

Received November 1, 2021, accepted November 22, 2021, date of publication December 6, 2021, date of current version December 14, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3132962

User Preference-Based Demand Response for Smart Home Energy Management Using Multiobjective Reinforcement Learning

SONG-JEN CHEN, WEI-YU CHIU^{ID}, (Member, IEEE), AND WEI-JEN LIU

Multi-Objective Control and Reinforcement Learning (MOCaRL) Laboratory, Department of Electrical Engineering, National Tsing Hua University, Hsinchu 300044, Taiwan

Corresponding author: Wei-Yu Chiu (chiuweiyu@gmail.com)

This work was supported by the Ministry of Science and Technology of Taiwan under Grant MOST 108-2221-E-007-100 and Grant MOST 110-2221-E-007-097-MY2.

ABSTRACT A well-designed demand response (DR) program is essential in smart home to optimize energy usage according to user preferences. In this study, we proposed a multiobjective reinforcement learning (MORL) algorithm to design a DR program. The proposed approach improved conventional algorithms by mitigating the effect of the change in user preferences and addressed the uncertainty induced by future price and renewable energy generation. Because two Q-tables were used, the proposed algorithm simultaneously considers electricity cost and user dissatisfaction; when user preference changes, the proposed MORL algorithm uses the previous experience to customize appliances' scheduling and swiftly achieve the best objective value. The generalizability of the proposed algorithm is high. Therefore, the algorithm can be implemented in a smart home equipped with an energy storage system, renewable energy source, and various types of appliances such as inflexible, time-flexible, and power-flexible ones. Numerical analysis using real-world data revealed that in case of price and renewable uncertainty, the proposed approach can deliver excellent performance after a change of user preference; it achieved 8.44% cost reduction as compared with mixed-integer nonlinear programming based DR while increasing the dissatisfaction level only by 1.37% on average.

INDEX TERMS Energy management system (EMS), reinforcement learning (RL), multiobjective reinforcement learning (MORL), demand response (DR), smart home.

I. INTRODUCTION

Demand response (DR) [1] refers to the changes in the energy demand of devices in response to time-varying energy prices. A DR program can reduce the cost of energy through adjustment of appliance operation. In a smart home, an energy management system (EMS) [2] plays a crucial role in a DR program. An EMS can shift certain demands for appliances from peak price hours to nonpeak price hours. This adjustment reduces the demand for electricity during peak load times [3]. A well-designed DR scheme with an EMS can benefit users, improve human comfort level, reduce energy consumption (or electricity cost), and reduce the system's peak demand.

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian^{ID}.

Incentive-based programs and price-based programs are two major types of DR programs [2]. In incentive-based programs, incentive payments are given to induce lower energy usage at times of high electricity prices. Users adjust the energy usage after incentive is provided. In price-based programs, users can obtain financial benefits and discounts in return for their peak demand shifted and consumption reduced at designated times. Users can respond to the price change promptly when adopting a price-based program [4], [5]. Price-based DR programs for smart homes have been widely studied in recent years [6], [7]. In price-based programs, the power utility provider sets a time-based electricity tariff to encourage users to voluntarily adjust their energy consumption. One of the electricity tariffs commonly used is real-time pricing (RTP) whose pricing signals change with time [8]–[10]. The peak price is set at the time of peak demand

to encourage users to curtail their energy usage. In smart homes, an EMS can shift the energy demand to nonpeak price hours and thus reduce the total electricity cost. However, although this method helps reduce the electricity cost, the user comfort level can also decrease, which consequently increases user dissatisfaction, because of partial satisfaction of energy demand required by users [11]. A well-designed DR scheme and an EMS can allow users to automatically manage demand and balance electricity cost and user dissatisfaction [12].

When a DR program is used to optimize energy resources, the objective is to minimize the electricity cost and user dissatisfaction. In [13]–[17], the authors applied mixed-integer nonlinear programming (MINLP) to achieve the optimal DR. Meanwhile, optimizing energy resources often involves the use of an EMS, which has been used by, for example, companies and academic institutions such as Aggregate Industries, Thorn-Zumtobel, and Sheffield Hallam University to save energy costs. These approaches require complete information, such as full-day electricity prices and renewable supply, to determine the optimal scheduling. In practice, however, users are only provided with forecast electricity prices and renewable supply; data uncertainty must be considered.

To address this uncertainty, machine learning technologies have been used for energy management [18], [19]. Reinforcement learning (RL) is a machine learning-based decision-making algorithm that can address an uncertain environment using limited information [20]. The use of RL involves the identification of agents and the environment. Agents interact with the environment to receive feedback from the environment; this reward (feedback) is then used to construct a Q-table, which stores Q-values for the evaluation of each state–action pair.

Conventional RL algorithms have been applied for energy management in smart homes over the past few years [21]. Ruelens *et al.* [22] developed a batch RL technique called fitted Q-iteration to perform day-ahead scheduling only for power-flexible appliances. Wen *et al.* [23] proposed an RL-based algorithm to operate home appliances automatically but considered only time-flexible appliances. Lu *et al.* [24] proposed a multiagent RL technique to schedule the operation of various home appliances in a decentralized manner. However, they did not include renewable energy sources and energy storage systems and the influence of the prediction error was not discussed. Xu *et al.* [25] developed a multiagent RL-based EMS but did not consider an energy storage system to reduce the energy bill. Remani *et al.* [26] proposed a RL-based model to control load balance problem but trained in single agent. Mathew *et al.* [27] developed a priced-based deep RL-based EMS. In modern smart homes, various home appliances and their interaction with the energy storage system as well as renewable energy sources should be considered.

In [22], [23], and [25], the authors formulated a DR problem as a single objective RL (SORL) problem. They used weighted-sum methods to combine multiple objectives.

Park *et al.* [28] and Lu *et al.* [24] devised the concepts of using various weights but they did not consider possible weight changes in response to user preferences. A user-friendly DR program was used to customize the scheduling of appliance operation according to user preferences. However, user preferences are assumed to be fixed, and possible changes over time are not considered. If user preferences change, learning algorithms can approximate an optimal scheduling after a certain period. However, the speed of approximating a new scheduling after change of preference depends on various factors. These factors have not been sufficiently investigated. To address the change of user preferences, a mechanism to expedite the learning process must be implemented in order to minimize the learning costs.

Multiobjective RL (MORL) [29] is preferred over the SORL to address the problem of multiple conflicting objectives mainly because of two reasons. First, MORL improves the performance of SORL by developing diverse Pareto optimal models with multiple objectives. Second, MORL is suitable for optimizing multiple objectives simultaneously for sequential decision-making problems. While the concept of MORL has emerged recently, the applications of MORL to the smart home energy management have not been explored. Therefore, further research is required in this field.

Due to the advantages of MORL over SORL, we proposed a user preference-based DR program that uses MORL for energy management in a smart home. We implemented the proposed DR program in a smart home equipped with various types of home appliances, an energy storage system, and a renewable energy source. Appliances and the energy storage system are considered as agents that aim to approach optimal energy scheduling in the presence of uncertainty from renewable energy generation and electricity price. To handle a possible change in user preferences, two Q-tables were used in the proposed approach to simultaneously consider electricity cost and user dissatisfaction. After the user preference changes, the previous experience can be used to swiftly approach the best objective value without undergoing any time-consuming learning process.

The main contributions of this paper are as follows. First, the proposed user preference-based DR program using MORL algorithm can swiftly approach the best objective value after the user preference changes. An EMS is used to rapidly customize energy scheduling according to a change in the user preference. The concept of using the MORL to improve the learning speed and its application to home energy management has not been investigated in the literature. Second, the proposed home energy management approach is flexible and can address a wide range of appliances and other smart home components, including inflexible, time-flexible, power-flexible appliances, an energy storage system, and a renewable energy source, whereas existing approaches are applicable to only some of these devices. This feature of the proposed approach is critical because users desire a DR program that can handle various types of appliances, and smart homes are often equipped with an energy storage system and

renewable energy source to improve the energy efficiency. Third, a numerical analysis was conducted using real-world data; the results revealed that the proposed approach outperformed existing methods in terms of electricity cost and user dissatisfaction while simultaneously considering the price and renewable energy uncertainty.

The remainder of this paper is organized as follows. Section II discusses the related work. Section III details the scheme of a smart home, system models of diverse categories of home appliances, and the energy storage system. In Section IV, a user preference-based DR program that uses the MORL algorithm is proposed with dedicated designs of states, actions and rewards. Section V presents our simulation results and a comparison of the existing smart home energy management methods. Finally, conclusions are drawn in Section VI.

II. RELATED WORK

In a smart home, a renewable system has been used to supply electricity to home appliances [30]–[32]. To predict future renewable supply, various weather forecasting methods have been developed [33]–[35]. For example, Cerrai *et al.* [36] proposed an outage prediction model for an electric distribution network.

A battery system is often used to reduce the variability of renewable energy. As such, battery models have been investigated [37], [38]. Price forecasting [39]–[41] or prediction of energy consumption [42]–[44] may be involved in a battery system to form an EMS.

Prediction errors were observed in conventional prediction algorithms [45]–[47]. To reduce the prediction error, for example, Kodaira *et al.* [48] proposed to use prediction intervals of estimated error based on prior predictions to approach optimal energy scheduling; Lee *et al.* [49] applied an error driven prediction modulation to evaluate the difference between the forecast load and actual load.

To improve energy efficiency, hybrid systems have been examined [50]–[52]. For example, Li *et al.* [53] considered photovoltaic battery energy storage systems as a black start power source that can improve the regional power grid and broaden the application of photovoltaic (PV) power generation; Yulong *et al.* [54] proposed a hybrid energy storage system which consists of ultra capacitors and battery packs. The hybrid system was able to prevent the battery from a large current impact and increase instantaneous battery capacity.

While energy scheduling within a smart home has been widely investigated, its impact can extend to a larger scale if transactive energy is applied. Transactive energy is related to economic and control techniques that can be employed to exchange energy or power flow within a region of interest, such as a multimicrogrid system or a community. For instance, Iqbal *et al.* [55] proposed a metaheuristic polar bear optimization method to solve a DR problem considering interconnected microgrids in a transactive energy market. Meanwhile, blockchain technologies can be employed to ensure the security of energy transactions [56]–[58].

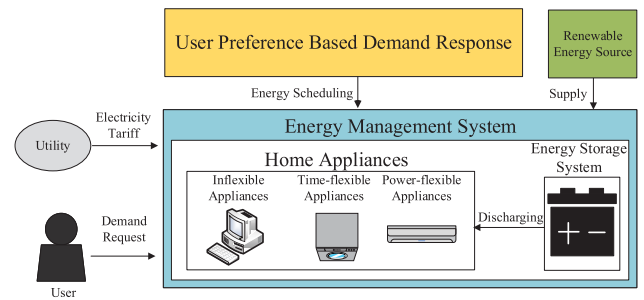


FIGURE 1. Home energy management system.

III. SYSTEM MODELS AND PROBLEM FORMULATION

Fig. 1 illustrates the smart home environment considered in this study. Typically, smart homes include an EMS, an energy storage system, a renewable energy source, and several home appliances. A user sends a full-day demand request list to the EMS.

In this study, we primarily focus on the price-based programs to optimize energy resources. The electricity can be bought from a utility or supplied from renewable energy sources. The price of electricity will fluctuate due to several factors, including fuels, power plant costs, transmission and distribution system, weather conditions, and regulations [59]. The electricity price is automatically updated periodically in the EMS to evaluate the energy bill and a renewable energy source is used to charge the energy storage system. The EMS controls home appliances and the energy storage system to improve energy usage and reduce the energy bill. Home appliances are categorized into three groups, namely inflexible, time-flexible, and power-flexible [24], [60]. The following subsections present models of home appliances, energy storage system, and our problem formulation.

A. INFLEXIBLE APPLIANCES

Inflexible appliances have critical energy demand that must be satisfied in each time slot. Inflexible appliances, such as refrigerators, must be operated continuously, and their energy demand cannot be adjusted. The energy consumption of inflexible appliance n in time slot h is denoted as $E_{n,h}^a$, where $n \in \{1, 2, 3, \dots, N^a\}$ and $h \in \{1, 2, 3, \dots, H\}$. Here, N^a and H represent the number of inflexible appliances and the optimization window, respectively. The energy consumption of all inflexible appliances is as follows:

$$E^a = \sum_{h=1}^H \sum_{n=1}^{N^a} E_{n,h}^a. \quad (1)$$

B. TIME-FLEXIBLE APPLIANCES

Time-flexible appliances, such as washing machines, can be scheduled using an EMS to operate at nonpeak hours to reduce the energy bill. For time-flexible appliances, users can send the request to the EMS and set an operation period. The EMS must finish the job in this operation period. The energy consumption of time-flexible

appliance n in time slot h is denoted as $E_{n,h}^b$, where $n \in \{1, 2, 3, \dots, N^b\}$ with N^b representing the number of time-flexible appliances. The energy consumption of all time-flexible appliances during the observation period is as follows:

$$E^b = \sum_{h=1}^H \sum_{n=1}^{N^b} E_{n,h}^b. \quad (2)$$

Shifting the operation time of a device can result in user dissatisfaction. The user dissatisfaction, denoted as $U_{n,h}^b$, for appliance n in time slot h increases with the increment of shifted operation time. Here, $h_{n,s}^b$ is the start time of operation conducted by the EMS, $h_{n,r}^b$ is the request start time, and $h_{n,e}^b$ is the end time. The EMS can shift the start time in the available operation period $[h_{n,r}^b, h_{n,e}^b]$. Typically, the priority for various time-flexible appliances is different. Here, θ_n^b denotes the device-dependent dissatisfaction parameter of time-flexible appliance n with $0 < \theta_n^b \leq 1$. A larger θ_n^b implies a higher priority for time-flexible appliance n . The user dissatisfaction is related to dissatisfaction parameters and the operation time.

We denote user dissatisfaction as $U_{n,h}^b(\theta_n^b, h_{n,s}^b)$ with the following variable constraints:

$$h_{n,r}^b \leq h_{n,s}^b \leq h_{n,e}^b. \quad (3)$$

For example, user dissatisfaction can be linearly related to the difference between the actual start time $h_{n,s}^b$ and the request start time $h_{n,r}^b$ [61]

$$U_{n,h}^b = \theta_n^b(h_{n,s}^b - h_{n,r}^b). \quad (4)$$

C. POWER-FLEXIBLE APPLIANCES

Power-flexible appliances, such as air conditioners, are common in a smart home. The use of power-flexible appliances can be adjusted to satisfy the energy demand during a requested time. Power-flexible appliances can operate with flexible power consumption between the minimum energy demand and maximum energy demand. The energy consumption of power-flexible appliance n in time slot h is denoted by $E_{n,h}^c$, where $n \in \{1, 2, 3, \dots, N^c\}$ and N^c is the number of power-flexible appliances. The energy consumption of all power-flexible appliances is calculated as follows:

$$E^c = \sum_{h=1}^H \sum_{n=1}^{N^c} E_{n,h}^c. \quad (5)$$

An EMS can reduce the energy cost through the adjustment of power-flexible appliances, but this adjustment can increase user dissatisfaction.

Here, $d_{n,\max}^c$ is the desired energy demand for the power-flexible appliance n . The minimum energy demand is denoted by $d_{n,\min}^c$. The EMS controls the energy consumption ranging from $d_{n,\min}^c$ to $d_{n,\max}^c$. Furthermore, θ_n^c denotes the device-dependent dissatisfaction parameter of power-flexible appliance n with $0 < \theta_n^c \leq 1$. A larger

θ_n^c implies a higher priority for power-flexible appliance n . The user dissatisfaction $U_{n,h}^c(\theta_n^c, E_{n,h}^c)$ of power-flexible appliance n is related to the energy consumption and dissatisfaction parameters, and it has the following variable constraints

$$d_{n,\min}^c \leq E_{n,h}^c \leq d_{n,\max}^c. \quad (6)$$

For example, $U_{n,h}^c(\theta_n^c, E_{n,h}^c)$ can be formulated as follows [24]:

$$U_{n,h}^c = \theta_n^c(E_{n,h}^c - d_{n,\max}^c)^2. \quad (7)$$

D. ENERGY STORAGE SYSTEM

In this study, a battery system was used as the energy storage system. A battery system can charge at nonpeak hours and discharge electrical energy to appliances at peak hours to reduce the energy bill. The energy level of the battery system in time slot h is denoted by B_h kWh, which satisfies the following equation [62]:

$$B^{\min} \leq B_h \leq B^{\max} \quad (8)$$

where B^{\min} and B^{\max} are the minimum battery level and maximum battery capacity, respectively. Here, B_h is mainly influenced by charging and discharging activities, denoted as P_h kW. The renewable energy source charges the battery system in time slot h and it is denoted by E_h^{res} kWh. Positive P_h represents charging and negative P_h represents discharging, E_h^{res} represents the energy supply from the renewable energy source to the battery system. The power leakage occurs with a power leakage rate β , where $0 < \beta \leq 1$. The dynamics of the battery system can be expressed as follows [62], [63]:

$$B_{h+1} = \beta B_h + \eta P_h \Delta h + E_h^{\text{res}} \quad (9)$$

where Δh is the length of a time slot and η is the charging/discharging efficiency.

Let the following equation be the total energy consumption of all home appliances:

$$E_h^{\text{HA}} = \sum_{n=1}^N (E_{n,h}^a + E_{n,h}^b + E_{n,h}^c). \quad (10)$$

Because of the energy demand of appliances and limited battery capacity, the charging/discharging power P_h should satisfy the following equation:

$$P_h^{\min} \leq P_h \leq P_h^{\max} \quad (11)$$

where maximum power P_h^{\max} and minimum power P_h^{\min} in charging or discharging events are calculated as follows:

$$P_h^{\max} = \frac{\eta(B^{\max} - B_{h-1}) - E_h^{\text{res}}}{\Delta h} \quad (12)$$

$$P_h^{\min} = \max\left\{\frac{\eta(B^{\min} - B_{h-1}) - E_h^{\text{res}}}{\Delta h}, -\frac{E_h^{\text{HA}}}{\Delta h}\right\}. \quad (13)$$

E. PROBLEM FORMULATION

The electricity cost of a smart home in time slot h , denoted as $f_{1,h}$, is the multiplication of energy bought from the utility and RTP λ_h^{RTP} cent/kWh:

$$\begin{aligned} f_{1,h} &= \lambda_h^{\text{RTP}} \times E_h^{\text{HA}} + \lambda_h^{\text{RTP}} \times P_h \Delta h \\ &= \lambda_h^{\text{RTP}} (E_h^{\text{HA}} + P_h \Delta h). \end{aligned} \quad (14)$$

For a charging event, $P_h > 0$ and the term $E_h^{\text{HA}} + P_h \Delta h$ represents the energy consumption of appliances and battery charging. For a discharging event, $P_h < 0$ and energy discharged from the battery, i.e., $P_h \Delta h < 0$, is used by the appliances; hence, the difference $E_h^{\text{HA}} + P_h \Delta h$ represents the total energy that should be bought from the utility.

The total user dissatisfaction is denoted as $f_{2,h}$. The two objectives must be minimized:

$$f_{2,h} = \sum_{n=1}^N \{\theta_n^b (h_{n,s}^b - h_{n,r}^b) + \theta_n^c (E_{n,h}^c - d_{n,\max}^c)^2\}. \quad (15)$$

Consider weight w with $0 \leq w \leq 1$ that represents user preference on objectives. A larger weight indicates the objective that the user prefers. A user preference-based DR in a smart home can be realized by solving

$$\begin{aligned} \min \quad & \mathbb{E} \left[\sum_{h=1}^H (w f_{1,h} + (1-w) f_{2,h}) \right] \\ \text{subject to} \quad & (3), (6), (8), \text{ and } (11). \end{aligned} \quad (16)$$

In this study, we focus on the dynamic adjustment in the EMS, where the agent can make decisions based on the user's preferences for the target, i.e., w and $1-w$. The weight may change with time; the EMS must reschedule energy when the user preference changes in the presence of uncertainty.

IV. USER PREFERENCE-BASED MULTIOBJECTIVE REINFORCEMENT LEARNING APPROACH

Owing to the uncertainty in electricity price and renewable generation, and possible change of user preference, this section develops a user preference-based DR approach that solves (16) using an MORL algorithm. The home appliances and battery system were considered as agents that interact with the environment (utility, renewable energy source, and the user). The EMS regulates the operation of appliances to reduce energy cost and dissatisfaction from the user. Furthermore, the EMS controls the charging/discharging of the battery. In this section, the framework of MORL is first presented, followed by our customized RL algorithms for energy control and scheduling of the charging/discharging of the battery and home appliances.

The framework of multiagent RL is based on a Markov decision process. Here, \mathcal{S}_n is the state set of agent n and $\mathbf{s}_{n,h}$ is the state in time slot h , where $\mathbf{s}_{n,h} \in \mathcal{S}_n$. Given $\mathbf{s}_{n,h}$, agent n selects an action $\mathbf{a}_{n,h}$ from its action set \mathcal{A}_n , and then proceeds to the next state $\mathbf{s}_{n,h+1}$ and receives reward $r_{n,h+1}$ given by the environment. The reward $r_{n,h+1}$ is used to evaluate action $\mathbf{a}_{n,h}$. The main goal of agent n is to determine

an optimal policy π_n^* or an optimal mapping from states to actions that maximizes the expected value of the cumulative reward. Given a state $\mathbf{s}_{n,h}$, an action $\mathbf{a}_{n,h}$ and a policy π_n , the action value of pair $(\mathbf{s}_{n,h}, \mathbf{a}_{n,h})$ is defined by the following expression:

$$q_{n,\pi_n}(\mathbf{s}_{n,h}, \mathbf{a}_{n,h}) = \mathbb{E}_{\pi_n} \left[\sum_{h=1}^H \gamma^{h-1} r_{n,h} | \mathbf{s}_{n,h}, \mathbf{a}_{n,h} \right], \quad \forall \mathbf{s}_{n,h} \in \mathcal{S}_n, \quad \forall \mathbf{a}_{n,h} \in \mathcal{A}_n \quad (17)$$

where γ is the discount factor. The action-value function under an optimal policy, denoted as $q_n^*(\mathbf{s}_{n,h}, \mathbf{a}_{n,h})$, is expressed as follows:

$$q_n^*(\mathbf{s}_{n,h}, \mathbf{a}_{n,h}) = \max_{\pi_n} q_{n,\pi_n}(\mathbf{s}_{n,h}, \mathbf{a}_{n,h}). \quad (18)$$

Q-learning is often used to achieve or approach the optimality in (18) [20], [64]. In Q-learning, an agent uses a Q-table to estimate q_n^* and has two strategies to select an action: for exploration, the agent randomly selects an action to explore the state space and action space with probability ϵ ; for optimization or exploitation, the agent selects an action on the basis of the current state that yields the highest Q-value with probability $1 - \epsilon$. This mechanism is termed ϵ -greedy action selection. The greedy action $\mathbf{a}_{n,h}^*$ can be determined by

$$\mathbf{a}_{n,h}^* = \arg \max_{\mathbf{a}} Q_n(\mathbf{s}_{n,h}, \mathbf{a}). \quad (19)$$

Given the action selection strategies, Q-learning can approach q_n^* through the following update rule:

$$\begin{aligned} Q_n(\mathbf{s}_{n,h}, \mathbf{a}_{n,h}) &\leftarrow (1 - \alpha) Q_n(\mathbf{s}_{n,h}, \mathbf{a}_{n,h}) \\ &+ \alpha (r_{n,h+1} + \gamma \max_{\mathbf{a}} Q_n(\mathbf{s}_{n,h+1}, \mathbf{a})) \end{aligned} \quad (20)$$

where $\alpha \in (0, 1]$ is the learning rate.

In our smart home scenario, a user desires to minimize the cost and user dissatisfaction; we thus interpret reward r as a penalty, and the aforementioned maximization is replaced by minimization. We refer to home appliance n as agent n and design its state $\mathbf{s}_{n,h}$ as follows:

$$\mathbf{s}_{n,h} = [\lambda_h^{\text{RTP}}, h]. \quad (21)$$

Let $a_{n,h}$ denote the action performed by agent n in time slot h . For inflexible appliances, the action space has only one option "on", because it cannot be scheduled. For time-flexible appliances, the action space has two options which are "on" and "off" to determine when the appliances are operated. For power-flexible appliances, the action space is a set of discrete numbers $\{1, 2, 3, \dots, E_{n,\max}^c\}$; each element in the set represents a level of energy consumption and $E_{n,\max}^c$ is the maximum level. Each agent can receive two reward signals as feedback from the environment. For time-flexible appliances, the reward is expressed as follows:

$$r_{1,n,h} = \lambda_h^{\text{RTP}} E_{n,h}^b, \quad r_{2,n,h} = U_{n,h}^b. \quad (22)$$

For power-flexible appliances, the reward is used as follows:

$$r_{1,n,h} = \lambda_h^{\text{RTP}} E_{n,h}^c, \quad r_{2,n,h} = U_{n,h}^c. \quad (23)$$

Both the aforementioned reward functions are designed to encourage agents to minimize objectives (14) and (15) by updating Q-value $Q_{k,n}(s_{n,h}, a_{n,h})$ as follows:

$$Q_{k,n}(s_{n,h}, a_{n,h}) \leftarrow (1 - \alpha)Q_{k,n}(s_{n,h}, a_{n,h}) + \alpha(r_{k,n,h+1} + \gamma \min_a Q_{k,n}(s_{n,h+1}, a)) \quad (24)$$

where $k \in \{1, 2\}$ represents the k th objective.

We proposed to use an MORL algorithm to expedite the learning process in response to possible changes of the user preference on objectives. Algorithm 1 presents the pseudocode of the proposed user preference-based MORL algorithm for agent n . Agent n selects the action derived from $wQ_{1,n} + (1 - w)Q_{2,n}$. After Q-tables $Q_{1,n}$ and $Q_{2,n}$ reach steady values, greedy action selection can be readily performed given any weight w . Thus, the best objective value $wQ_{1,n} + (1 - w)Q_{2,n}$ can be achieved without further learning when a change in user preference occurs.

Algorithm 1 Proposed User Preference Based MORL for Energy Scheduling of Home Appliance n

Require: dissatisfaction parameters θ_n^b, θ_n^c ; user's request $t_{n,r}, t_{n,e}, d_{n,\min}, d_{n,\max}$; weight w .

Initialize $Q_{k,n}(s_{n,h}, a_{n,h})$ arbitrarily.

Ensure: scheduling policy for home appliance n derived from $wQ_{1,n} + (1 - w)Q_{2,n}$.

- 1: **Loop** for each episode
 - 2: **For** $h = 1, 2, \dots, H$ **do**
 - 3: Choose action $a_{n,h}$ on the basis of current state $s_{n,h}$ using ϵ -greedy derived from $wQ_{1,n} + (1 - w)Q_{2,n}$.
 - 4: Take action $a_{n,h}$, obtain rewards $r_{1,n,h+1}, r_{2,n,h+1}$ and next state $s_{n,h+1}$.
 - 5: Update $Q_{1,n}(s_{n,h}, a_{n,h})$ and $Q_{2,n}(s_{n,h}, a_{n,h})$ using (24).
 - 6: $s_{n,h} \leftarrow s_{n,h+1}$.
 - 7: **End for**
-

The EMS also considers the battery system as an agent in a smart home. Its state is designed as follows:

$$s_h^{\text{ess}} = [B_h, \lambda_h^{\text{RTP}}, h]. \quad (25)$$

The action of battery system agent, denoted by a_h^{ess} , is P_h and subject to (11). Unlike the agents associated with the home appliances, this agent has one reward signal as feedback from the environment:

$$r_h^{\text{ess}} = \lambda_h^{\text{RTP}} P_h \Delta h. \quad (26)$$

The reward is the electricity cost because the goal of the agent is to reduce the energy bill. The agent's Q-table is updated by the following expression:

$$Q^{\text{ess}}(s_h^{\text{ess}}, a_h^{\text{ess}}) \leftarrow (1 - \alpha)Q^{\text{ess}}(s_h^{\text{ess}}, a_h^{\text{ess}}) + \alpha(r_{h+1}^{\text{ess}} + \gamma \min_a Q^{\text{ess}}(s_{h+1}^{\text{ess}}, a)). \quad (27)$$

Algorithm 2 presents the pseudocode of the Q-learning for the battery system. The input contains the battery system parameters. The renewable energy source energy supply is stored in the battery system. The battery system agent observes the battery state and energy demand of home appliances to charge/discharge energy to home appliances.

Algorithm 2 Q-Learning for Control of Battery System

Require: battery system parameters $B^{\min}, B^{\max}, \beta, \eta$.

Initialize Q^{ess} arbitrarily.

Ensure: scheduling policy for battery system derived from Q^{ess} .

- 1: **Loop** for each episode
 - 2: **For** $h = 1, 2, \dots, H$ **do**
 - 3: Store energy E_h^{res} , observe home appliance energy demand and state of battery.
 - 4: Choose action a_h^{ess} on the basis of current state s_h^{ess} using ϵ -greedy derived from Q^{ess} .
 - 5: Take action a_h^{ess} , obtain r_{h+1}^{ess} and next state s_{h+1}^{ess} .
 - 6: Calculate battery dynamics in (9).
 - 7: Update Q^{ess} using (27).
 - 8: $s_h^{\text{ess}} \leftarrow s_{h+1}^{\text{ess}}$.
 - 9: **End for**
-

Remark 1: The variation of renewable generation can affect the total energy bought from the utility. Because the electricity prices change with time, the total energy bought from the utility can affect the energy cost, which is minimized as a goal for a learning agent. Fortunately, renewable generation generally has a periodic pattern related to time; the agent can learn how to perform energy scheduling by accessing state h (time index) as well as other important states such as the current electricity price λ_h^{RTP} and battery level B_h . All that information is thus encoded as a state vector in (25) along with historical information stored in the Q-table for the agent to plan for coming activities.

V. NUMERICAL RESULTS

We examined energy scheduling in a smart home by using the proposed approach. In our analysis, the scheduling horizon starts from $h = 1$ and ends at $H = 24$, and one episode consists of 24 hours. The RTP signals were broadcast periodically by the utility and day-ahead prices were used as forecast prices. Price uncertainty resulted from the difference between the real-time and forecast prices [65]. The prediction error of prices was evaluated using the following expression of mean absolute percentage error (MAPE):

$$MAPE = \frac{100}{H} \sum_{h=1}^H \frac{|\lambda_h^{\text{RTP}} - \lambda_h^{\text{D}}|}{\lambda_h^{\text{RTP}}} \% \quad (28)$$

where λ_h^{D} cent/kWh represents the day-ahead pricing signal.

Fig. 2 illustrates sample data, including the solar energy generation and energy price, from the electric industry leader PJM [66]. By using the data set of PJM and analyzing total

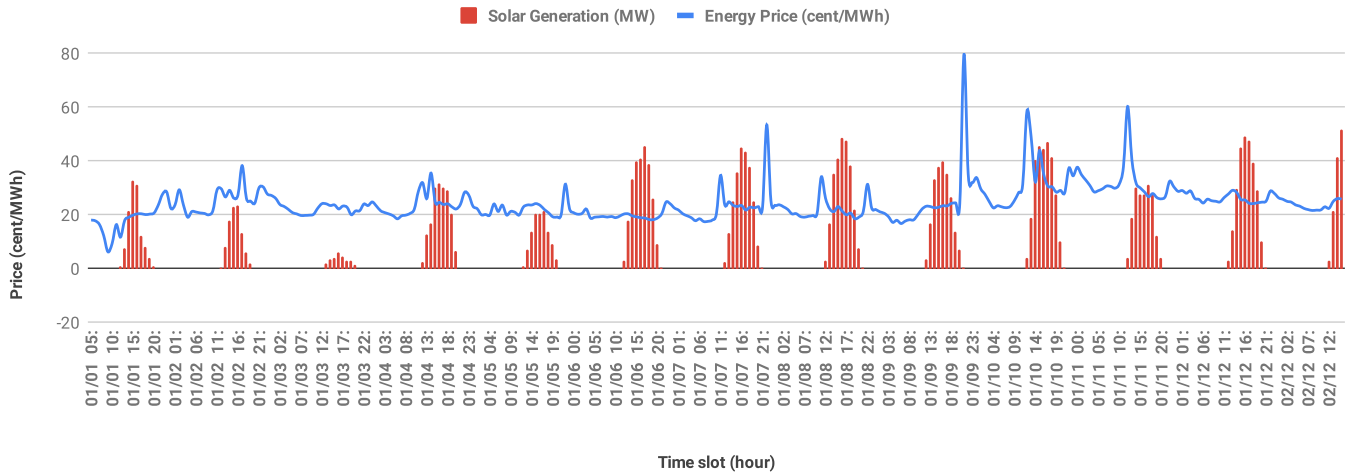


FIGURE 2. Pricing and renewable energy generation from PJM.

TABLE 1. Home appliances list.

Inflexible Appliances			
Device ID	Requested Time Range	Power	-
Refrigerator	[1, 24]	0.4 kW	-
Time-Flexible Appliances			
Device ID	Requested Time Range	Operation Duration	Power
Washing Machine	[17, 20]	3 hours	0.5 kW
Power-Flexible Appliances			
Device ID	Requested Time Range	Minimum Power	Maximum Power
AC1	[1, 24]	0 kW	1.4 kW
AC2	[1, 24]	0 kW	1.3 kW
AC3	[1, 24]	0 kW	1.5 kW

electricity prices of each month in 2019, we discovered that the highest prediction errors of prices occurred in January. The following three days were selected: January 2, 2019 (maximum MAPE), January 12, 2019 (50th percentile), and January 27, 2019 (minimum MAPE). Fig. 3 displays the RTP and day-ahead pricing of selected days. A PV system was used as a renewable energy source. Fig. 4 illustrates the corresponding PV power output. A prediction error of 10% for power generation was considered in our analysis [67].

The smart home in the study was equipped with an inflexible appliance, a time-flexible appliance, three power-flexible appliances, a battery system, and a PV system [2]. Table 1 displays the requested appliances operation list of the smart home user. The Li-Ion battery could have about 90% of the efficiency of charging/discharging and power leakage was about 0.1% to 0.2%; $\eta = 0.99$ and power leakage rate $\beta = 0.99$ in (9) were set [68]. The minimum battery level B^{\min} and maximum battery capacity B^{\max} in (8) were 1 and 12 kWh, respectively. For the learning algorithms, parameters were tuned for the best performance: $\epsilon = 0.1$, $\alpha = 0.9$ and $\gamma = 0.8$.

The proposed MORL was used to realize the user preference-based DR. Fig. 5 displays the appliances' scheduling using the proposed DR program on January 2, 2019. The energy consumption in peak price time slots was reduced. The peak electricity price occurs at $t = 13$ and all power-flexible appliances' energy consumptions were reduced in that time

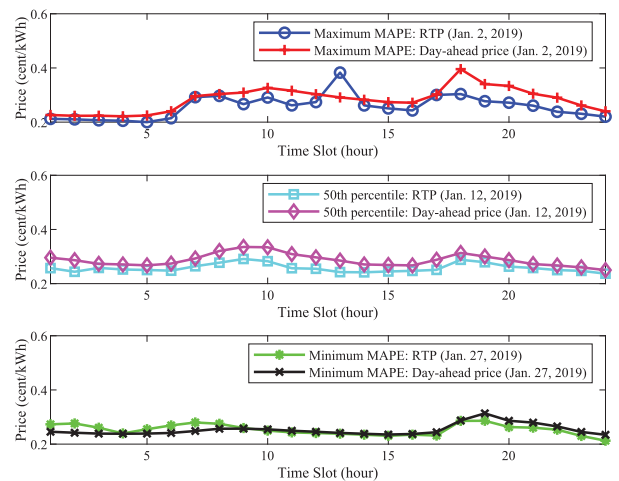


FIGURE 3. Pricing signals on three selective days.

slot. Compared with the normal operation from $t = 17$ to $t = 19$, the washing machine operated from $t = 18$ to $t = 20$, which reduced the energy bill. The proposed DR program was effective in reducing the energy bill.

The proposed approach was compared with SORL [24], normal operation, and MINLP algorithms [16]. An average of 10 simulation runs were performed to smooth all learning curves. In MINLP-based DR, prior information regarding RTP and PV supply was used for all future time slots to produce ideal performance. For MINLP-based DR without prior information, SORL-based DR, and the proposed MORL-based DR, they used forecast prices and forecast renewable energy generation for energy scheduling. For SORL, the reward function of agent n in time slot h is the weighted-sum of rewards in (22) and (23) and is expressed as follows:

$$r_{n,h} = w\lambda_h^{\text{RTP}} E_{n,h} + (1-w)U_{n,h}. \tag{29}$$

The associated Q-table was updated using (20). The normal operation was the operation of appliances according to

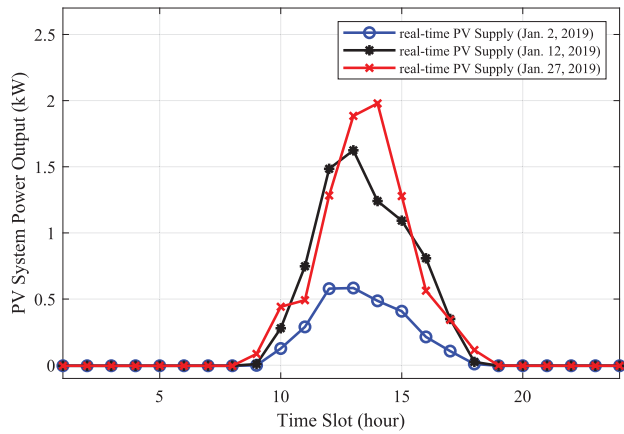


FIGURE 4. 24-hour PV supply.

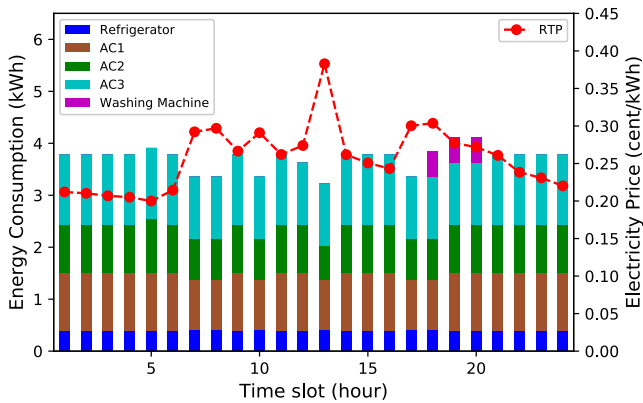
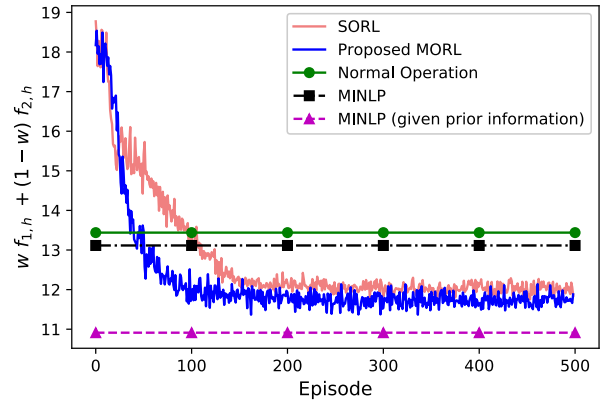


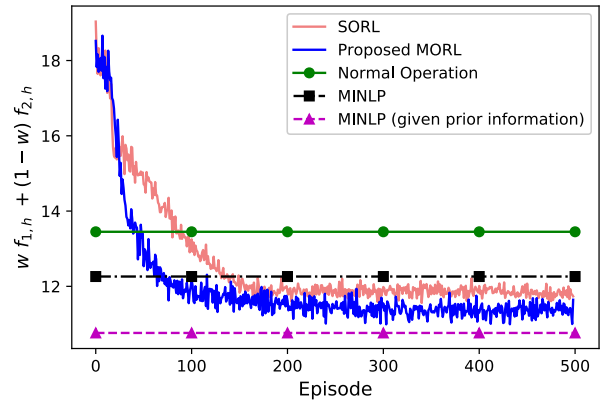
FIGURE 5. Scheduling for home appliances on January 2, 2019 ($w = 0.5$).

Table 1. MINLP-based DR used day-ahead pricing signals as the prediction of RTP signals to evaluate the energy cost for the day for energy scheduling. Fig. 6 displays the learning curves of comparable methods on the selected days. The performance of the proposed MORL-based DR was closest to the ideal performance, followed by the SORL-based DR. The MINLP-based DR program performed slightly better than the normal operation, but was susceptible to prediction errors. From Fig. 6(a) with the highest prediction error to Fig. 6(c) with the lowest prediction error, the performance of MINLP-based DR improved but was still worse than that produced by the proposed approach. As illustrated in Fig. 6(b) and Fig. 6(c), the MINLP-based DR exhibited better performance when the prediction errors were low.

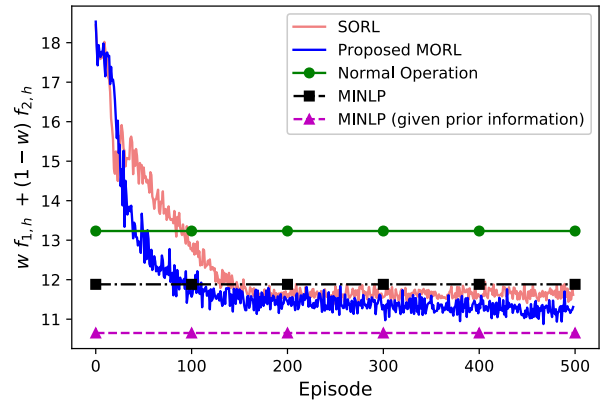
To verify the effectiveness and efficiency of the proposed approach addressing user preference changes, Fig. 7 illustrates the learning curves of the SORL-based and proposed MORL-based DRs. Two weight settings of $w = 0.2$ and $w = 0.7$ were considered. In Fig. 7, the user changed the preference at episode 300 (dotted vertical line). After the change in the user preference, the objective value achieved by the SORL-based DR required approximately 300 episodes training to converge to a steady value. By contrast, once the objective value achieved by the MORL-based DR reached a steady value, a jump to another steady objective value was



(a)



(b)



(c)

FIGURE 6. Comparison of learning curves of selected days ($w = 0.5$) with (a) maximum MAPE (2 January, 2019), (b) 50th percentile (12 January, 2019), and (c) minimum MAPE (27 January, 2019).

observed when a change in user preference occurred. This quick convergence was attributed to the use of two Q-tables, which mitigated the effect of the change of weights.

Finally, the cost and user dissatisfaction under various preference weights of the proposed MORL-based and MINLP-based DRs were compared. Table 2 presents a summary of the average electricity cost and average user dissatisfaction level in January, 2019. The normal operation had the same cost (£ 32.34) because appliances were not scheduled in response to prices. For other energy scheduling

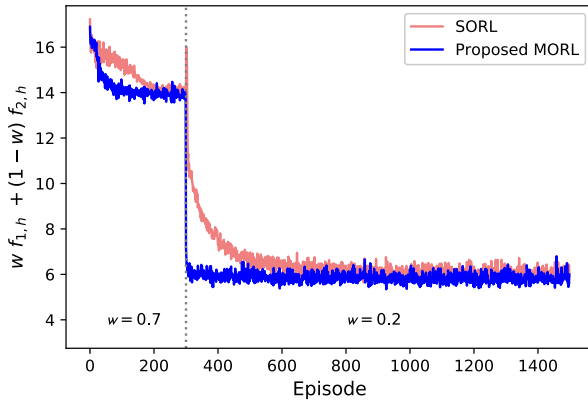


FIGURE 7. Learning curves on January 2, 2019, considering a change in user preference (episode 1 to episode 300: $w = 0.7$, episode 301 to episode 1500: $w = 0.2$). When the weight changes, the SORL-based DR undergoes additional learning processes to update the Q-table until it reaches a steady objective value. For the proposed MORL, the weight change does not affect the update function (24) using rewards (22) and (23). Thus a steady objective value is achieved.

TABLE 2. Costs and user dissatisfaction level under different user preference weights.

Average Cost			
Weight	$w = 0.3$	$w = 0.5$	$w = 0.7$
Proposed MORL based DR	¢ 23.81	¢ 19.24	¢ 11.35
MINLP based DR [16]	¢ 26.44	¢ 21.15	¢ 12.12
Normal Operation	¢ 32.34	¢ 32.34	¢ 32.34
Average User Dissatisfaction Level			
Weight	$w = 0.3$	$w = 0.5$	$w = 0.7$
Proposed MORL based DR	0.939	4.482	17.858
MINLP based DR [16]	0.903	4.441	18.001
Normal Operation	0	0	0

methods, a larger weight w yielded a lower electricity cost and higher user dissatisfaction. As compared with MINLP, cost reductions achieved by our approach were 9.95%, 9.03%, and 6.35% for $w = 0.3, 0.5$, and 0.7 , respectively (cost reduction by 8.44% on average), while the dissatisfaction level was slightly increased by 3.99%, 0.92%, and -0.79% for $w = 0.3, 0.5$, and 0.7 , respectively (dissatisfaction level increased by 1.37% on average). The proposed MORL-based DR outperformed the MINLP-based DR in the presence of prediction errors of pricing signals.

VI. CONCLUSION

This study investigated the user preference-based DR program for smart home users to optimize home appliances' scheduling by shifting or reducing energy consumption. In the literature, user preference has been assumed to be fixed without considering possible changes over time. To address possible changes in user preference, the MORL-based DR was proposed to achieve fast convergence to a steady objective value in the presence of price and renewable uncertainty. Relevant numerical analysis using real-world data, including price signals and renewable energy generation, was conducted to illustrate the advantages of the proposed approach. For SORL-based DR, an extra learning period was observed as compared with the proposed approach. For MINLP-based

DR using an optimization algorithm instead of a learning algorithm, its performance was not affected by a change in user preference, but susceptible to the uncertainty induced by electricity prices and renewable energy generation. The numerical analyses show that the proposed MORL-based DR outperformed the MINLP-based DR in terms of energy cost reductions by 8.44% while sacrificing little user dissatisfaction only by 1.37% on average. Our future work includes the investigation of Blockchain-enabled transactive energy markets for residential DR programs.

REFERENCES

- [1] R. Deng, Z. Yang, M. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 570–582, Jun. 2015.
- [2] H. Shareef, M. S. Ahmed, A. Mohamed, and E. A. Hassan, "Review on home energy management system considering demand responses, smart technologies, and intelligent controllers," *IEEE Access*, vol. 6, pp. 24498–24509, 2018.
- [3] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 152–178, Mar. 2015.
- [4] A. Etxegarai, A. Bereziartua, J. A. Dañobeitia, O. Abarrategi, and G. Saldaña, "Impact of price-based demand response programs for residential customers," in *Proc. IEEE Medit. Electrotech. Conf.*, Marrakesh, Morocco, May 2018, pp. 204–208.
- [5] S. L. Arun and M. P. Selvan, "Intelligent residential energy management system for dynamic demand response in smart buildings," *IEEE Syst. J.*, vol. 12, no. 2, pp. 1329–1340, Jun. 2018.
- [6] N. G. Paterakis, O. Erdiç, A. G. Bakirtzis, and J. P. S. Catalão, "Optimal household appliances scheduling under day-ahead pricing and load-shaping demand response strategies," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1509–1519, Dec. 2015.
- [7] K. McKenna and A. Keane, "Residential load modeling of price-based demand response for network impact studies," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2285–2294, Sep. 2016.
- [8] S. Kim and G. B. Giannakis, "An online convex optimization approach to real-time energy pricing for demand response," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2784–2793, Nov. 2017.
- [9] Z. Zhang, F. Li, and H. Shi, "A pricing strategy reflecting the cost of power volatility to facilitate decentralized demand response," *IEEE Access*, vol. 7, pp. 105863–105871, 2019.
- [10] Z. Zhao, W. C. Lee, Y. Shin, and K.-B. Song, "An optimal power scheduling method for demand response in home energy management system," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1391–1400, Sep. 2013.
- [11] M. Shafie-Khah and P. Siano, "A stochastic home energy management system considering satisfaction cost and response fatigue," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 629–638, Feb. 2018.
- [12] V. Pradhan, V. S. K. Balijepalli, and S. A. Khaparde, "An effective model for demand response management systems of residential electricity consumers," *IEEE Syst. J.*, vol. 10, no. 2, pp. 434–445, Jun. 2016.
- [13] S. Althaher, P. Mancarella, and J. Mutale, "Automated demand response from home energy management system under dynamic pricing and power and comfort constraints," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1874–1883, Jul. 2015.
- [14] A. Heidari, Z. Y. Dong, D. Zhang, P. Siano, and J. Aghaei, "Mixed-integer nonlinear programming formulation for distribution networks reliability optimization," *IEEE Trans. Ind. Informat.*, vol. 14, no. 5, pp. 1952–1961, May 2018.
- [15] M. Alipour, K. Zare, and M. Abapour, "MINLP probabilistic scheduling model for demand response programs integrated energy hubs," *IEEE Trans. Ind. Informat.*, vol. 14, no. 1, pp. 79–88, Jan. 2018.
- [16] S. Moon and J.-W. Lee, "Multi-residential demand response scheduling with multi-class appliances in smart grid," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2518–2528, Jul. 2018.
- [17] L. Bhamidi and S. Sivasubramani, "Optimal planning and operational strategy of a residential microgrid with demand side management," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2624–2632, Jun. 2020.

- [18] P. Kou, D. Liang, and L. Gao, "Stochastic energy scheduling in microgrids considering the uncertainties in both supply and demand," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2589–2600, Sep. 2018.
- [19] T. Pinto, R. Faia, M. Navarro-Caceres, G. Santos, J. M. Corchado, and Z. Vale, "Multi-agent-based CBR recommender system for intelligent energy management in buildings," *IEEE Syst. J.*, vol. 13, no. 1, pp. 1084–1095, Mar. 2019.
- [20] Y. Yang, R. Luo, M. Li, M. Zhou, W. Zhang, and J. Wang, "Mean field multi-agent reinforcement learning," in *Proc. 35th ICML*, Stockholm, Sweden, vol. 80, Jul. 2018, pp. 5571–5580.
- [21] E. C. Kara, M. Berges, B. Krogh, and S. Kar, "Using smart devices for system-level management and control in the smart grid: A reinforcement learning framework," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Tainan, Taiwan, Nov. 2012, pp. 85–90.
- [22] F. Ruelens, B. J. Claessens, S. Vandael, B. De Schutter, R. Babuska, and R. Belmans, "Residential demand response of thermostatically controlled loads using batch reinforcement learning," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2149–2159, Sep. 2017.
- [23] Z. Wen, D. O'Neill, and H. Maei, "Optimal demand response using device-based reinforcement learning," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2312–2324, Sep. 2015.
- [24] R. Lu, S. H. Hong, and M. Yu, "Demand response for home energy management using reinforcement learning and artificial neural network," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6629–6639, Nov. 2019.
- [25] X. Xu, Y. Jia, Y. Xu, Z. Xu, S. Chai, and C. S. Lai, "A multi-agent reinforcement learning-based data-driven method for home energy management," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3201–3211, Jul. 2020.
- [26] T. Remani, E. A. Jasmin, and T. P. I. Ahamed, "Residential load scheduling with renewable generation in the smart grid: A reinforcement learning approach," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3283–3294, Sep. 2019.
- [27] A. Mathew, A. Roy, and J. Mathew, "Intelligent residential energy management system using deep reinforcement learning," *IEEE Syst. J.*, vol. 14, no. 4, pp. 5362–5372, Dec. 2020.
- [28] L. Park, Y. Jang, S. Cho, and J. Kim, "Residential demand response for renewable energy resources in smart grid systems," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3165–3173, Dec. 2017.
- [29] C. Liu, X. Xu, and D. Hu, "Multiobjective reinforcement learning: A comprehensive overview," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 45, no. 3, pp. 385–398, Mar. 2015.
- [30] V. Pilloni, A. Floris, A. Meloni, and L. Atzori, "Smart home energy management including renewable sources: A QoE-driven approach," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 2006–2018, May 2018.
- [31] Y. Liu and S. Hu, "Renewable energy pricing driven scheduling in distributed smart community systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 5, pp. 1445–1456, May 2017.
- [32] W. Yi, Y. Zhang, Z. Zhao, and Y. Huang, "Multiobjective robust scheduling for smart distribution grids: Considering renewable energy and demand response uncertainty," *IEEE Access*, vol. 6, pp. 45715–45724, 2018.
- [33] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Trans. Ind. Appl.*, vol. 48, no. 3, pp. 1064–1069, May/June 2012.
- [34] H. Sangrody, M. Sarailoo, N. Zhou, N. Tran, M. Motalleb, and E. Foruzan, "Weather forecasting error in solar energy forecasting," *IET Renew. Power Gener.*, vol. 11, no. 10, pp. 1274–1280, Aug. 2017.
- [35] L. An, J. Duan, M.-Y. Chow, and A. Duel-Hallen, "A distributed and resilient bargaining game for weather-predictive microgrid energy cooperation," *IEEE Trans. Ind. Informat.*, vol. 15, no. 8, pp. 4721–4730, Aug. 2019.
- [36] D. Cerrai, D. W. Wanik, M. A. E. Bhuiyan, X. Zhang, J. Yang, M. E. B. Frediani, and E. N. Anagnostou, "Predicting storm outages through new representations of weather and vegetation," *IEEE Access*, vol. 7, pp. 29639–29654, 2019.
- [37] D. M. Rosewater, D. A. Copp, T. A. Nguyen, R. H. Byrne, and S. Santoso, "Battery energy storage models for optimal control," *IEEE Access*, vol. 7, pp. 178357–178391, 2019.
- [38] M. T. Lawder, B. Suthar, P. W. Northrop, S. De, C. M. Hoff, O. Leitermann, M. L. Crow, S. Santhanagopalan, and V. R. Subramanian, "Battery energy storage system (BESS) and battery management system (BMS) for grid-scale applications," *Proc. IEEE*, vol. 102, no. 6, pp. 1014–1030, Jun. 2014.
- [39] Y. Wang, X. Lin, and M. Pedram, "Adaptive control for energy storage systems in households with photovoltaic modules," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 992–1001, Mar. 2014.
- [40] H. Chitsaz, P. Zamani-Dehkordi, H. Zareipour, and P. P. Parikh, "Electricity price forecasting for operational scheduling of behind-the-meter storage systems," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6612–6622, Nov. 2018.
- [41] M. Ross, C. Abbey, F. Bouffard, and G. Joos, "Microgrid economic dispatch with energy storage systems," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3039–3047, Jul. 2018.
- [42] Y. Wang, X. Lin, and M. Pedram, "A near-optimal model-based control algorithm for households equipped with residential photovoltaic power generation and energy storage systems," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 77–86, Jan. 2016.
- [43] B.-R. Ke, T.-T. Ku, Y.-L. Ke, C.-Y. Chuang, and H.-Z. Chen, "Sizing the battery energy storage system on a university campus with prediction of load and photovoltaic generation," *IEEE Trans. Ind. Appl.*, vol. 52, no. 2, pp. 1136–1147, Apr. 2016.
- [44] A. Fukushima, T. Yano, S. Imahara, H. Aisu, Y. Shimokawa, and Y. Shibata, "Prediction of energy consumption for new electric vehicle models by machine learning," *IET Intell. Transp. Syst.*, vol. 12, no. 9, pp. 1174–1180, Nov. 2018.
- [45] Z. D. Asher, D. A. Baker, and T. H. Bradley, "Prediction error applied to hybrid electric vehicle optimal fuel economy," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 6, pp. 2121–2134, Nov. 2017.
- [46] K. Mahmud, J. Ravishanker, M. J. Hossain, and Z. Y. Dong, "The impact of prediction errors in the domestic peak power demand management," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4567–4579, Jul. 2020.
- [47] M. Jia, F. Li, Y. Pu, and Z. Chen, "A lossless electrocardiogram compression system based on dual-mode prediction and error modeling," *IEEE Access*, vol. 8, pp. 101153–101162, 2020.
- [48] D. Kodaira, W. Jung, and S. Han, "Optimal energy storage system operation for peak reduction in a distribution network using a prediction interval," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2208–2217, May 2020.
- [49] X. Lee, M. Yan, F. Y. Xu, Y. Wang, Y. Fan, Z. Lee, Y. Wen, M. Shahidepour, and L. L. Lai, "Virtual storage-based DSM with error-driven prediction modulation for microgrids," *IEEE Access*, vol. 7, pp. 71109–71118, 2019.
- [50] R. Singh, R. C. Bansal, A. R. Singh, and R. Naidoo, "Multi-objective optimization of hybrid renewable energy system using reformed electric system cascade analysis for islanding and grid connected modes of operation," *IEEE Access*, vol. 6, pp. 47332–47354, 2018.
- [51] B. S. Sami, "Intelligent energy management for off-grid renewable hybrid system using multi-agent approach," *IEEE Access*, vol. 8, pp. 8681–8696, 2020.
- [52] C.-T. Tsai, T. M. Beza, E. M. Molla, and C.-C. Kuo, "Analysis and sizing of mini-grid hybrid renewable energy system for islands," *IEEE Access*, vol. 8, pp. 70013–70029, 2020.
- [53] K. Li, P. Zhang, G. Li, F. Wang, Z. Mi, and H. Chen, "Day-ahead optimal joint scheduling model of electric and natural gas appliances for home integrated energy management," *IEEE Access*, vol. 7, pp. 133628–133640, 2019.
- [54] Y. Zhao, W. Wang, C. Xiang, H. Liu, and R. Langari, "Research and bench test of nonlinear model predictive control-based power allocation strategy for hybrid energy storage system," *IEEE Access*, vol. 6, pp. 70770–70787, 2018.
- [55] M. M. Iqbal, M. F. Zia, K. Beddiar, and M. Benbouzid, "Optimal scheduling of grid transactive home demand responsive appliances using polar bear optimization algorithm," *IEEE Access*, vol. 8, pp. 222285–222296, 2020.
- [56] F. S. Ali, O. Bouachir, O. Özkasap, and M. Aloqaily, "SynergyChain: Blockchain-assisted adaptive cyber-physical P2P energy trading," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5769–5778, Aug. 2021.
- [57] O. Bouachir, M. Aloqaily, L. Tseng, and A. Boukerche, "Blockchain and fog computing for cyberphysical systems: The case of smart industry," *Computer*, vol. 53, no. 9, pp. 36–45, Sep. 2020.
- [58] M. Aloqaily, A. Boukerche, O. Bouachir, F. Khalid, and S. Jangsher, "An energy trade framework using smart contracts: Overview and challenges," *IEEE New.*, vol. 34, no. 4, pp. 119–125, Jul. 2020.
- [59] U.S. Energy Information Administration. (Apr. 2021). *Electricity Explained: Factors Affecting Electricity Prices*. Accessed: Oct. 28, 2021. [Online]. Available: <https://www.eia.gov/energyexplained/electricity/prices-and-factors-affecting-prices.php>
- [60] F. Alfaverh, M. Denai, and Y. Sun, "Demand response strategy based on reinforcement learning and fuzzy reasoning for home energy management," *IEEE Access*, vol. 8, pp. 39310–39321, 2020.

- [61] K. Ma, T. Yao, J. Yang, and X. Guan, "Residential power scheduling for demand response in smart grid," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 320–325, Jun. 2016.
- [62] H.-H. Chang, W.-Y. Chiu, H. Sun, and C.-M. Chen, "User-centric multi-objective approach to privacy preservation and energy cost minimization in smart home," *IEEE Syst. J.*, vol. 13, no. 1, pp. 1030–1041, Mar. 2019.
- [63] F. Luo, W. Kong, G. Ranzi, and Z. Y. Dong, "Optimal home energy management system with demand charge tariff and appliance operational dependencies," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 4–14, Jan. 2020.
- [64] C. J. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 30, no. 4, pp. 279–292, 1992.
- [65] X. Huang, S. H. Hong, and Y. Li, "Hour-ahead price based energy management scheme for industrial facilities," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2886–2898, Dec. 2017.
- [66] *PJM Interconnection*. Accessed: Oct. 28, 2021. [Online]. Available: <http://www.pjm.com>
- [67] F. Wang, Z. Mi, S. Su, and H. Zhao, "Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters," *Energies*, vol. 5, no. 5, pp. 1355–1370, May 2012.
- [68] H. Sun, Z. Fu, F. Tao, L. Zhu, and P. Si, "Data-driven reinforcement-learning-based hierarchical energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles," *J. Power Sources*, vol. 455, pp. 1–12, Apr. 2020.



SONG-JEN CHEN received the B.S. degree in mechanical engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2018, and the M.S. degree in electrical engineer from National Tsing Hua University, Hsinchu, in 2020. His research interests include reinforcement learning algorithms and smart grids.



WEI-YU CHIU (Member, IEEE) received the Ph.D. degree in communications engineering from National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2010. He is currently an Associate Professor of electrical engineering with NTHU. His research interests include multiobjective optimization and reinforcement learning, and their applications to control systems, robotics, and smart energy systems. He was a recipient of the Youth Automatic Control Engineering Award bestowed by Chinese Automatic Control Society, in 2016, the Outstanding Young Scholar Academic Award bestowed by Taiwan Association of Systems Science and Engineering, in 2017, the Erasmus+ Program Fellowship funded by European Union (staff mobility for teaching), in 2018, and the Outstanding Youth Electrical Engineer Award bestowed by the Chinese Institute of Electrical Engineering, in 2020. From 2015 to 2018, he was an Organizer or the Chair of the International Workshop on Integrating Communications, Control, and Computing Technologies for Smart Grid (ICT4SG). He is a Subject Editor for *IET Smart Grid*.



WEI-JEN LIU received the B.S. degree in electronic engineering from the National Taipei University of Technology, Taipei, Taiwan, in 2020. He is currently pursuing the master's degree in electronic engineering with National Tsing Hua University, Hsinchu, Taiwan. He research interests include blockchain technologies, network security, quantitative trading, and renewable energy.

...