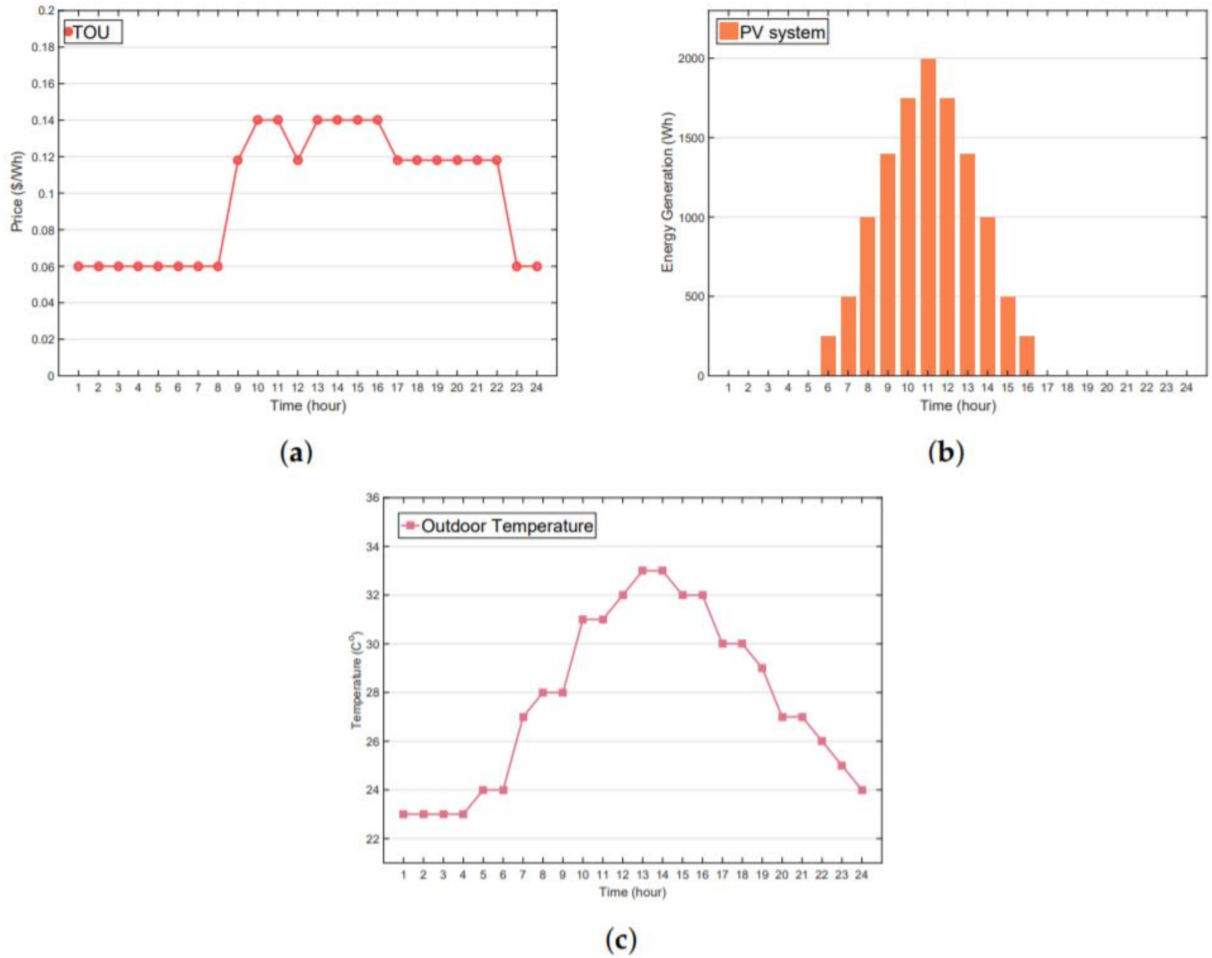


Figure 5. Profiles of electricity price and weather. (**a**) time-of-use (TOU) price; (**b**) solar photovoltaic (PV) generation; (**c**) outdoor temperature.

Numerical Examples

☐ Performance of the Proposed RL-Based HEMS Algorithm

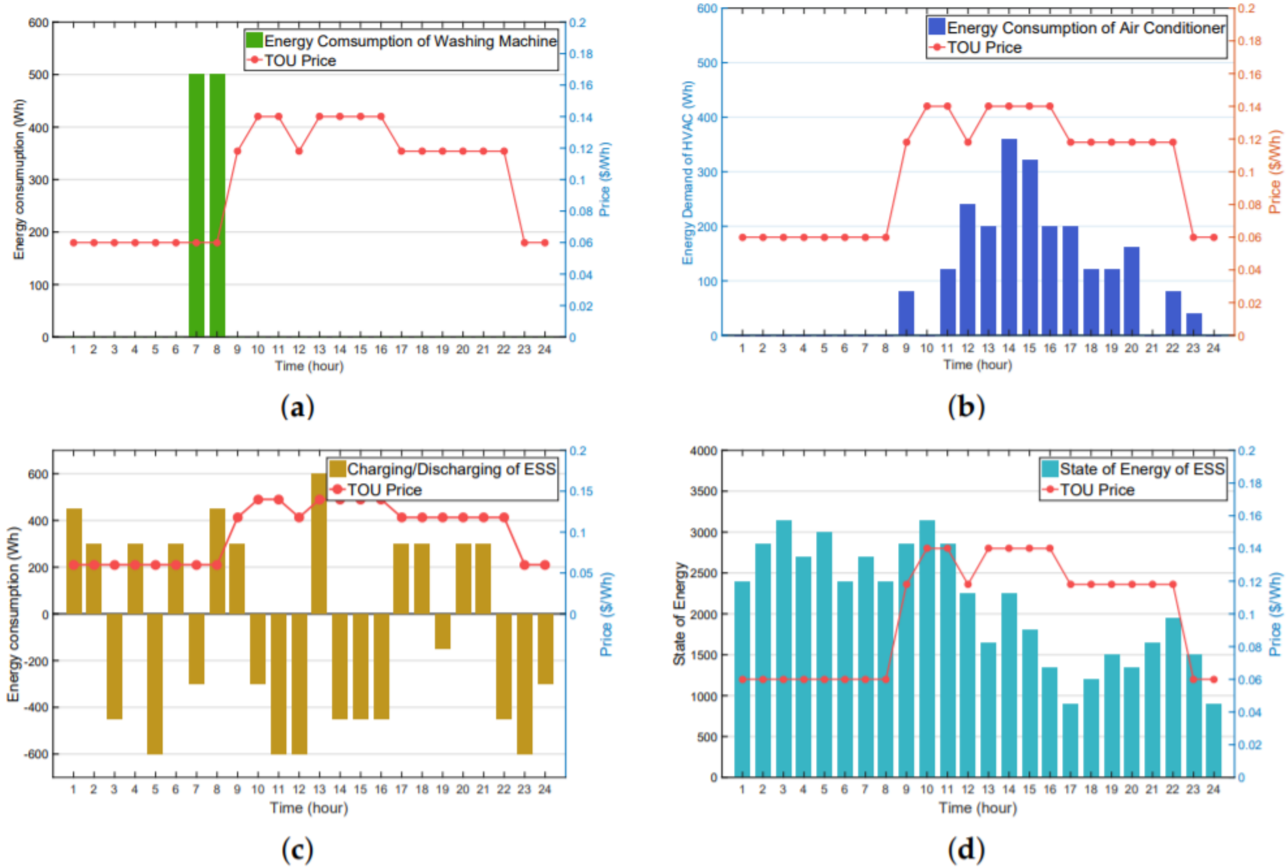


Figure 6. RL-based day-ahead operating schedule of appliance under TOU pricing tariff. (**a**) energy consumption of washing machine (WM); (**b**) energy consumption of air conditioner (AC); (**c**) charging and discharging of energy storage system (ESS); (**d**) state of energy (SOE) of ESS.

Numerical Examples

☐ Impact of Different Parameters in Reward Function on the Proposed Algorithm

$$r_t^{AC} = \begin{cases} -\left[\pi_t E_t^{AC} + \kappa\left(T^{\min} - T_t^{in}\right)\right], & \text{if } T_t^{in} < T^{\min}, \\ -\left[\pi_t E_t^{AC} + \kappa\left(T_t^{in} - T^{\max}\right)\right], & \text{if } T_t^{in} > T^{\max}, \\ -\pi_t E_t^{AC}, & \text{otherwise}, \end{cases}$$

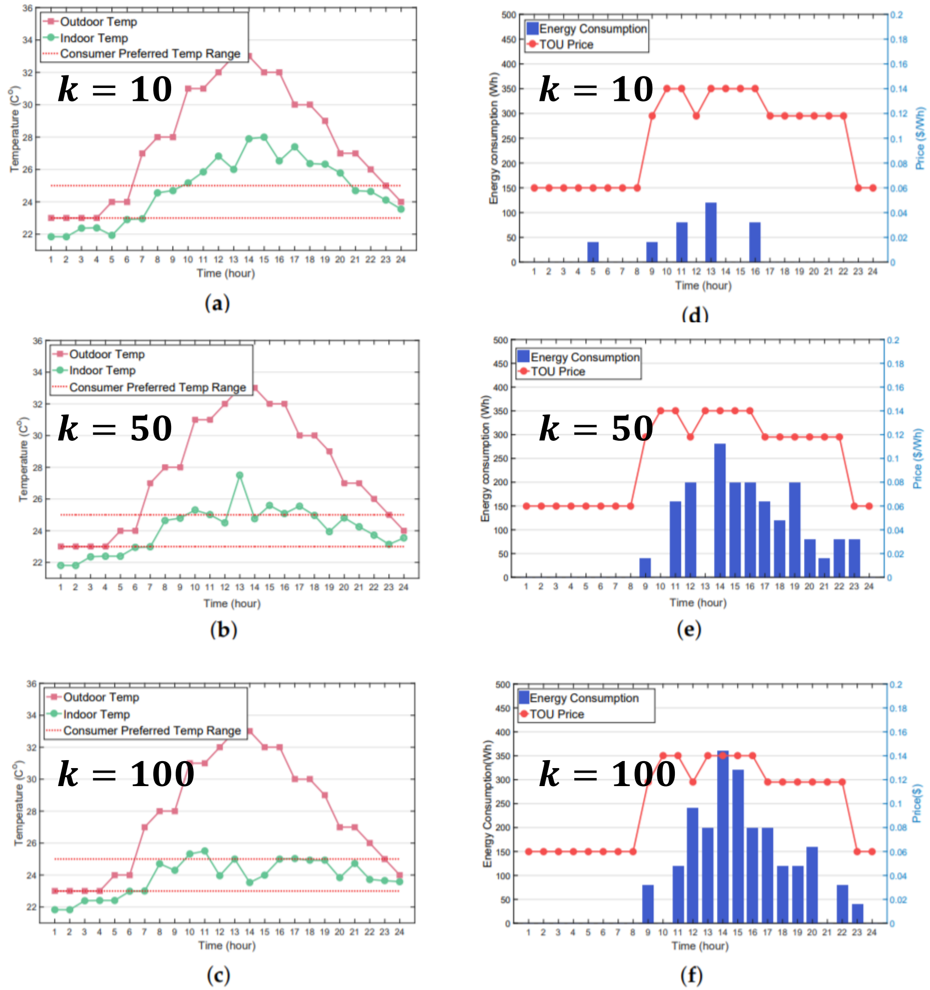


Figure 7. Impact of different penalties (κ) for AC scheduling on indoor temperature (T_t^{in}) and energy consumption ($E_{a,t}$). **(a)** T_t^{in} with $\kappa = 10$; **(b)** T_t^{in} with $\kappa = 50$; **(c)** T_t^{in} with $\kappa = 100$; **(d)** $E_{a,t}$ with $\kappa = 10$; **(e)** $E_{a,t}$ with $\kappa = 50$; **(f)** $E_{a,t}$ with $\kappa = 100$.

NICELAB
Networking for Intelligence Communications and Energy

Numerical Examples

☐ Impact of Different Parameters in Reward Function on the Proposed Algorithm

$$
r_t^{\mathrm{WM}} = \begin{cases} -[\pi_t E_t^{\mathrm{WM}} + \overline{\delta}(\omega_s^{\mathrm{pref}} - t)], & \text{if } t < \omega_s^{\mathrm{pref}}, \\ -[\pi_t E_t^{\mathrm{WM}} + \underline{\delta}(t - \omega_f^{\mathrm{pref}})], & \text{if } t > \omega_f^{\mathrm{pref}}, \\ -\pi_t E_t^{\mathrm{WM}}, & \text{otherwise}, \end{cases}
$$

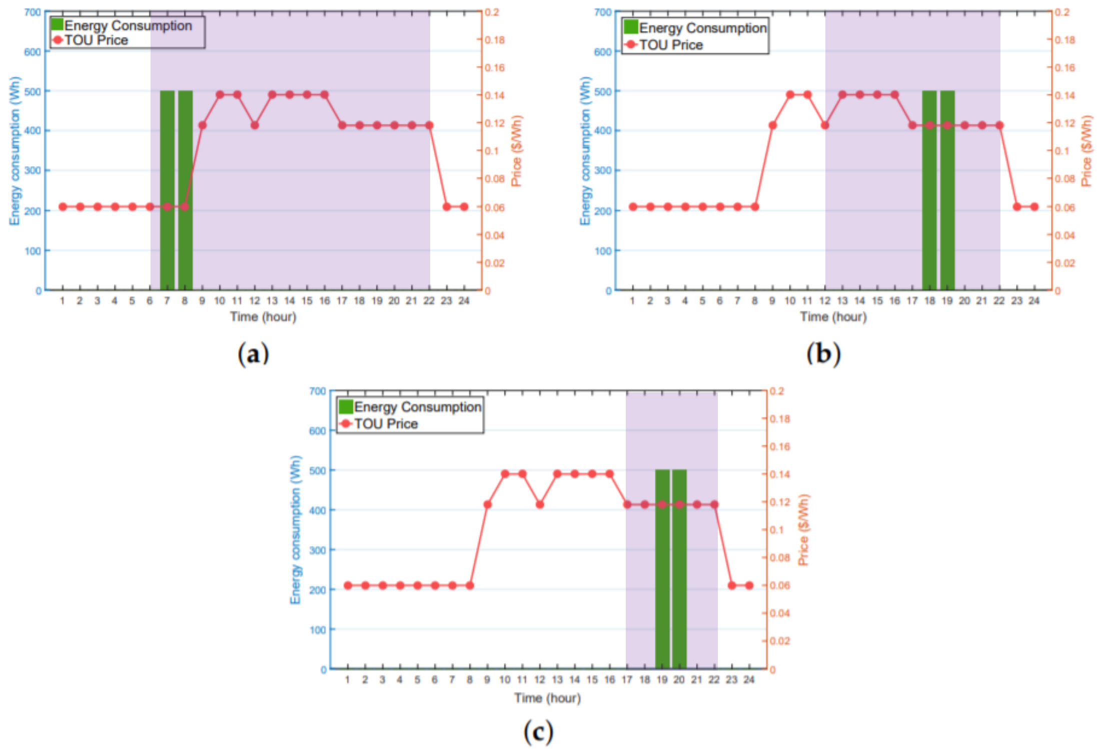


Figure 8. Impact of different preferred operating time interval $[\omega_s^{\mathrm{pref}}, \omega_f^{\mathrm{pref}}]$ for WM scheduling on energy consumption ($E_{a,t}$). **(a)** (6:00 a.m., 10:00 p.m.); **(b)** (12:00 p.m., 10:00 p.m.); **(c)** (5:00 p.m., 10:00 p.m.).

Numerical Examples

☐ Impact of ANN on AC Agent Performance

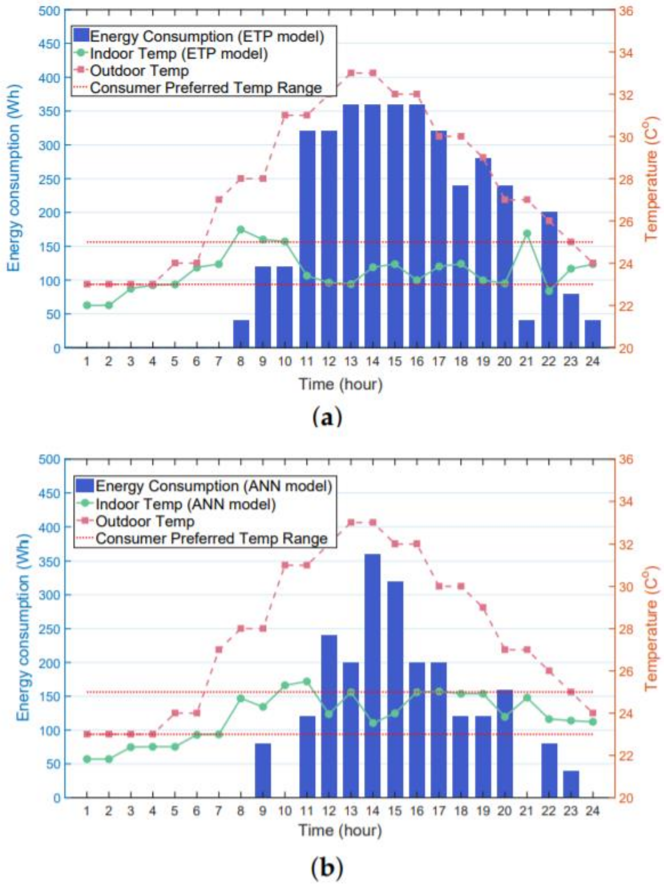


Figure 9. RL-based day-ahead energy consumption schedule of air conditioner through indoor temperature prediction using: (**a**) equivalent thermal parameters (ETP) model; (**b**) ANN model.

Numerical Examples

☐ Performance Comparison between MILP- and RL-Based HEMS

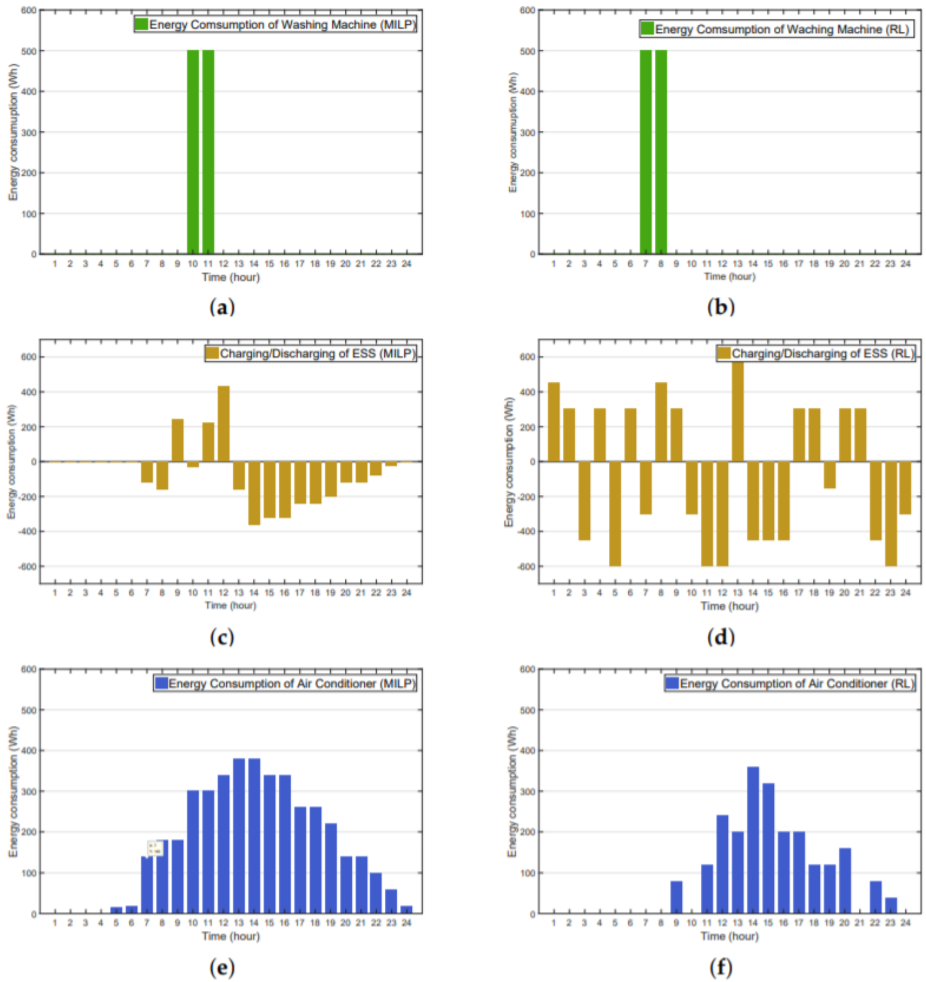


Figure 10. Comparison of energy consumption and charing/discharging energy between mixed-integer linear programming (MILP) and RL methods. **(a)** WM with MILP; **(b)** WM with RL; **(c)** ESS with MILP; **(d)** ESS with RL; **(e)** air conditioner (AC) with MILP; **(f)** AC with RL.

Numerical Examples

☐ Performance Comparison between MILP- and RL-Based HEMS

$$\frac{X^{\text{bill,MILP}} - X_p^{\text{bill,RL}}}{X^{\text{bill,MILP}}} \times 100(\%),$$

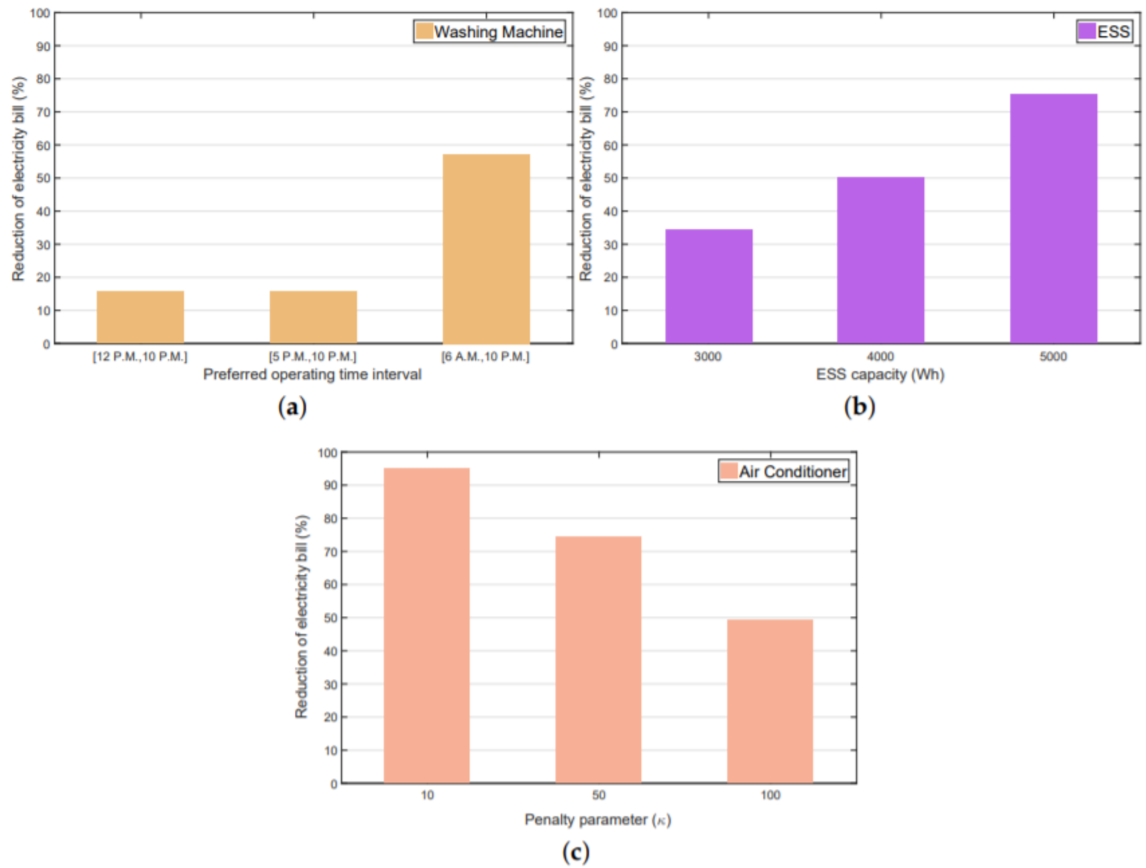


Figure 11. Reduction in electricity bill with different operating conditions of appliances. (**a**) WM under varying preferred operating time intervals; (**b**) ESS under varying capacity; (**c**) AC under varying penalty parameters.

Numerical Examples

☐ Performance Comparison between MILP- and RL-Based HEMS

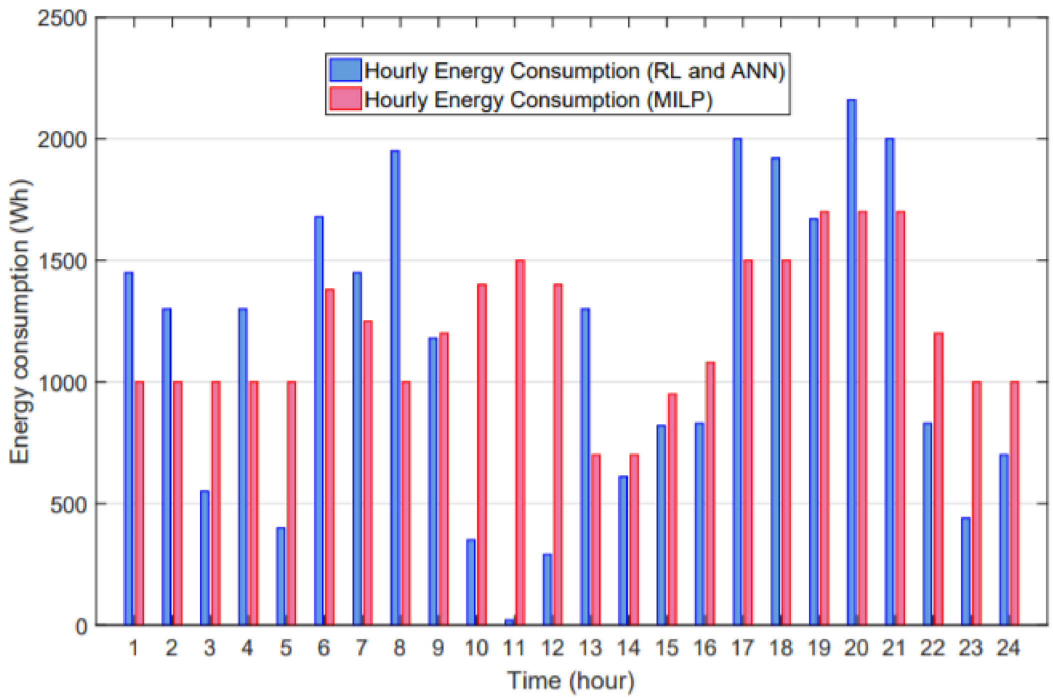


Figure 12. Comparison of hourly energy consumption schedule between MILP method and RL methods.

Discussion / Future work

☐ Discussion

- Wholesale and Retail Electricity Markets under Real-Time Pricing (RTP)
- Electric Vehicle (EV) Integration
- Constraint of the Lifetime for ESS

- *Since the data used in the simulation is a deterministic value, it is different from reality.*

☐ Future work

- Develop A **multi-agent reinforcement learning algorithm**
- The practical implementation of the developed algorithm should be **tested in large-scale realistic electric power networks.**
- Plan to integrate advanced neural network models such as **recurrent neural networks and long short-term memory** in the proposed framework to **improve the prediction accuracy of the indoor temperature.**

Code implementation

☐ Visualization

- "AC_agent" 에 넘기는 reward에 따라 결과가 다름

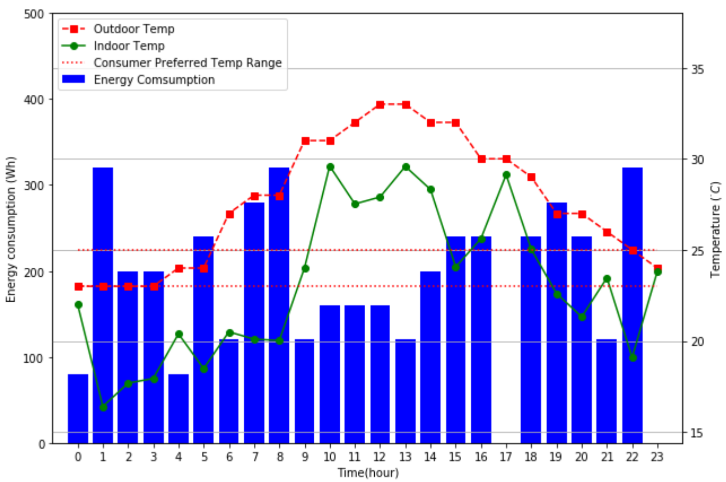
$$T_t^{\text{in}} = \hat{f}(T_{t-1}^{\text{in}}, T^{\max}, T^{\min}, \hat{T}_t^{\text{out}}, E_t^{\text{AC}})$$ X

수집된 data 가 없는 관계로
***ETP** 식으로 대체*

$$T_t^{\text{in}} = T_{t-1}^{\text{in}} + \alpha(\hat{T}_{t-1}^{\text{out}} - T_{t-1}^{\text{in}}) + \beta E_{a,t},$$ O

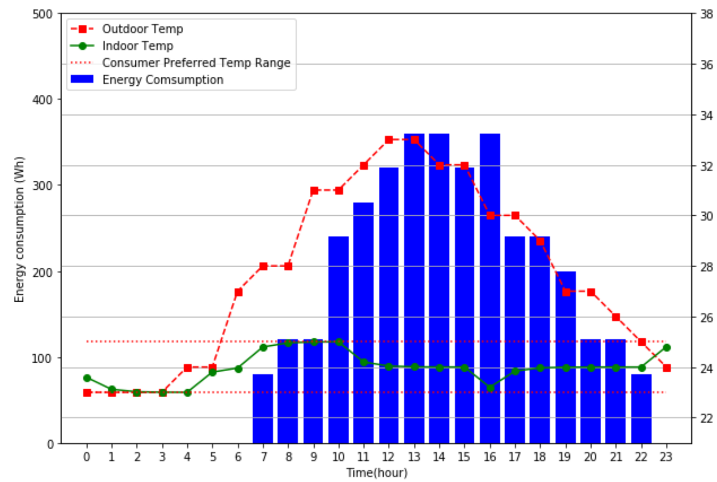
paper

$$reward = r_t^{WM} + r_t^{AC} + r_t^{ESS}$$



our code

$$reward = r_t^{AC}$$

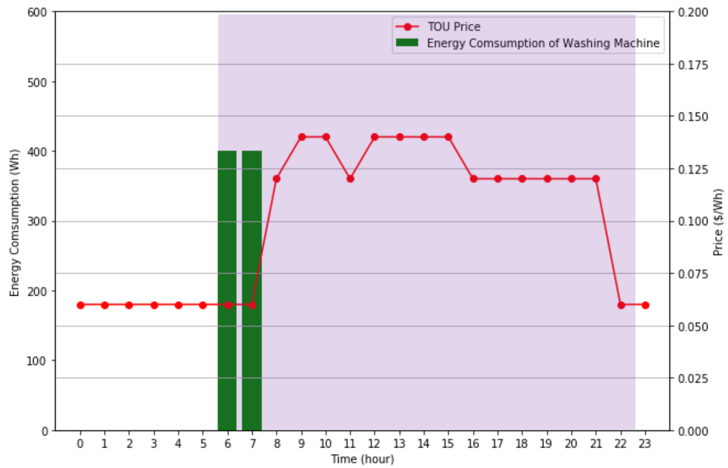


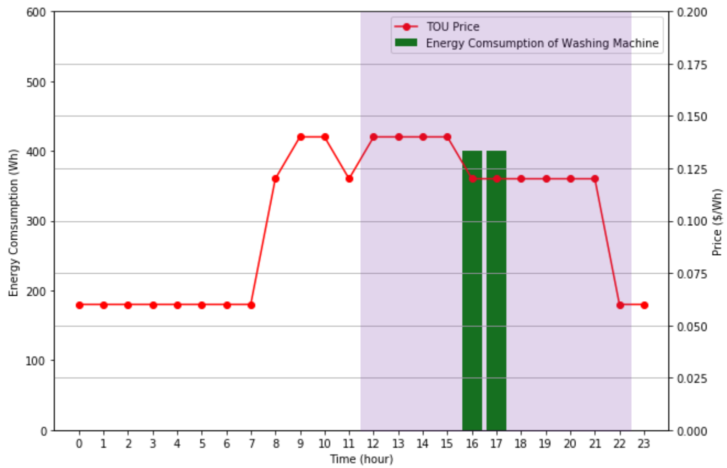
- ETP : the equivalent thermal parameters (ETP) model Equation

Code implementation

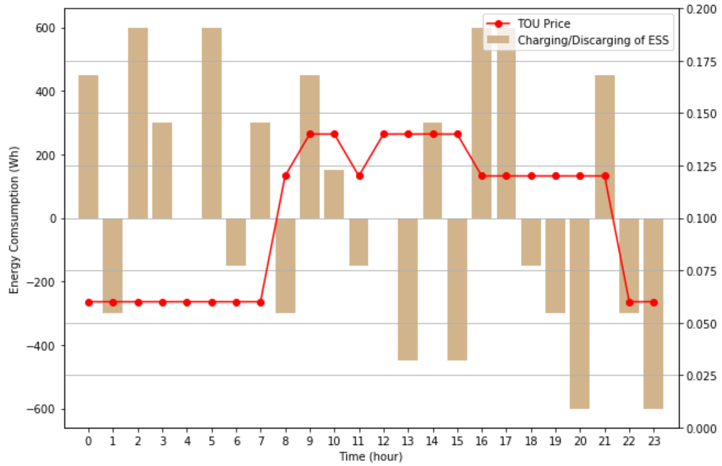
☐ Visualization

- $\left[\omega_s^{pref}, \omega_f^{pref}\right]$를 $[6, 22]$ 및 $[12, 22]$로 설정한 결과 TOU에 따라 스케줄이 다름을 확인





- 최종 “ESS_agent”로 부터 도출된 시간 및 TOU에 따른 ESS 스케줄 시각화



Networking
Next

Intelligence
Innovative

Communications
Creative

Energy
Envisioning

NICELAB
Networking for Intelligence Communications and Energy