Energy Conversion and Management 103 (2015) 553–561

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Action dependent heuristic dynamic programming based residential energy scheduling with home energy inter-exchange

Yancai Xu, Derong Liu, Qinglai Wei *

The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 3 February 2015
Accepted 17 June 2015
Available online 11 July 2015

Keywords:
Action dependent heuristic dynamic programming
Residential energy management
Energy scheduling
Control sequence
Control strategy

ABSTRACT

Residential energy scheduling is a hot topic nowadays in the background of energy saving and environmental protection worldwide. To achieve this objective, a new residential energy scheduling algorithm is developed for energy management, based on action dependent heuristic dynamic programming. The algorithm works under the circumstance of residential real-time pricing and two adjacent housing units with energy inter-exchange, which can reduce the overall cost and enhance renewable energy efficiency after long-term operation. It is designed to obtain the optimal control policy to manage the directions and amounts of electricity energy flux. The algorithm's architecture is mainly constructed based on neural networks, denoting the learned characteristics in the linkage of layers. To get close to real situations, many constraints such as maximum charging/discharging power of batteries are taken into account. The absent energy penalty cost is developed for the first time as a part of the performance index function. When the environment changes, the residential energy scheduling algorithm gains new features and keeps adapting in real-time operations. Simulation results show that the developed algorithm is beneficial to energy conversation.

1. Introduction

In the past several decades, the energy crisis has been plaguing humanity with the rising price of oil and other fossil fuels. Furthermore, traditional fossil fuels have limited reserves, which becomes the bottleneck for the economic development of a vast majority of countries in the world [1]. More and more nations are taking actions to promote the use of renewable resources and energy management. European Commission proposed the smart grid vision for the aim of encouraging manufacturers to produce smart appliances to prevent power blackouts by isolating disturbances in the grid, reduce power usage peaks and lower total energy usage in 2006 [2]. A detailed research agenda was given for guiding the development of Europe's future smart grid networks [3]. American government proposed the development of smart gird to create pilot projects to promote the manufacturing of advanced metering devices so as to reduce energy cost [4]. Some countries constructed demonstration projects and gained many research achievements.

Besides government promotions, intelligent power grid has attracted many researchers' attention. The research fields vary from the communication standards to communication security, from unidirectional energy transmission to bidirectional transmission, from centralized adjustment of electrical power system to decentralized adjustment, from smart community with renewable resources to smart community with plugin vehicles. Zakariazadeh et al. proposed a stochastic multi-objective mathematical scheduling method for a smart grid distribution system [5]. Figueiredo and Martins integrated the building's renewable energy production into the demand-side energy management [6]. Lin and Chen studied how to solve optimal power flow problems in a smart grid transmission system [7]. Qureshi et al. focused on the impact of energy storage in buildings [8]. While Huang and Liu studied on the energy end-user side [9]. Hannan et al. showed how light electric vehicles can achieve the multi-resources control [10]. Wei et al. proposed a novel dual iterative *Q*-learning method for energy control [11]. Most of the countries attach great importance to the development of smart grid and take it as a strategy to occupy a commanding height in industry. The objective that all scientists pursue, is a robust system, inside of which the energy is transmitted unidirectionally on the generators' side and bidirectionally on the users' side. Thus, many power failures such as blackouts can be avoided.

* Corresponding author.
E-mail addresses: yancai.xu@ia.ac.cn (Y. Xu), derong.liu@ia.ac.cn (D. Liu), qinglai.wei@ia.ac.cn (Q. Wei).

http://dx.doi.org/10.1016/j.enconman.2015.06.048
0196-8904/

With the development of smart grid, many intelligent algorithms have been proposed in the last decades and some have come into use. Adaptive dynamic programming (ADP) is an efficient way to solve this problem, which was developed by Werbos in the last century [12]. Great efforts have been made to the research of ADP ever after. Prokhorov and Wunsch proposed adaptive critic designs in 1997 [13]. Liu et al. proposed action-dependent adaptive critic designs in 2001 [14]. Later, Liu and Wei utilized ADP to solve the multi-person zero-sum differential games [15]. Wei et al. achieved dual iterative adaptive dynamic programming for time-delay discrete-time nonlinear systems [16]. Bhasin et al. achieved a reinforcement learning-based adaptive critic controller for tracking problems [17]. Generally, there are three frameworks of ADP: heuristic dynamic programming (HDP), dual heuristic programming (DHP), and globalized dual heuristic programming (GDHP). For a linear quadratic regulator problem, it often falls to solve a Riccati equation. On the other hand, for a nonlinear system, its optimal state feedback control depends on the Hamilton–Jacobi–Bellman (HJB) equation [18]. Dynamic programming (DP) is widely used to tackle optimal control problems for nonlinear systems. However, the implementation of DP backward-in-time makes the computation expensive for nonlinear systems, which is the "curse of dimensionality" [19]. To overcome this difficulty, a "critic network" is built to approximate the solution and avoid the computationally awkward curse [20].

The research of integrating ADP and smart grid is quite hot recently. Mohagheghi et al. used ADHDP for the design of a static compensator [21]. Lu et al. applied the HDP method to a large power system stability control [22]. Fuselli et al. achieved home energy management with ADP [23]. Besides ADP, other intelligent algorithms including genetic algorithm (GA), ant colony algorithm, particle swarm optimization (PSO) approach, simulated annealing method, and fuzzy optimization, can also be employed to smart grid. Keçebaş et al. achieved artificial neural network modeling for energy analysis [24]. Sousa et al. used annealing approach for energy resource management [25]. A PSO-GA approach was used in [26] by Yu et al. for energy demand projection. Then, a PSO-based fuzzy logic controller was developed by Safari et al. for green power energy system [27]. Fuzzy logic controller was also applied in the microgrid battery management [28]. Traditional solutions like mixed integer linear programming (MILP), can also be used to solve energy management problem [29]. One concrete residential microgrid building came into use in Italy to improve the performance of energy usage [30].

In this paper, two-household energy management is incorporated into the smart grid residential terminals' applications for end-users. The ultimate goal consists of managing the usage of all energy resources in an optimal way [31]. The absent energy penalty cost is developed for the first time within the performance index function. This paper tackles the problem of optimal energy management and cost-saving within two housing units. The action dependent adaptive dynamic programming (ADHDP) based residential energy scheduling algorithm has the property of adapting to the frequent and swift changing systems based on the algorithm's own characteristics; meanwhile, it avoids other algorithms' shortcomings, such as the inability to adapt to dynamical and rapid changing environments.

The organization of this paper is as follows. In Section 2, the overall structure of the residential management system is given and each component is briefly described. In Section 3, the ADHDP principle is introduced. Section 4 focuses on the detailed residential energy control strategy with energy inter-exchange. Section 5 presents the experimental simulations. Finally, in Section 6 the paper concludes with a few remarks and future work.

2. Residential energy management

2.1. Notation

The used notations are listed as follows.

Parameters:	
Δt	Time interval
$C(t)$	Residential real-time price at time t (cents)
$C_p(t)$	Residential penalty price at time t (cents)
$E_{b1}(t)$	Remaining capacity of battery 1 at time t (kW h)
E_{b1}^{max}	Maximum capacity of battery 1 (given at 25 °C) (kW h)
E_{b1}^{min}	Minimum remaining capacity of battery 1 (given at 25 °C) (kW h)
$P_{load1}(t)$	Residential load of housing unit 1 at time t (kW)
$P_{load2}(t)$	Residential load of housing unit 2 at time t (kW)
$P_{grid}(t)$	Power from the utility gird (kW)
$P_{b1}(t)$	Charging/discharging power of battery 1 at time t (kW)
$P_{b2}(t)$	Charging/discharging power of battery 2 at time t (kW)
P_{b1}^{max}	Positive, maximum discharging power of battery 1 (kW)
P_{b2}^{max}	Positive, maximum discharging power of battery 2 (kW)
P_{b1}^{min}	Negative, maximum charging power of battery 1 (kW)
P_{b2}^{min}	Negative, maximum charging power of battery 2 (kW)
P_{rate1}	Rated output power of the battery 1 (kW)
$\eta(P_{b1}(t))$	Total efficiency in the battery 1 storage system, considering all auxiliary devices
γ	Discount factor, $0 < \gamma \le 1$
$u_1(t)$	Initial control action of battery 1 at time $t,\ u_1(t) \in \{1, 0, -1, -2, -3\}$
$u_2(t)$	Initial control action of battery 2 at time $t,\ u_2(t) \in \{1, 0, -1, -2, -3\}$
$\hat{u}_1(t)$	Adjusted control action of battery 1 at time $t,\ u_1(t) \in \{1, 0, -1, -2, -3\}$
$\hat{u}_2(t)$	Adjusted control action of battery 2 at time $t,\ u_2(t) \in \{1, 0, -1, -2, -3\}$
$C_{Ab1}(t)$	Absent capacity of battery 1 at time t (kW h)
$C_{Ab2}(t)$	Absent capacity of battery 2 at time t (kW h)
$x(\cdot)$	State vector of the system, $x(\cdot) \in \mathbb{R}^n$
$u(\cdot)$	Control action, $u(\cdot) \in \mathbb{R}^n$
$F(\cdot)$	System transition function
$U(\cdot)$	Utility function
$J(\cdot)$	Performance index/cost
$u^*(\cdot)$	Optimal control
$J^*(\cdot)$	Optimal performance index/cost
$Q(\cdot)$	Output of the critic network

2.2. Components of residential energy management

In this paper, the residential energy system with home energy inter-exchange is composed of a public utility grid, two housing units with storage systems and power management units (PMU). The PMU plays an important role in energy conversion and management. Fig. 1 shows the power flows of residential energy scheduling, in which the AC utility grid is the primary source of electric energy. Each of the two housing units is equipped with a storage system whose capacity can be the same as or different from the other. Inside each housing unit, there are some household

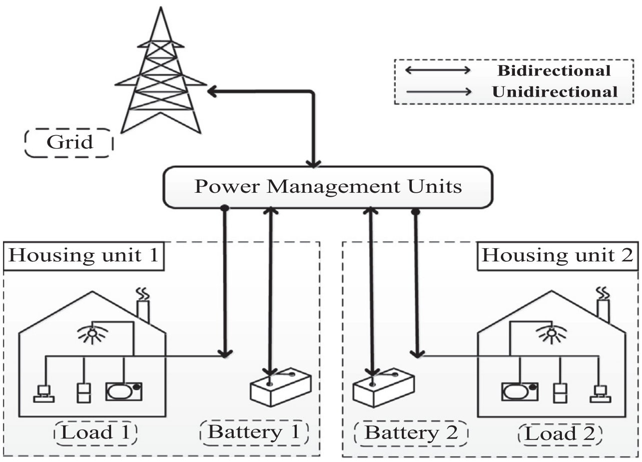

Fig. 1. Power flows of the residential energy scheduling with home energy interexchange.

electric appliances such as televisions, refrigerators and lamps. The PMU allows electric power from different sources, which include battery 1, battery 2 and the utility grid, to be pooled and distributed over the housing units. During this process, the power flows can be unidirectional or bidirectional depending on actual conditions. The utility grid, battery 1 and battery 2 can either supply/get electric power to/from the PMU separately. As for residential loads 1 and 2, they only consume electricity, transmitted via the PMU.

One of the major roles of the PMU is about electric power conversion. With thyristor and MOSFET devices, engineers have designed pulse width modulation (PWM) switch to achieve AC&DC reversible conversion [32]. Batteries can only handle direct current (DC) whenever charging from/discharging to the PMU. The storage systems are mainly made up of lead–acid batteries in consideration of economy cost factor.

2.3. Residential real-time pricing

Residential real-time pricing is a widely used pricing strategy in many countries' electric power industry. As a load management policy, it pushes the end-users to shift electricity use from peak load hours to light load hours [33]. The policy greatly improves power system efficiency and allows new power system construction projects [34]. The hourly change of wholesale electricity price is mainly influenced by consumer demand. High demand for electricity usually means high electric price. Low demand means low price.

2.4. Load profile

Though the residential load is a continuous time process, discretization can greatly simplify the computational model and save the computation time. The residential real-time price (RRTP) is usually published with one-hour step. Therefore, the residential load profile is calculated hourly for convenience [35]. Fig. 3 indicates a typical residential load profile [36]. There are likely to have a small spike in the morning and a big one in the evening. The two spikes correspond to the RRTP curve in Fig. 2, indicating that the residential load plays an important role in the pricing strategy.

2.5. Model of batteries

2.5.1. Battery energy

Energy loss exists in the process of AC → DC and DC → AC conversions, which could not be eliminated. The efficiency equation of these devices in the battery model was derived by Yau et al. [37] as

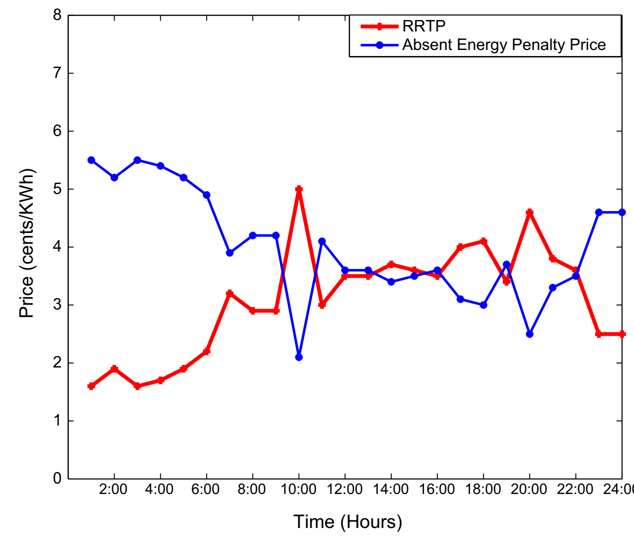

Fig. 2. A daily residential real-time hourly prices.

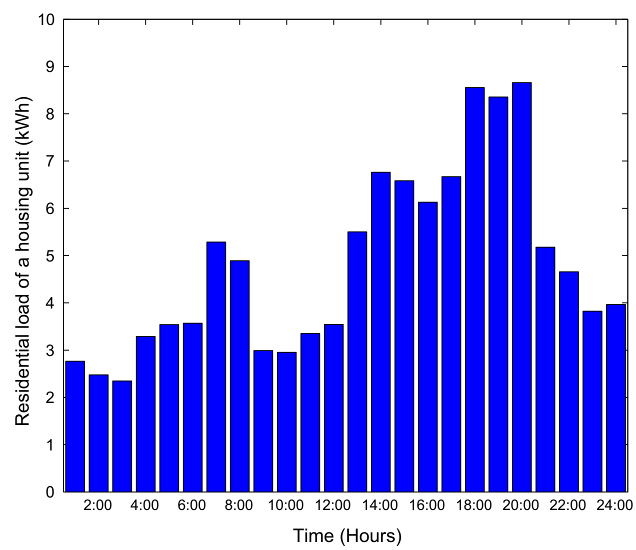

Fig. 3. A typical daily residential load profile.

$$\eta(P_{b1}(t)) = 0.898 - 0.173|P_{b1}(t)|/P_{rate1}, \quad P_{rate1} > 0. \tag{1}$$

Suppose the efficiencies of charging and discharging are the same, calculated by

$$E_{b1}(t+1) = E_{b1}(t) - P_{b1}(t+1)\Delta t \cdot \eta(P_{b1}(t+1)), \quad P_{b1}(t+1) < 0, \tag{2}$$

$$E_{b1}(t+1) = E_{b1}(t) - P_{b1}(t+1)\Delta t/\eta(P_{b1}(t+1)), \quad P_{b1}(t+1) > 0. \tag{3}$$

It is worth pointing out that the battery output power is divided by the efficiency $\eta(P_{b1}(t+1))$ during the time interval $t+1$. Given the fact that the calculating time interval is one hour, the time interval Δt can be omitted.

2.5.2. Battery constraints

In practical applications, the batteries need to be handled carefully to avoid adverse conditions such as overcharge, overdischarge or overcurrent.

a. Upper and lower energy bounds for batteries

$$E_{b1}^{min} \leqslant E_{b1}(t) \leqslant E_{b1}^{max}. \quad (4)$$

b. Maximum charge/discharge power limits

$$P_{b1}^{min} \leqslant P_{b1}(t) \leqslant P_{b1}^{max}. \quad (5)$$

c. Residential load balance. All the energy provided by the AC utility grid, batteries 1 and 2, must meet the fluctuating demand of appliances in housing units 1 and 2.

$$P_{load1}(t) + P_{load2}(t) = P_{grid}(t) + P_{b1}(t) + P_{b2}(t). \quad (6)$$

3. Action dependent heuristic dynamic programming

In discrete-time nonlinear systems, ADHDP employs the Bellman principle of optimality to solve complex optimization problems [38]. The Bellman equation describes the relationship between the value function of one period and that of the next period [39]. Suppose the discrete-time nonlinear dynamic system as below:

$$x(k+1) = F(x(k), u(k), k), \quad k = 0, 1, \ldots. \quad (7)$$

With transition function $F(\cdot)$ and control action $u(k)$, the next state $x(k+1)$ can be obtained. Assume the associated performance index/cost function as

$$J(x(i), i) = \sum_{k=i}^{\infty} \gamma^{k-i} U(x(k), u(k), k). \quad (8)$$

The solution is to minimize the function $J(x(i), i)$ through choosing an appropriate control sequence $u(k), k = i, i+1, \ldots$. According to the Bellman principle of optimality, the optimal control sequence $u^*(k)$ at time k can be expressed as

$$u^*(k) = \arg\min_{u(k)} \{U(x(k), u(k)) + \gamma J^*(x(k+1))\}. \quad (9)$$

ADHDP, as one framework of ADP, is an effective method to solve dynamic programming problems with discrete formulations by using policy evaluation and policy improvement. The neural network based adaptive critic approach is used to solve ADHDP formulations [9,38]. The residential energy scheduling has already been discretized with the real-time pricing at a one-hour interval. The ADHDP architecture of residential energy scheduling used in this paper is shown in Fig. 4.

4. Residential energy scheduling strategy

Residential energy scheduling is designed based on ADHDP with home energy inter-exchange. This section introduces an energy scheduling strategy to boost residential energy efficiency and promote energy conservation.

4.1. States of storage system (batteries)

There are five kinds of states for each battery inside the storage systems. Furthermore, the states of batteries 1 and 2 must not be in conflict with each other. Besides, each battery is primarily responsible for the appliances in its own side housing unit. Take the states of batteries 1 and 2 as u_1 and u_2.

1. Charging state $(u_1, u_2 = 1)$
   The storage systems charge the batteries as $u_1,\ u_2 = 1$ usually in the night when RRTP and batteries' remaining capacities are low.
2. Idle state $(u_1, u_2 = 0)$
   Idle state appears when charging or discharging is not quite so necessary. Idle state can decrease the usage frequency of batteries. The AC utility grid provides all the required energy when $u_1,\ u_2 = 0$.
3. Discharging state A $(u_1, u_2 = -1)$
   Discharging state A means that each battery can only discharge to the appliances inside its own housing unit. It appears when RRTP is high and/or the other battery energy is low.
4. Discharging state B $(u_1, u_2 = -2)$
   Discharging state B is opposite to discharging state A, in which the batteries are forced to output electric energy to appliances in the other housing unit.
5. Discharging state C $(u_1, u_2 = -3)$
   When RRTP is high and one battery has low remaining capacity, the other battery is likely in discharging state C. Thus, the battery outputs electric energy to appliances in both housing units at that time period.

One important thing to note is that one battery's state cannot affect the other's. If battery 1 is in one state, then battery 2 can be in any state except which may cause potential conflicts, and vice versa.

4.2. Batteries' control strategy

The residential energy scheduling is a discrete problem here. Thus, the batteries have discrete processes. Two batteries correspond to two housing units. Battery 1 is primarily responsible for the energy supply of housing unit 1. The detailed energy-related components within the energy storage system are listed in Table 1. Batteries 1 and 2 are the bond which connects both the

Table 1
Energy-related components within the energy storage system.

PMU components	AC utility grid
	Battery 1
	Battery 2
	Household load 1
	Household load 2

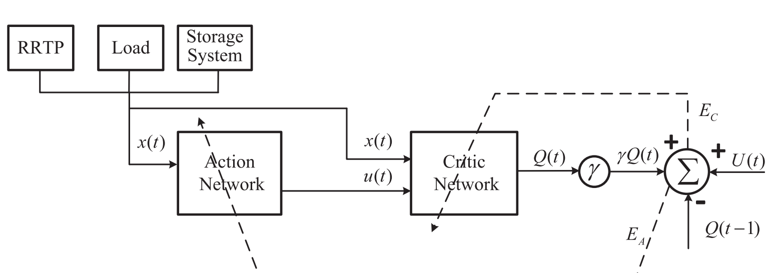

Fig. 4. ADHDP architecture of residential energy scheduling.

Table 2
Control strategy of batteries 1 and 2.

Battery state	Corresponding control action $u_1(t)$ & $u_2(t)$
Charging state	1
Idle state	0
Discharging state A	−1
Discharging state B	−2
Discharging state C	−3

Table 3
Batteries' control scheme.

Control scheme	$u_1(t)$	$u_2(t)$	Battery 1	Battery 2
Feasible control action	1	1	Charging state	Charging state
	1	0	Charging state	Idle state
	1	−1	Charging state	Discharging state A
	1	−2	Charging state	Discharging state B
	1	−3	Charging state	Discharging state C
	0	1	Idle state	Charging state
	0	0	Idle state	Idle state
	0	−1	Idle state	Discharging state A
	0	−2	Idle state	Discharging state B
	0	−3	Idle state	Discharging state C
	−1	1	Discharging state A	Charging state
	−1	0	Discharging state A	Idle state
	−1	−1	Discharging state A	Discharging state A
	−2	1	Discharging state B	Charging state
	−2	0	Discharging state B	Idle state
	−2	−2	Discharging state B	Discharging state B
	−3	1	Discharging state C	Charging state
	−3	0	Discharging state C	Idle state

household loads and the outside AC utility grid. For each battery, detailed states are shown in Table 2 with their corresponding control commands. The control action $u_1(t)$, $u_2(t)$ can only choose from the five fixed integer numbers with specialized control meanings, i.e., $u_1(t), u_2(t) \in \{1, 0, -1, -2, -3\}$. However, the two batteries are likely to interact with each other, so the control strategy must be designed carefully to avoid conflicts. Table 3 has shown all the feasible matched control strategies when controlling both batteries at the same time. As we have assumed that one household load can only be supplied by one energy resource at the same time, it can be seen that there are 7 infeasible matches. Conflicts occur when infeasible combinations are chosen. Future work will cover the multi-resource problem.

Some potential conflicts still happen sometimes. The control strategies are given in advance, without considering the batteries' actual situation. Therefore, the batteries' control scheme has to be adjusted to meet actual needs afterwards. If there is any clash between $u_1(t)$ and $u_2(t)$, $u_1(t)$ has the default priority over $u_2(t)$. Tables 4 and 5 show the adjustment data of control action $u_1(t)$ in different situations. When the initial control action given as $u_1(t) = 1$, it is necessary to identify the remaining capacity of battery 1. If battery 1 is full, there is no need to charge just as the adjusted control action $\hat{u}_1(t) = 0$. If $C_{\mathrm{Ab1}}(t)$ is less than $|P_{\mathrm{b1}}^{\min}|$, then $|P_{\mathrm{b1}}(t)|$ equals $C_{\mathrm{Ab1}}(t)$ as $\hat{u}_1(t) = 1$. Otherwise P_{b1} is equal to $P_{\mathrm{b1}}^{\min}$ with a negative $P_{\mathrm{b1}}^{\min}$. When $u_1(t) = -1$, we first need to identify whether $P_{\mathrm{load1}}(t)$ has already been powered by battery 2/AC utility grid or not. If load 1 has been powered, then no discharge is needed for battery 1 as $\hat{u}_1(t) = 0$. Otherwise, further identification will be carried out whether enough energy is left in battery 1 to satisfy household load 1. With enough energy, battery 1 discharges to household load 1 as $\hat{u}_1(t) = -1$. When $u_1(t) = -2$, the PMU needs to check the household load 2, which has already been powered as $\hat{u}_1(t) = 0$ or not been powered as $\hat{u}_1(t) = -2$. Then it decides to discharge to appliances in housing unit 2 or not, according to $\hat{u}_1(t)$. As for $u_1(t) = -3$, the most complicated adjustment is shown in Table 5. The PMU first checks in a sequence of household load 1 and household load 2. Then PMU chooses the appropriate adjustment control action based on verifications above. Other cases are adjusted in the same way.

4.3. Optimization goal

The ADHDP-based residential energy scheduling is to help end users to reduce cost over a long time. Hence, the primary optimization portion is the cost of the AC utility grid electricity when employing RRTP. The other portion is the penalty cost which values the usage of batteries' energy. It is assumed that residential real-time pricing and storage devices are employed in the energy management system. Appropriate scheduling of batteries in both two housing units can be arranged to minimize the total cost C_{ost}.

$$C_{\mathrm{ost}} = \sum_{i=0}^{t} C(i)P_{\mathrm{grid}}(i) + \sum_{i=0}^{t} C_{\mathrm{p}}(i)[C_{\mathrm{Ab1}}(i) + C_{\mathrm{Ab2}}(i)] \tag{10}$$

The performance index function (10) is developed for the first time with the absent energy penalty cost. The first term of the total cost is about buying electricity from AC utility grid; the less, the better. The second term is the developed absent energy penalty cost, which refers to energy balance with the measure to improve batteries' recharge. The penalty energy price for batteries is acquired from the RRTP. They tend to change within the same limited range but in a reverse trend as shown in Fig. 2.

4.4. ADHDP-based residential energy scheduling algorithm

Our developed residential energy scheduling algorithm is a special version of ADP, or more precisely, ADHDP. The training process of the newly developed algorithm is simplified with training mainly for the critic network as shown in Fig. 4.

Initial data is the basis for the ADHDP-based residential energy scheduling algorithm. Then data includes the system states and

Table 4
Adjustment data of control action $u_1(t) \in \{1, 0, -1, -2\}$.

Initial control action $u_1(t)$	Condition		Charging/discharging power $P_{\mathrm{b1}}(t)$	Adjusted control action $\hat{u}_1(t)$
1	$E_{\mathrm{b1}}(t) = E_{\mathrm{b1}}^{\max}$		Idle: 0	0
	$C_{\mathrm{Ab1}}(t) \leqslant \lvert P_{\mathrm{b1}}^{\min}\rvert$		Charging: $C_{\mathrm{Ab1}}(t)$	1
	$C_{\mathrm{Ab1}}(t) > \lvert P_{\mathrm{b1}}^{\min}\rvert$		Charging: $\lvert P_{\mathrm{b1}}^{\min}\rvert$	1
0	0		Idle: 0	0
−1	$P_{\mathrm{load1}}(t) = P_{\mathrm{b2}}(t)$ or $P_{\mathrm{grid}}(t)$		Idle: 0	0
	else	$P_{\mathrm{load1}}(t) > E_{\mathrm{b1}}(t) - E_{\mathrm{b1}}^{\min}$	Idle: 0	0
		$P_{\mathrm{load1}}(t) \leqslant E_{\mathrm{b1}}(t) - E_{\mathrm{b1}}^{\min}$	Discharging: $P_{\mathrm{load1}}(t)$	−1
−2	$P_{\mathrm{load2}}(t) = P_{\mathrm{b2}}(t)$ or $P_{\mathrm{grid}}(t)$		Idle: 0	0
	else	$P_{\mathrm{load2}}(t) > E_{\mathrm{b1}}(t) - E_{\mathrm{b1}}^{\min}$	Idle: 0	0
		$P_{\mathrm{load2}}(t) \leqslant E_{\mathrm{b1}}(t) - E_{\mathrm{b1}}^{\min}$	Discharging: $P_{\mathrm{load2}}(t)$	−2

Table 5
Adjustment data of control action $u_1(t) = -3$.

Condition				Discharging power $P_{b1}(t)$	Adjusted control action $\hat{u}_1(t)$
$P_{load1}(t) = P_{b2}(t)$ or $P_{grid}(t)$	$P_{load2}(t) = P_{b2}(t)$ or $P_{grid}(t)$			0	0
	else	$P_{load2}(t) \geqslant E_{b1}(t) - E_{b1}^{min}$		0	0
		$P_{load2}(t) < E_{b1}(t) - E_{b1}^{min}$		$P_{load2}(t)$	−2
else	$P_{load1}(t) > E_{b1}(t) - E_{b1}^{min}$	$P_{load2}(t) = P_{b2}(t)$ or $P_{grid}(t)$		0	0
		else	$P_{load2}(t) > E_{b1}(t) - E_{b1}^{min}$	0	0
			$P_{load2}(t) \leqslant E_{b1}(t) - E_{b1}^{min}$	$P_{load2}(t)$	−2
	$P_{load1}(t) \geqslant E_{b1}(t) - E_{b1}^{min}$	$P_{load2}(t) = P_{b2}(t)$ or $P_{grid}(t)$		$P_{load1}(t)$	−1
		else	$P_{load1}(t) + P_{load2}(t) > E_{b1}(t) - E_{b1}^{min}$	$P_{load1}(t)$	−1
			$P_{load1}(t) + P_{load2}(t) \leqslant E_{b1}(t) - E_{b1}^{min}$	$P_{load1}(t) + P_{load2}(t)$	−3

other pretreatment data such as RRTP, housing units' loads, batteries' remaining energy capacity and the utility function. The utility function $U(t)$ is acquired by Eq. (12). Eq. (11) shows how the cost to be computed at time t. Inside the scheduling process, the critic network output $Q(t-1)$ is obtained as $Q(t-1) = \gamma Q(t) + U(t)$. While training, $Q(t-1)$ can be given by utility $U(t) + \gamma Q(t)$ with the mapping: $\{x(t-1); u(t-1)\} \longrightarrow \{Q(t-1)\}$.

$$C_{ost}(t) = C(t)P_{grid}(t) + C_p(t)[C_{Ab1}(t) + C_{Ab2}(t)] \tag{11}$$

$$U(t) = \frac{C_{ost}(t)}{\max\{C_{ost}(1), C_{ost}(2), \ldots\}} \tag{12}$$

The control sequences of batteries are given from the control parameters $\{1, 0, -1, -2, -3\}$ equally. The system states needed for the networks are in Table 5. The overall residential energy scheduling algorithm is implemented by four steps. Afterwards, the detailed control strategy can be obtained. The ADHDP-based residential energy scheduling algorithm implemented by neural networks is explained step by step in Algorithm 1.

First, the ADHDP-based residential energy scheduling networks need to be well trained. With the preexisting system states and other data, the training strategy is implemented according to [40], which is used in critic and action neural networks training.

Second, select the best control strategy of batteries. The well-trained network can be used to select the best control sequences for batteries. All the feasible control sequences are in Table 3. The well-trained network combs through the whole feasible control sequences and finds out the best control strategy.

Third, compute the total cost C_{ost} of electricity. The total cost is the sum of the AC utility grid cost and the penalty cost. Then the cost can be used to compare with the previous one.

Four, improve the networks continuously. Afterwards, the residential energy scheduling has already been established. During the application, new data comes at each hour. The network can be retrained and changes to become more and more skilled at local energy scheduling.

Algorithm 1. ADHDP-based residential energy scheduling algorithm

Part I: Initialization
1: Give the basic parameters for the ADHDP-based residential energy scheduling algorithm.
2: Collect data including system states x and control actions u_1, u_2, and load them into the algorithm.
3: Give the computation precision $\varepsilon > 0$ and the number of loops.

Part II: Iteration
1: Initialize two critic networks which are exactly the same, $\text{Cnetwork}_1 = \text{Cnetwork}_2$.
2: Compute the output Q of the network Cnetwork_1.
3: Compute Q' as $Q' = rQ + U$.
4: Train Cnetwork_2 with the desired output Q'.
5: Copy Cnetwork_2 to Cnetwork_1, as $\text{Cnetwork}_1 = \text{Cnetwork}_2$.
6: If the training has met the requirements (the pre-set number of loops or the computation precision), then goto the next step. Otherwise, goto step 2 and repeat the process.
7: Pick the best critic network and the control strategy.

5. Simulations

The developed ADHDP-based residential energy scheduling algorithm needs to be evaluated in this section. Our ultimate goal of the simulations is to minimize the cost of electricity for users as

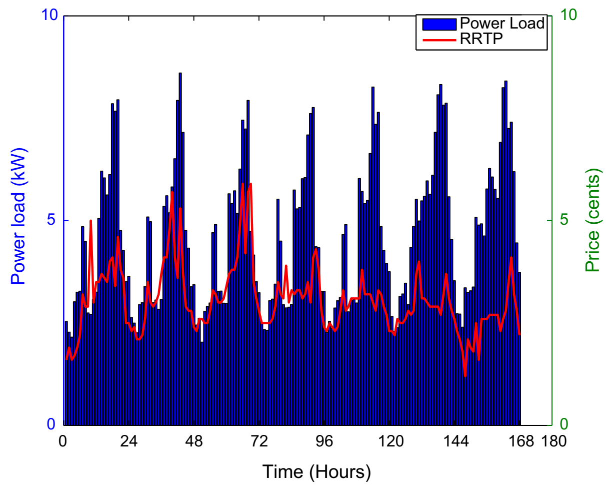

Fig. 5. RRTP and typical residential power loads of housing unit 1.

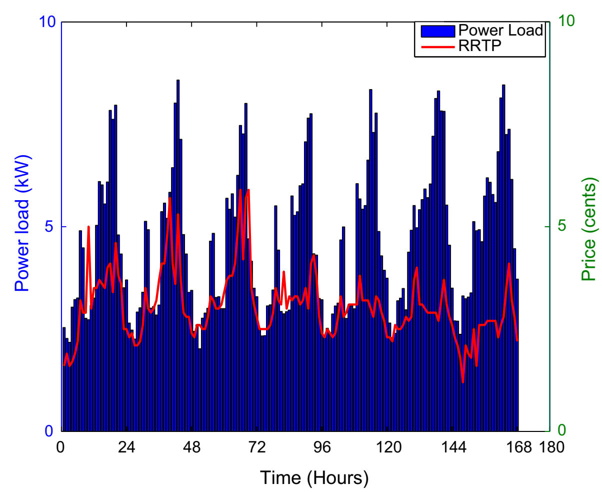

Fig. 6. RRTP and typical residential power loads of housing unit 2.

much as possible. Meanwhile, the control strategy satisfies the local restrictions for practical applications.

The simulation scenario is designed with two housing units producing local appliances load, two batteries to be controlled, and one AC utility grid. All the parts are connected via the PMU. Two housing units have similar appliances and their maximum power loads are the same as 10 kW. Their typical local power loads can be seen in Figs. 5 and 6, which are one-week horizon loads of the two housing units. It is clear that during weekdays the peak loads appear in the morning at about 8:00 and in the evening at about 19:00. The phenomenon shows that people get up and start moving in the morning, while they do cooking and recreational activities in the evening. However, during weekends people do "at home" activities over all the day and the loads increase slowly until the evening peak. Two batteries are initially the same with each 100 kW h. Their maximum charging rate is 10 kW and maximum discharging rate 20 kW. The initial capacities of the batteries are 90% of the total, while the minimum remaining capacities for them to keep are 10% of the total so as to reduce damage to batteries. The AC utility grid is assumed to have enough power for all loads at any time.

In Figs. 5 and 6, their curves are similar from the appearance, but their numeric values are definitely different. With the added 12% random noise for the two residential power loads, the numeric values change randomly within the agreed limits [36]. Therefore the detailed control strategies for the two batteries might have been vastly different during the one-week horizon. Obviously, the RRTP has positive correlation with the specific load values.

The ADHDP-based residential energy scheduling simulations mainly refer to the critic network, which is a multilayer feedforward neural network. The network's structure is 7-16-1 (7 neurons in the input layer, 16 neurons in the hidden layer, and 1 neuron in the output layer). Seven input neurons are the RRTP, two batteries' energy states, two housing units loads, and two control variables of the batteries. The activation function uses hyperbolic tangent function to compute. After that, the critic network outputs the $Q(t)$ as a value approaching the ideal optimal performance index $J^*(x(k))$.

The simulation is carried out in the environment of the MathWorks Matlab R2013b, which works on a DELL computer running a Windows 7 64-bit OS, with a Intel Core i5 CPU @3.2 GHz and 8GB ram. The simulation results are shown next with some figures.

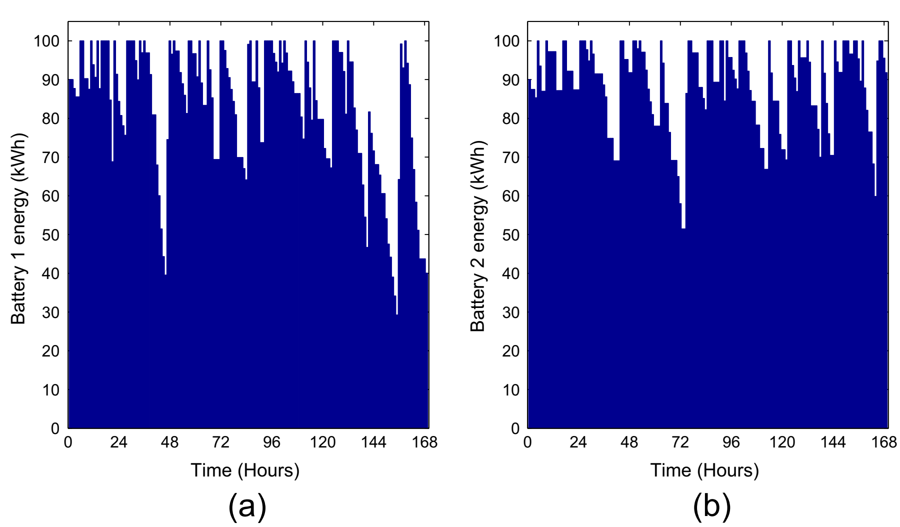

Fig. 7. Energy remaining capacities of batteries 1 and 2. (a) Energy changes in battery 1. (b) Energy changes in battery 2.

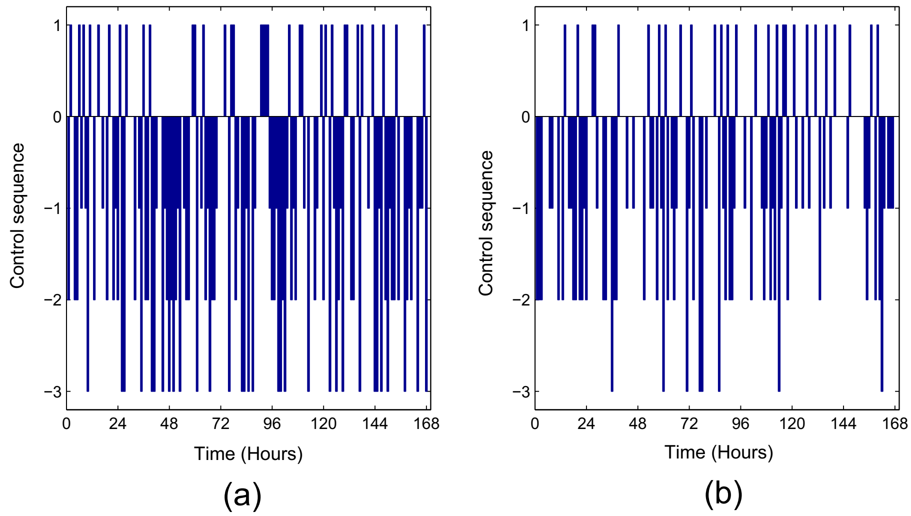

Fig. 8. Optimal control sequences of batteries 1 and 2. (a) Optimal control sequence of battery 1. (b) Optimal control sequence of battery 2.

Fig. 7 shows the energy levels of the two batteries during the one-week horizon. The initial capacities of both batteries are 90%. After running, energy capacities change at every hour. Nevertheless, batteries are at full-capacity states in the night. It is wise for the batteries to discharge during the daytime when the RRTP is high and to be charged during night when the RRTP is low. Therefore, the residents can reduce cost on electricity without interfering with the normal life concerning power consumption. Meanwhile, it helps the AC utility grid to reduce peak loads and shift the "peaks" to electricity "troughs" to make the whole load for the AC utility grid smoother.

The adjustment of control actions will be described to satisfy housing load 1. Housing load 2 will be handled similarly. Therefore, two control actions and batteries' remaining capacities are different most of the time. As battery 1 has to discharge and charge much deeper to satisfy its own housing unit demand due to the nearby priority principle. Battery 1 discharges more energy than battery 2 in total during the week. In Fig. 8, the optimal control sequences are given. The optimal scheduling of the two batteries can be achieved, when the PMU gives the control orders in the way of {1: during trough-load time; $-1/-2/-3$: during peak-load time; 0: during middle-load time}. The control action {1} mainly appears in nighttime. While the actions $\{-1/-2/-3\}$ often appear in the day.

The batteries tend to charge at trough-load time and discharge at peak-load time. It has to be noticed that when the RRTP is low and the batteries charge at the maximum charging rate so as to save money and time for next day. Due to the amounts of calculation, the results shown in the figures are not regularly the same for each day, not to mention the energy-consuming differences between weekdays and weekends. Roughly speaking, the batteries discharge during the period of 6:00–20:00, and charge during the period of 22:00–02:00 during weekdays. While at weekends, they discharge between 7:00 and 19:00, and charge between 21:00 and 04:00. More specifically, residents consume much more energy than usual during weekends and the remaining energy capacities drop fast. Therefore the discharging time becomes shorter and the charging time turns out to be a little longer.

Overall, the developed ADHDP-based residential energy scheduling algorithm can reduce cost for residents. After the simulation part, the total cost using the new algorithm over a one-week period is 5070.45 cents. In contrast, the original expenditure without any energy storage devices is 8244.60 cents. It is obvious that 38.5% of the original expenditure can be saved for the residents after one-week operation. Consequently, the goal of reducing cost on electricity has been achieved. Besides, the developed algorithm has the ability to adapt to new environments. Whenever the habits of consuming electricity are changed, whenever there are appliances added or removed, or whenever the residents are replaced, the ADHDP-based residential energy scheduling algorithm can always learn new characteristics and evolve in a timely fashion.

6. Conclusion

In this paper, a new ADHDP-based residential energy scheduling algorithm with home energy inter-exchange is developed. The algorithm is used to solve the optimization problem of reducing electricity cost between two housing units under the circumstance of RRTP. Neural networks are trained online or offline with the preexisting concrete data at first. Once the networks gain the characteristics of the new environments, they can contribute to the actual money-saving work. It is shown that the developed absent energy penalty cost has good application effects as part of the performance index function (10). Besides, the algorithm has the ability to adapt to new environment timely. Simulations have shown that the algorithm has reduced a big part of the original total electricity cost. The reduction further illustrates the practicality of the new algorithm in micro smart grid management. Future work will be focused on adding more devices and mixed algorithms.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 61273140, 61374105, and 61233001, and in part by Beijing Natural Science Foundation under Grant 4132078.

References

[1] Zhao Z, Yan H, Zuo J, Tian Y, Zillante G. A critical review of factors affecting the wind power generation industry in China. Renew Sustain Energy Rev 2013;19:499–508.
[2] European Commission [EC]. European technology platform smartgrids: vision and strategy for Europes electricity networks of the future (EUR 22040), Directorate-General for Research Cooperation, Energy, Brussels; 2006.
[3] European Commission [EC], European technology platform smartgrids: strategic research agenda for Europes electricity networks of the future (EUR 22580). Directorate-General for Research Cooperation, Energy, Brussels; 2007.

[4] 111th U.S. Congress [USC]. The American Recovery and Reinvestment Act. The American Recovery and Reinvestment Act of 2009. H.R.1., S.1., Washington (DC); 2009.
[5] Zakariazadeh A, Jadid S, Siano P. Economic-environmental energy and reserve scheduling of smart distribution systems: a multiobjective mathematical programming approach. Energy Convers Manage 2014;78:151–64.
[6] Figueiredo J, Martins J. Energy production system management renewable energy power supply integration with building automation system. Energy Convers Manage 2010;51(6):1120–6.
[7] Lin S, Chen J. Distributed optimal power flow for smart grid transmission system with renewable energy sources. Energy 2013;56:184–92.
[8] Qureshi WA, Nair NKC, Farid MM. Impact of energy storage in buildings on electricity demand side management. Energy Convers Manage 2011;52(5):2110–20.
[9] Huang T, Liu D. A self-learning scheme for residential energy system control and management. Neural Comput Appl 2013;22(2):259–69.
[10] Hannan MA, Azidin FA, Mohamed A. Multi-sources model and control algorithm of an energy management system for light electric vehicles. Energy Convers Manage 2012;62:123–30.
[11] Wei Q, Liu D, Shi G. A novel dual iterative Q-learning method for optimal battery management in smart residential environments. IEEE Trans Ind Electron 2015;62(4):2509–18.
[12] Werbos PJ. Approximate dynamic programming for real-time control and neural modeling. In: White DA, Sofge DA, editors. Handbook of intelligent control: neural, fuzzy and adaptive approaches. New York: Van Nostrand; 1992.
[13] Prokhorov DV, Wunsch DC. Adaptive critic designs. IEEE Trans Neural Netw 1997;8(5):997–1007.
[14] Liu D, Xiong X, Zhang Y. Action-dependent adaptive critic designs. In: Proceedings of the INNS-IEEE international joint conference on neural networks. Washington; 2001.
[15] Liu D, Wei Q. Multi-person zero-sum differential games for a class of uncertain nonlinear systems. Int J Adaptive Control Signal Process 2014;28(3–5):205–31.
[16] Wei Q, Wang D, Zhang D. Dual iterative adaptive dynamic programming for a class of discrete-time nonlinear systems with time-delays. Neural Comput Appl 2013;23(7–8):1851–63.
[17] Bhasin S, Sharma N, Patre P, Dixon W. Asymptotic tracking by a reinforcement learning-based adaptive critic controller. J Control Theory Appl 2011;9(3):400–9.
[18] Lewis FL, Syrmos VL. Optimal control. New York (NY): Wiley; 1995.
[19] Wei Q, Liu D, Xu Y. Neuro-optimal tracking control for a class of discrete-time nonlinear systems via generalized value iteration adaptive dynamic programming approach. Soft Comput 2014:1–10.
[20] Wei Q, Liu D. Neural-network-based adaptive optimal tracking control scheme for discrete-time nonlinear systems with approximation errors. Neurocomputing 2015;149(3):106–15.
[21] Mohagheghi S, Del Valle Y, Venayagamoorthy GK, Harley RG. A proportional-integrator type adaptive critic design-based neurocontroller for a static compensator in a multimachine power system. IEEE Trans Ind Electron 2006;54(1):86–96.
[22] Lu C, Si J, Xie X. Direct heuristic dynamic programming for damping oscillations in a large power system. IEEE Trans Syst, Man, Cyber, Part B: Cyber 2008;38(4):1008–13.
[23] Fuselli D, De Angelis F, Boaro M, Squartini S, Wei Q, Liu D, et al. Action dependent heuristic dynamic programming for home energy resource scheduling. Int J Electr Power Energy Syst 2013;48:148–60.
[24] Keçebaş A, Yabanova İ, Yumurtacı M. Artificial neural network modeling of geothermal district heating system thought exergy analysis. Energy Convers Manage 2012;64:206–12.
[25] Sousa T, Morais H, Vale Z, Faria P, Soares J. Intelligent energy resource management considering vehicle-to-grid: a simulated annealing approach. IEEE Trans Smart Grid 2012;3(1):535–42.
[26] Yu S, Zhu K, Zhang X. Energy demand projection of China using a path-coefficient analysis and PSO-GA approach. Energy Convers Manage 2012;53(1):142–53.
[27] Safari S, Ardehali MM, Sirizi MJ. Particle swarm optimization based fuzzy logic controller for autonomous green power energy system with hydrogen storage. Energy Convers Manage 2013;65:41–9.
[28] Arcos-Aviles D, Vega C, Guinjoan F, Marroyo L, Sanchis P. Fuzzy logic controller design for battery energy management in a grid connected electro-thermal microgrid. In: IEEE 23rd international symposium on industrial electronics (ISIE). Istanbul; 2014.
[29] Morais H, Kádár P, Faria P, Vale ZA, Khodr HM. Optimal scheduling of a renewable micro-grid in an isolated load area using mixed-integer linear programming. Renew Energy 2009;35(1):151–6.
[30] Comodi G, Giantomassi A, Severini M, Squartini S, Ferracuti F, Fonti A, et al. Multi-apartment residential microgrid with electrical and thermal storage devices: experimental analysis and simulation of energy management strategies. Appl Energy 2015;137:854–66.
[31] Chuang A, McGranaghan M. Functions of a local controller to coordinate distributed resources in a smart grid. In: Power and energy society general meeting-conversion and delivery of electrical energy in the 21st century. Pittsburgh (PA); 2008.
[32] Liu X, Wang P, Loh PC. A hybrid AC/DC microgrid and its coordination control. IEEE Trans Smart Grid 2011;2(2):278–86.
[33] Lee T. Operating schedule of battery energy storage system in a time-of-use rate industrial user with wind turbine generators: a multipass iteration particle swarm optimization approach. IEEE Trans Energy Convers 2007;22(3):774–82.
[34] Qureshi A, Weber R, Balakrishnan H, Guttag J, Maggs B. Cutting the electric bill for internet-scale systems. ACM SIGCOMM Comput Commun Rev 2009;39(4):123–34.
[35] Beccali M, Cellura M, Lo Brano V, Marvuglia A. Forecasting daily urban electric load profiles using artificial neural networks. Energy Convers Manage 2004;45(18):2879–900.
[36] OpenEI. <http://en.openei.org/community> [accessed August 2014].
[37] Yau T, Walker LN, Gupta HL, Raithel R. Effects of battery storage devices on power system dispatch. IEEE Trans Power Apparatus Syst 1981;PAS-100(1):375–83.
[38] Wang F, Zhang H, Liu D. Adaptive dynamic programming: an introduction. IEEE Comput Intell Mag 2009;4(2):39–47.
[39] Bellman RE. Dynamic programming. Princeton (NJ): Princeton University Press; 1957.
[40] Lendaris GG, Paintz C. Training strategies for critic and action neural networks in dual heuristic programming method. In: International conference on neural networks. Houston (TX); 1997.