Soft Comput
DOI 10.1007/s00500-016-2194-y

METHODOLOGIES AND APPLICATION

Optimization of electricity consumption in office buildings based on adaptive dynamic programming

Guang Shi[1] · Qinglai Wei[1] · Derong Liu[2]

Abstract In this paper, an optimization method based on adaptive dynamic programming is developed to improve the electricity consumption of rooms in office buildings through optimal battery management. Rooms in office buildings are generally divided into office rooms, computer rooms, storage rooms, meeting rooms, etc., and each category of rooms have different characteristics of electricity consumption, which is divided into electricity consumption from sockets, lights and air-conditioners in this paper. The developed method based on action-dependent heuristic dynamic programming is explained in detail, and different optimization strategies of electricity consumption in different categories of rooms are proposed in accordance with the developed method. Finally, a detailed case study on an office building is given to demonstrate the practical effect of the developed method.

Keywords Office buildings · Electricity consumption optimization · Battery management · Optimal control · Adaptive dynamic programming · Neural networks

1 Introduction

Over the past years, humans have become increasingly dependent on electricity both in life and work. The constantly rising cost, growing environmental pollution and severe resource shortage have posed new opportunities and challenges to the development of efficient control and management strategies for energy consumption. Smart grid, as an intelligent power grid, has attracted widespread attention in recent years. Extensive research has been conducted in both theory and practice. Severini et al. (2013) developed a hybrid algorithmic framework including genetic, neural network and deterministic optimization algorithms to optimize energy consumption in smart homes. Arsuaga-Rios and Vega-Rodriguez (2015) presented a multi-objective brain storm algorithm (MOBSA) to improve energy use in grid systems. Ma et al. (2014) proposed a distributed algorithm for energy consumers to control their energy consumption. Li et al. (2013) studied a data-driven strategy based on the type-2 fuzzy method to model and optimize energy consumption in smart homes and intelligent buildings. Anvari-Moghaddam et al. (2015) presented a management scheme to improve the efficiency of residential energy consumption in a typical smart micro-grid. With in-depth development of smart grid, increasing intelligence is required in the design of efficient energy management systems. Therefore, optimal battery management has become an important approach to saving expense on electricity in smart grid.

Proposed by Werbos (1977, 1991), adaptive dynamic programming (ADP) (Wei and Liu 2014; Song et al. 2013), also known as "adaptive critic designs" (Prokhorov and Wunsch 1997; Ni et al. 2013), "approximate dynamic programming" (Xu et al. 2014a; Molina et al. 2013), "neural dynamic programming" (Enns and Si 2003), "neuro-dynamic programming" (Bertsekas and Tsitsiklis 1996; Xu and Jagannathan 2013) and "reinforcement learning" (Ni et al. 2013; Xu et al. 2014b), has been verified with strong ability to solve the optimization problem of complex nonlinear systems by means of its strong self-learning capacity. The

Communicated by V. Loia.

✉ Qinglai Wei
qinglai.wei@ia.ac.cn

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

2 The School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Published online: 27 May 2016

method of ADP approximates the optimal performance index function and optimal controller by using function approximation structures (Wang et al. 2009) and circumvents the "curse of dimensionality" in dynamic programming (DP) by using the forward-in-time approach to solve the Hamilton-Jacobi-Bellman equation (Werbos 1991). Recent years witnessed extensive research on ADP (Liu and Wei 2012; Ni and He 2013; Wei et al. 2014, 2016a; Na and Herrmann 2014; Zhao et al. 2014). In Werbos (1991), ADP was divided into four major schemes, i.e., heuristic dynamic programming (HDP), action-dependent heuristic dynamic programming (ADHDP), dual heuristic dynamic programming (DHP) and action-dependent dual heuristic dynamic programming (ADDHP). In Prokhorov and Wunsch (1997), two more schemes of ADP were proposed, namely globalized dual heuristic dynamic programming (GDHP) and action-dependent globalized dual heuristic dynamic programming (ADGDHP). As one of the typical schemes of ADP, ADHDP has been effectively used in optimal battery control of home energy management systems (Huang and Liu 2013; Boaro et al. 2013), in which renewable resources, including wind and solar energies, were introduced into the energy systems.

However, most of previous research on management of energy consumption based on ADP focused on residential energy systems (Huang and Liu 2013; Boaro et al. 2013; Fuselli et al. 2013; Wei et al. 2015a, b, 2016b), rather than energy consumption in office buildings. Nevertheless, as a significant component of urban structure, office buildings account for a great proportion of social energy consumption, in which electricity consumption plays the key role. Moreover, with the rapid development of electricity storage technology, optimal management based on electricity storage has been widely concerned (Amjadi and Williamson 2010; Guerrero et al. 2013). Therefore, it is of great importance to improve the electricity consumption of office buildings based on electricity storage.

In our previous work (Shi et al. 2015), a data-driven method based on echo state network (ESN) is developed to classify rooms in office buildings into different categories, including office rooms, computer rooms, storage rooms and meeting rooms. Hence, it is necessary to further develop corresponding optimization strategies to improve the electricity consumption of rooms in office buildings in accordance with different characteristics of different categories of rooms and therefore save the total expense on electricity from the power grid. As far as we know, no research has been conducted in this respect, which motivates our research.

The rest of the paper is arranged as follows. Problem formulation of the electricity consumption management system of a room in an office building is given in Sect. 2. The developed optimization algorithm of electricity consumption based on ADP is elaborated in Sect. 3, and implementation by neural networks is explained. In Sect. 4, a detailed case study is presented to show the effectiveness and superiority of the developed algorithm. Finally, in Sect. 5, the conclusion is drawn and future work is proposed.

2 Problem formulation

In this section, the electricity consumption management system of a room in an office building is described, and the optimization target is presented.

2.1 Electricity consumption management system

As shown in Fig. 1, the electricity consumption management system consists of the power grid, a battery system (composed of a battery and an inverter), a power management unit and electricity demand, in which the electricity demand is divided into electricity demand from sockets, lights and air-conditioners for a typical room in an office building.

Based on Huang and Liu (2013), the model of the battery applied in this paper is described as

$$E_b(t+1) = E_b(t) - P_b(t) \times \eta(P_b(t)), \tag{1}$$

where $E_b(t)$ denotes the energy of the battery at time t with a time step of 1 h, $P_b(t)$ denotes the output power of the battery, while $P_b(t) > 0$ denotes battery discharging, $P_b(t) < 0$ denotes battery charging and $P_b(t) = 0$ denotes an idle state of the battery. The charging/discharging efficiency $\eta(P_b(t))$ of the battery can be derived as

$$\eta(P_b(t)) = 0.898 - 0.173|P_b(t)|/P_{\text{rate}}, \tag{2}$$

where P_{rate} denotes the rated output power of the battery.

Specifically, electricity consumption management in this paper is treated as a discrete-time battery control problem

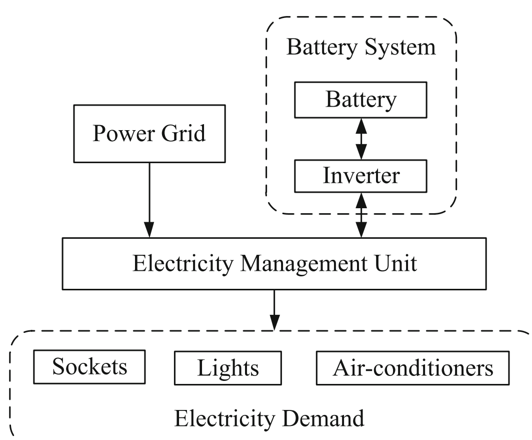

Fig. 1 Electricity consumption management system of a room in an office building

with a time step of 1 h as the data of electricity consumption are collected hourly, so the power output $P_b(t)$ of the battery satisfies $P_b(t)(\mathrm{kW}) \times 1(\mathrm{h}) = P_b(t)(\mathrm{kWh})$, which equals the value of power. Therefore, the battery model expressed in (1) makes sense.

2.2 Optimization target

In this paper, electricity flow from the battery to the power grid is forbidden, i.e., $P_g(t) \geq 0$ is defined as electricity from the grid. To facilitate our analysis, a 1-h delay is introduced in $P_b(t)$ and $P_L(t)$, where $P_L(t)$ denotes electricity demand at time t with $P_L(t) = P_{Ls}(t) + P_{Ll}(t) + P_{La}(t)$, where $P_{Ls}(t)$, $P_{Ll}(t)$ and $P_{La}(t)$ denote electricity demand from sockets, lights and air-conditioners, respectively. Then, the demand balance equation is expressed as

$$P_L(t-1) = P_b(t-1) + P_g(t), \tag{3}$$

which indicates that the electricity supply (from the battery and the power grid) should necessarily balance the electricity demand at each hour. It is also assumed that electricity from the grid is enough to satisfy the electricity demand.

Given the electricity demand and the electricity price which is denoted by $C(t)$, the optimization target of electricity consumption is to obtain the optimal charging/discharging/idle strategy of the battery at each time step to minimize the total performance index function

$$J_T = \sum_{t=0}^{\infty} C(t) \times P_g(t) \tag{4}$$

while meeting the demand balance Eq. (3) and other conditions including $P_g(t) \geq 0$ for the electricity from the grid and $|P_b(t)| \leq P_{\text{rate}}$ for the charging/discharging power of the battery. J_T refers to the total expense from the grid incurred over time. Let $x_1(t) = P_g(t)$, $x_2(t) = E_b(t)$ and $u(t) = P_b(t)$, the equation of the electricity consumption management system can be derived as

$$x(t+1) = F(x(t), u(t), t) = \begin{pmatrix} P_L(t) - u(t) \\ x_2(t) - u(t)\eta(u(t)) \end{pmatrix}, \tag{5}$$

where $x(t) = [x_1(t), x_2(t)]^T$.

Adaptive dynamic programming (ADP), which solves dynamic programming (DP) by approximating optimal solutions, can be applied to obtain the optimal control $u^*(t)$ of the above nonlinear system. Furthermore, given the optimal control $u^*(t)$, we can calculate $u_s^*(t) = \gamma_s(t) \cdot u^*(t)$, $u_l^*(t) = \gamma_l(t) \cdot u^*(t)$ and $u_a^*(t) = \gamma_a(t) \cdot u^*(t)$, to satisfy the electricity demand from sockets, lights and air-conditioners, respectively, where $\gamma_s(t) = P_{Ls}(t)/P_L(t)$, $\gamma_l(t) = P_{Ll}(t)/P_L(t)$ and $\gamma_a(t) = P_{La}(t)/P_L(t)$.

3 Optimization algorithm of electricity consumption based on ADP

In this section, the optimization algorithm of electricity consumption based on ADP is developed to find optimal control strategies for the electricity consumption management system of a room in an office building.

3.1 Adaptive dynamic programming

In accordance with Bellman's principle of optimality (Bellman 1957), the method of DP is applicable to obtaining optimal control actions to solve complex and nonlinear optimization problems. Given the discrete-time nonlinear system in (5), where $x(t)$ denotes the state vector, $u(t)$ denotes the control vector and $F(\cdot)$ denotes the system function, the performance index function (4) of the system can be derived as

$$J[x(t), t] = \sum_{l=t}^{\infty} \gamma^{l-t} U[x(l), u(l), l], \tag{6}$$

where $U[x(l), u(l), l] = C(l) \cdot x_1(l)$ is the utility function, γ is the discount factor satisfying $0 < \gamma \leq 1$, while $J(\cdot)$ depends on the initial state $x(l)$ and the initial time l. DP aims to obtain a series of control actions $u(l)$, $l = t, t+1, \ldots$, which minimize the performance index function in (6). Based on Bellman's principle of optimality (Bellman 1957), the optimal performance satisfies the Hamilton–Jacobi–Bellman (HJB) equation as follows

$$J^*[x(t), t] = \min_{u(t)} (U[x(t), u(t), t] + \gamma J^*[x(t+1), t+1]). \tag{7}$$

The optimal control $u^*(t)$ which achieves the minimum cost at time t is given by

$$u^*(t) = \arg\min_{u(t)} (U[x(t), u(t), t] + \gamma J^*[x(t+1), t+1]). \tag{8}$$

ADP is a method based on the iteration between policy improvement and value approximation of solutions to DP. Compared with traditional ADP schemes including heuristic dynamic programming (HDP) and dual heuristic dynamic programming (DHP), action-dependent heuristic dynamic programming (ADHDP) does not explicitly require a model network in its design, and the control is included in the input of the critic network besides the state, so that the computation precision is higher (Werbos 1991). Therefore, ADHDP is adopted to solve the problem in this paper. Next, the design of ADHDP will be elaborated.

3.2 Action-dependent heuristic dynamic programming

For the optimal control problem concerned in this paper, the method of ADHDP is adopted. Figure 2 shows a typical scheme of ADHDP.

As shown in Fig. 2, an explicit model network is not required, while the critic network is trained to minimize the following error:

$$E_{\mathrm{q}} = \sum_{t=0}^{\infty} E_{\mathrm{q}}(t) = \sum_{t=0}^{\infty} [Q(t-1) - U(t) - \gamma Q(t)]^2, \qquad (9)$$

where $Q(t)$ denotes the output of the critic network at time t, and the critic network follows the input–output relationship denoted by

$$Q(t) = Q[x(t), u(t)], \qquad (10)$$

where $x(t)$ is the state vector and $u(t)$ is the control vector.

If $E_{\mathrm{q}}(t) = 0$ at all time t, it is implied by (9) that

$$\begin{aligned} Q(t-1) &= U(t) + \gamma Q(t) \\ &= U(t) + \gamma[U(t+1) + \gamma Q(t+1)] \\ &= \cdots \\ &= \sum_{l=t}^{\infty} \gamma^{l-t} U(l). \end{aligned} \qquad (11)$$

By comparing (6) and (11), we have $Q(t-1) = J[x(t), t]$.

Based on the error function (9), the critic network is trained with the forward-in-time approach as follows.

Given the output target $Q(t-1) = U(t) + \gamma Q(t)$, the critic network is trained at time $t-1$. That is, the critic network is trained to achieve the mapping as follows

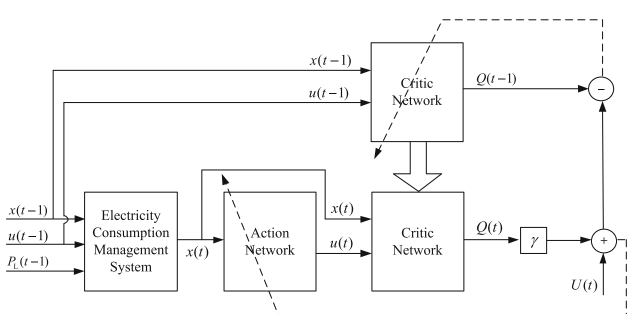

Fig. 2 A typical scheme of ADHDP

$$\left\{ \begin{matrix} x(t-1) \\ u(t-1) \end{matrix} \right\} \rightarrow \{Q(t-1)\}, \qquad (12)$$

where $x(t-1)$ and $u(t-1)$ denote the inputs of the network and $Q(t-1)$ denotes the output of the network. The target output for the network training is calculated with the output at time t as presented in (12). The objective of approximating the mapping denoted by (12) is to satisfy the output of the critic network as

$$Q(t-1) \approx U(t) + \gamma Q(t), \qquad (13)$$

which is required by (11) for approximation of solutions to DP.

After the training of the critic network is completed, the action network is then trained to obtain the control action $u(t)$ which minimizes the output of the critic network $Q(t)$. Therefore, the action network is trained to achieve the mapping as follows

$$\{x(t)\} \rightarrow \{u(t)\}, \qquad (14)$$

where $u(t)$ denotes the target output of the action network. As shown in Fig. 2, the action network is linked to the critic network during the process of training.

After the training of the action network is completed, the performance of the system is checked to terminate or continue the training by returning to the training of the critic network if the performance is unsatisfactory.

Remark 1 In previous research, relevant methods of ADP were developed and successfully applied to the optimization of energy consumption in residential systems. Huang and Liu (2013) proposed a self-learning scheme based on ADP to control residential energy consumption. Boaro et al. (2013) used ADP for renewable energy scheduling and battery management in a residential system. Fuselli et al. (2013) managed home energy sources by using ADHDP. Wei et al. (2015a, b) optimized energy consumption in residential environments via two different iterative ADP methods, respectively. Based on these results, we adopt the scheme of ADHDP or Q-learning (Lewis et al. 2012; Prokhorov and Wunsch 1997) to manage energy consumption in office buildings. The convergence analysis of the algorithm is given in Wei et al. (2016c), which lays a theoretical basis for our research.

3.3 Neural network implementation

Both the above-mentioned neural networks, i.e., critic and action networks, are established as three-layer back-propagation (BP) networks.

3.3.1 Critic network

The target of training the critic network is to obtain $Q(t-1)$ in accordance with (12). The critic network can be established with three input neurons, 14 hidden neurons and one linear output neuron. The three input neurons represent $x(t-1) \in R^2$ and $u(t-1) \in R$, respectively. The output neuron represents $Q(t-1) \in R$. The number of hidden neurons is decided by trial and error. With $k = 0, 1, \ldots$ as the training step, and $Z_{\mathrm{c}}(t-1) = [x(t-1)^{\mathrm{T}}, u(t-1)]^T$ as the input vector of the critic network, the output of the critic network is expressed as $\hat{Q}_{\mathrm{k}}(t-1) = W_{\mathrm{c}}^T(k)\sigma(\mathcal{Z}_{\mathrm{c}}(t-1))$, where $\mathcal{Z}_{\mathrm{c}}(t-1) = Y_{\mathrm{c}}^T Z_{\mathrm{c}}(t-1)$ and $\sigma(\cdot)$ is a sigmoid function (Si and Wang 2001). To improve the efficiency of training, the input-hidden weight matrix Y_{c} is fixed during the training, and only the hidden-output weight matrix $W_c(k)$ is updated. Based on Si and Wang (2001), the weight matrix of the critic network is updated as

$$W_{\mathrm{c}}(k+1) = W_{\mathrm{c}}(k) - \alpha_{\mathrm{c}} \left[\frac{\partial E_{\mathrm{c}}(k)}{\partial W_{\mathrm{c}}(k)} \right], \tag{15}$$

where $E_{\mathrm{c}}(k) = \frac{1}{2}(e_{\mathrm{c}}(k))^2$, $e_{\mathrm{c}}(k) = \hat{Q}_{\mathrm{k}}(t-1) - Q_{\mathrm{k}}(t-1)$, and $\alpha_{\mathrm{c}} > 0$ denotes the learning rate of the critic network.

3.3.2 Action network

The target of training the action network is to determine the control action $u(t)$ which minimizes the output of the critic network $Q(t)$ as indicated in (14). The action network can be established with two input neurons, nine hidden neurons and one linear output neuron. The two input neurons represent $x(t) \in R^2$, the output neuron represents $u(t) \in R$ and the number of hidden neurons is also decided by trial and error. The output of the action network can be expressed as $\hat{u}_{\mathrm{k}}(t) = W_{\mathrm{a}}^T(k)\sigma(\mathcal{Z}_{\mathrm{a}}(t))$, where $\mathcal{Z}_{\mathrm{a}}(t) = Y_{\mathrm{a}}^T x(t)$ and $\sigma(\cdot)$ is a sigmoid function (Si and Wang 2001). Similarly, the input-hidden weight matrix Y_{a} is fixed during the training, and only the hidden-output weight matrix $W_{\mathrm{a}}(k)$ is updated. Based on Si and Wang (2001), the weight matrix of the action network is updated as

$$W_{\mathrm{a}}(k+1) = W_{\mathrm{a}}(k) - \alpha_{\mathrm{a}} \left[\frac{\partial E_{\mathrm{a}}(k)}{\partial W_{\mathrm{a}}(k)} \right], \tag{16}$$

where $E_{\mathrm{a}}(k) = \frac{1}{2}(e_{\mathrm{a}}(k))^2$, $e_{\mathrm{a}}(k) = \hat{u}_{\mathrm{k}}(t) - u_{\mathrm{k}}(t)$, and $\alpha_{\mathrm{a}} > 0$ denotes the learning rate of the action network.

4 Case study

In this section, a detailed case study is given to illustrate the effectiveness and superiority of the developed method. The case study is based on an office building in one of our practical applications. The building is composed of 14 floors in total, each of which contains 6 rooms except the first floor, since it is used as the entrance hall of the entire building. The entire building adopts a central air-conditioning system, where each room is allowed to control air-conditioning by several switches. The data of each room are divided into electricity consumption from sockets, lights and air-conditioners, which are, respectively, measured on site by three electricity meters installed inside the room. The three types of electricity consumption can basically cover the entire electricity consumption in a room.

In our previous work (Shi et al. 2015), a data-driven classification method based on echo state network (ESN) is developed to classify rooms in office buildings into different categories, including office rooms, computer rooms, storage rooms and meeting rooms. Proposed by Jaeger (2001) and Jaeger and Haas (2004), ESN is a recurrent neural network (RNN) which has achieved extensive applications in chaotic time series prediction and classification. The method based on ESN developed in Shi et al. (2015) is divided into two steps. Given the data of electricity consumption in a room, the first step is to reconstruct the behavior of electricity consumption in three types by using three ESNs. The second step is to classify the room into a certain category by establishing another ESN.

The purpose of reconstructing electricity consumption is to show that a certain category of rooms in an office building corresponds to a fixed pattern of electricity consumption. Based on different characteristics of electricity consumption in a room, the room can be classified into a certain category. Generally, the electricity consumption in a room of a certain category does not frequently vary in view of relatively fixed working routines for personnel who work in the room, which lays the foundation for our room classification. With rooms classified into different categories, we aim to develop different strategies to optimize electricity consumption in different rooms.

Based on the results in Shi et al. (2015), we apply the developed method to optimize the electricity consumption of each room by installing a battery in each room (if necessary), so as to reduce the expense on electricity from the power grid. For the reason that stepped electricity price rather than real-time electricity price is implemented in China, we refer to typical real-time electricity price in non-summer seasons in the USA from ComEd Company. Combined with the real-time electricity price shown in Fig. 3, results of different categories of rooms are, respectively, presented as follows.

4.1 Office room

As given in Shi et al. (2015), Room 3 on the 4th floor is an office room, whose original electricity consumption in three

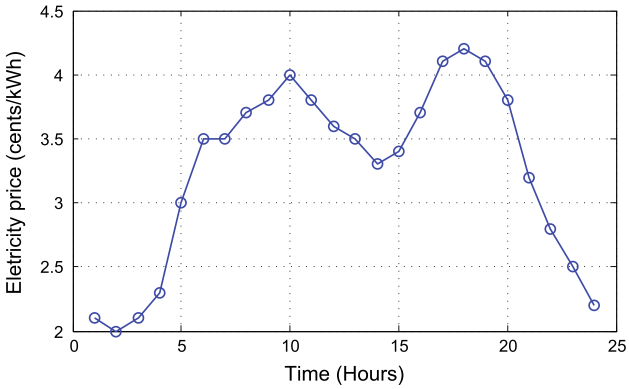

Fig. 3 Typical electricity price in non-summer seasons

types in 5 working days and results of electricity consumption behavior reconstructed by ESNs are shown in Fig. 4. It can be seen that all the three types of electricity consumption display a typical "double-peak" characteristic. On one hand, all the three curves reach their peaks in mid-morning around 11:00 and mid-afternoon around 16:00 on a working day. On the other hand, they achieve a low point at noon because part of personnel in the office room who usually go out for lunch then may switch off some electrical appliances using sockets, turn off some lights or adjust the temperature set for the air-conditioners, while some others who have their lunch inside the room may still consume some electricity. However, due to no special requirements on electricity consumption before and after work, all the appliances consuming electricity are turned off when nobody stays in the room, so the electricity consumption in non-working hours is close to zero.

Next, in accordance with the electricity demand in three types in Fig. 4 and the real-time electricity price in Fig. 3, the electricity consumption of the office room is optimized with the developed method based on ADHDP. A battery with a capacity of 15 kWh and a rated power output of 2 kW is installed in the room. Given the performance index function (4), the optimization method based on ADHDP is implemented by neural networks for 50 iterations to guarantee the computation precision of the entire algorithm as 10^{-4}. Both the critic and action networks are trained with a learning rate of 0.01 and a network precision of 10^{-6}. Meanwhile, carried out in the MATLAB R2012a environment on an Intel Core 2, the simulation of the office room only takes less than 1 min, which could meet practical demands especially when the number of rooms increases. Based on the electricity demand and electricity price in 5 working days, optimal control strategies of the battery are shown in Fig. 5, from which we can see that the control strategies for all the electricity demand in three types follow the same pattern given a similar pattern of demand, i.e., the battery is generally charged when the electricity price is low during a day and discharged to satisfy the demand when the electricity price is high. The battery

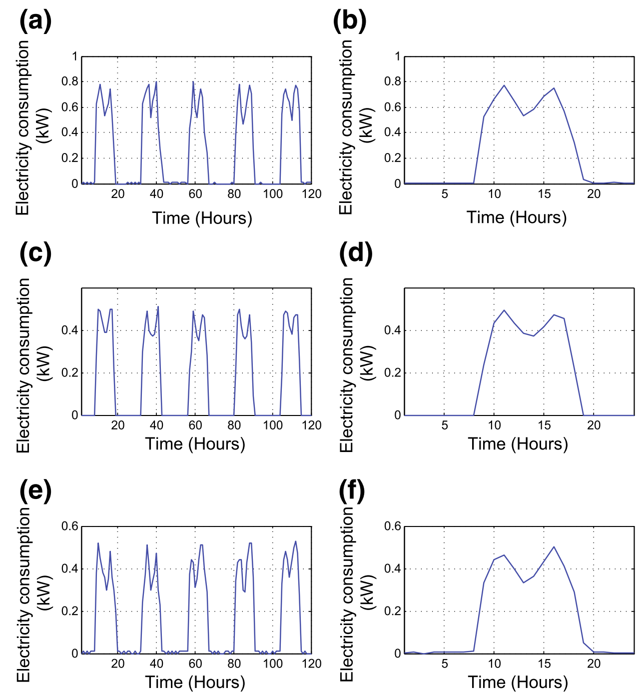

Fig. 4 Electricity consumption of an office room. **a** Electricity consumption from sockets in 5 working days. **b** Electricity consumption from sockets reconstructed by ESN. **c** Electricity consumption from lights in 5 working days. **d** Electricity consumption from lights reconstructed by ESN. **e** Electricity consumption from air-conditioners in 5 working days. **f** Electricity consumption from air-conditioners reconstructed by ESN

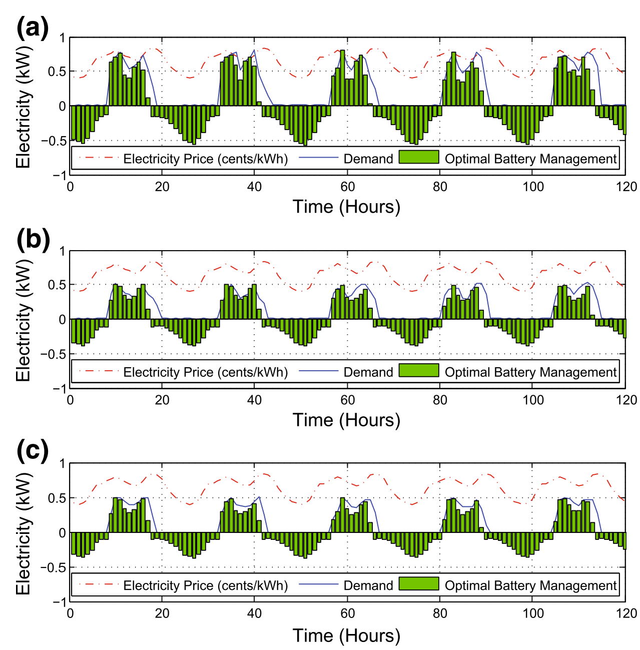

Fig. 5 Electricity management of the office room. **a** Electricity management of sockets in 5 working days. **b** Electricity management of lights in 5 working days. **c** Electricity management of air-conditioners in 5 working days

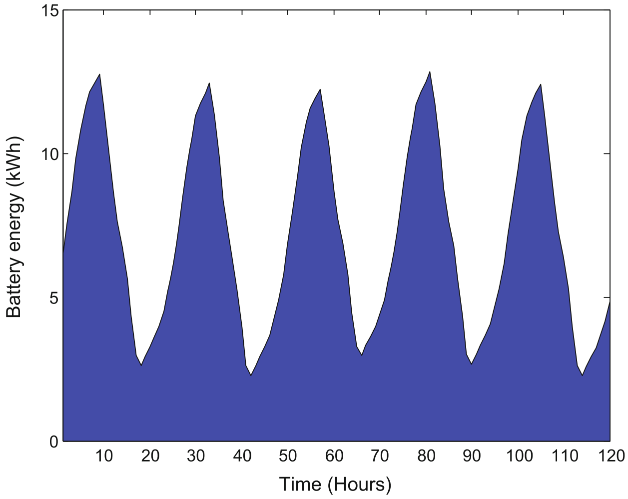

Fig. 6 Battery level of the office room

level is shown in Fig. 6, which indicates proper changes in the level of the battery by using the optimization method. In addition, the total expense on electricity from the grid in the office room in 5 working days, i.e., 120 h, is originally 262.78 cents and reduced to 209.07 cents after optimization with a total saving of 20.44 %.

4.2 Computer room

For the computer room of Room 4 on the 6th floor, which contains some computer equipments including hosts, servers, switches, etc., its original electricity consumption in three types in 5 working days and results of electricity consumption behavior reconstructed by ESNs are shown in Fig. 7. The most remarkable difference of the curves from those of the office room above is shown by the curve of electricity consumption from sockets, which almost remains unchanged during a whole working day, since all the computer equipments using sockets in the room require stable running in 24 h. However, in terms of the curves of electricity consumption from lights and air-conditioners, both of them are almost in the same form as those in the office room, with the "double-peak" characteristic specifically, due to similar working schedules of personnel in the computer room. It is noteworthy that electricity consumption from air-conditioners in the computer room remains at a constant nonzero value at night given the requirement on temperature from the computer equipments.

Then, a battery with the same parameters as the one in the office room is installed in the computer room. Initialized by the same performance index function and neural network parameters, the optimization method based on ADHDP is implemented to improve the electricity consumption in the computer room. Optimal control strategies of the battery in 5 working days are shown in Fig. 8. It can be seen that

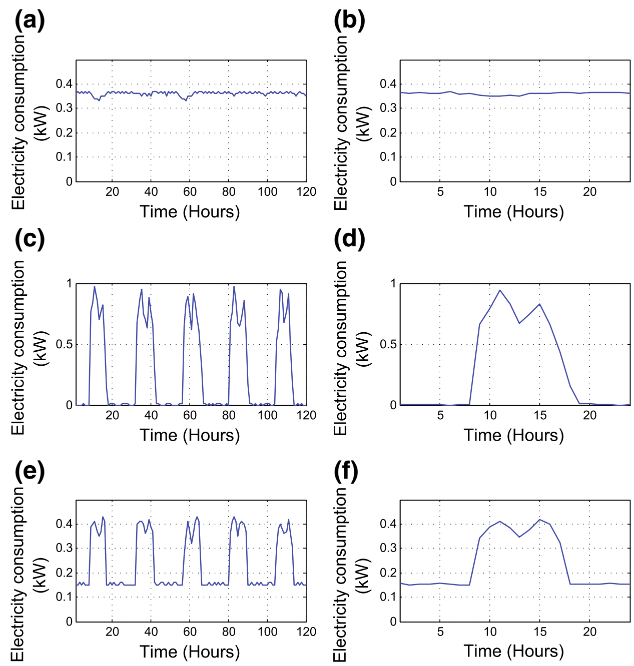

Fig. 7 Electricity consumption of a computer room. **a** Electricity consumption from sockets in 5 working days. **b** Electricity consumption from sockets reconstructed by ESN. **c** Electricity consumption from lights in 5 working days. **d** Electricity consumption from lights reconstructed by ESN. **e** Electricity consumption from air-conditioners in 5 working days. **f** Electricity consumption from air-conditioners reconstructed by ESN

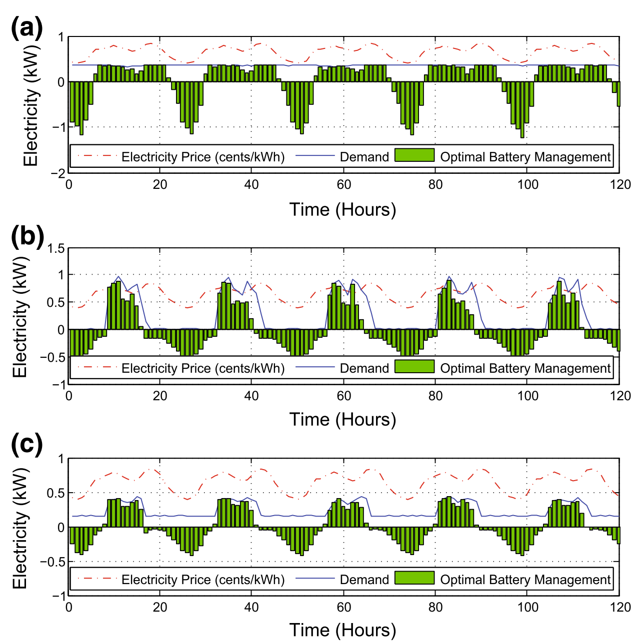

Fig. 8 Electricity management of the computer room. **a** Electricity management of sockets in 5 working days. **b** Electricity management of lights in 5 working days. **c** Electricity management of air-conditioners in 5 working days

except similar control strategies for electricity demand from lights and air-conditioners, the battery in the computer room is charged more intensely when the electricity price is low given the stable electricity demand from sockets. Moreover, the total expense on electricity from the grid in the computer room in 5 working days, i.e., 120 h, is originally 362.71 cents and reduced to 285.14 cents after optimization with a total saving of 21.39 %.

4.3 Storage room

Room 3 on the 13th floor is a storage room, where articles requiring a constant temperature for storage are stored. Its original electricity consumption in three types in 5 working days and results of electricity consumption behavior reconstructed by ESNs are shown in Fig. 9. It can be seen that all the three curves present entirely different characteristics, none of which still takes on the "double-peak" characteristic, but the electricity consumption from air-conditioners remains constant due to the special storage requirements of articles stored inside, while the curves of both two other types of electricity consumption are close to zero for the reason that nobody regularly works in the storage room, thus generally almost no electricity consumption from sockets and lights is incurred.

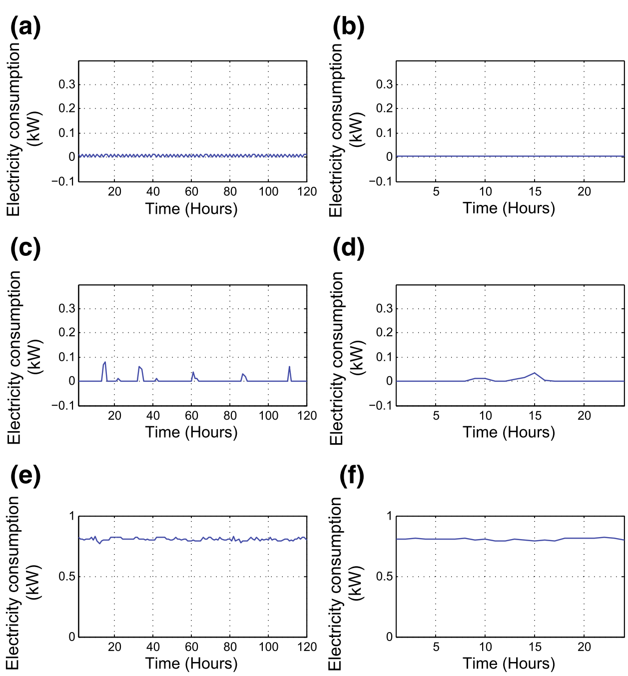

Fig. 9 Electricity consumption of a storage room. **a** Electricity consumption from sockets in 5 working days. **b** Electricity consumption from sockets reconstructed by ESN. **c** Electricity consumption from lights in 5 working days. **d** Electricity consumption from lights reconstructed by ESN. **e** Electricity consumption from air-conditioners in 5 working days. **f** Electricity consumption from air-conditioners reconstructed by ESN

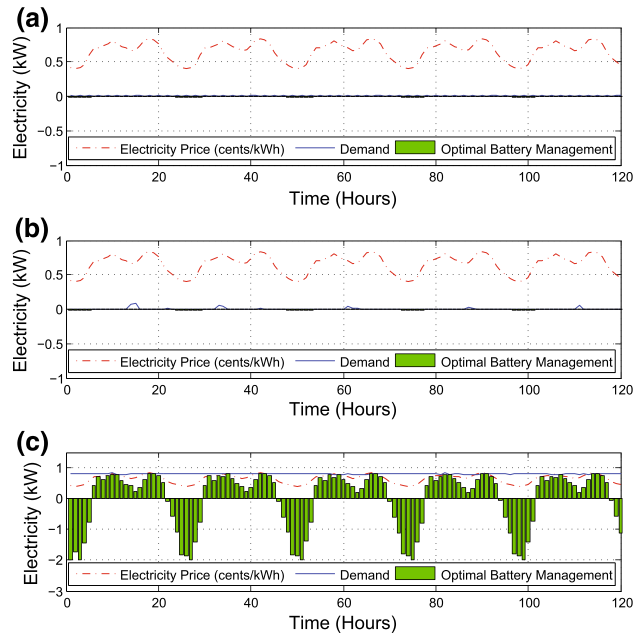

Fig. 10 Electricity management of the storage room. **a** Electricity management of sockets in 5 working days. **b** Electricity management of lights in 5 working days. **c** Electricity management of air-conditioners in 5 working days

Similarly, with a same battery installed, the optimization method based on ADHDP is implemented to improve the electricity consumption in the storage room. Optimal control strategies of the battery in 5 working days are shown in Fig. 10. Given almost no electricity demand from sockets and lights in the storage room, the output of the battery is not linked to the two demands but only satisfies the demand from air-conditioners, and given the similar stable demand from air-conditioners, the battery is intensely charged as well when the electricity price is low during a day. In addition, the total expense on electricity from the grid in the storage room in 5 working days, i.e., 120 h, is originally 315.17 cents and reduced to 243.69 cents after optimization with a total saving of 22.68 %.

4.4 Meeting room

Finally, the meeting room of Room 5 on the 8th floor is given as an example, whose original electricity consumption in three types in 5 working days and results of electricity consumption behavior reconstructed by ESNs are shown in Fig. 11. Since the meeting room is occasionally used without a fixed pattern, we can see that all the three curves of electricity consumption reconstructed by the ESNs are close to zero.

Since the electricity demand from sockets, lights and air-conditioners in the meeting room almost equals zero, it is unnecessary to install a battery in the room and therefore the

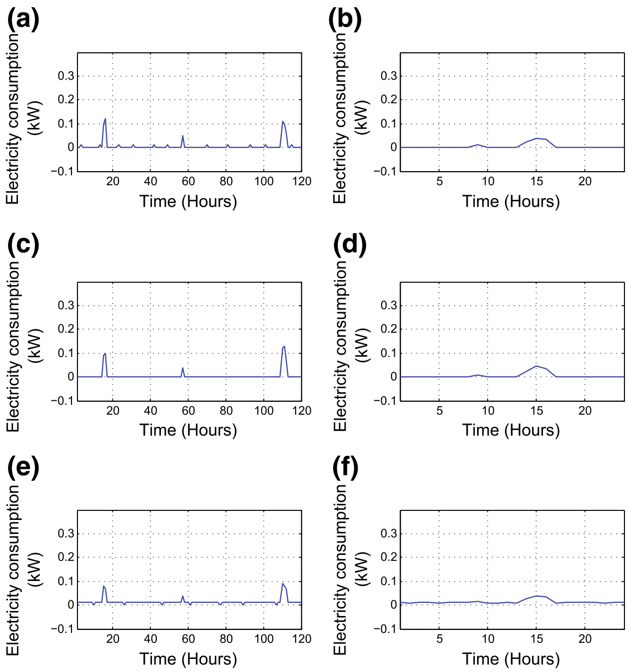

Fig. 11 Electricity consumption of a meeting room. **a** Electricity consumption from sockets in 5 working days. **b** Electricity consumption from sockets reconstructed by ESN. **c** Electricity consumption from lights in 5 working days. **d** Electricity consumption from lights reconstructed by ESN. **e** Electricity consumption from air-conditioners in 5 working days. **f** Electricity consumption from air-conditioners reconstructed by ESN

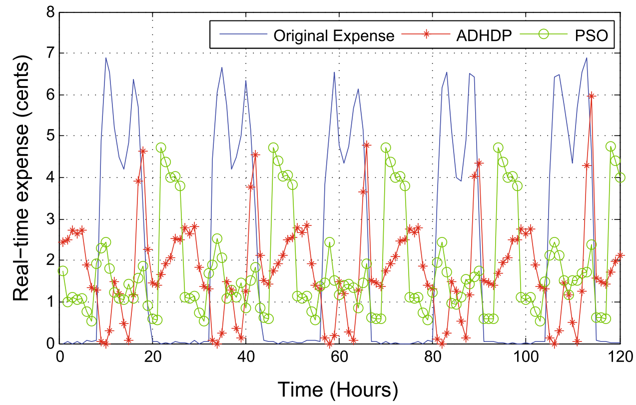

Fig. 12 Real-time expense comparison between ADHDP and PSO algorithms

Table 1 Total expense comparison

	Original	PSO	ADHDP
Total expense (cents)	262.78	220.10	209.07
Savings (%)		16.24	20.44

optimization method becomes meaningless. In other words, the cost of installing batteries in rooms classified as meeting rooms in the office building can be saved.

4.5 Expense comparison

To evaluate the superiority of the developed method, we compare it with the particle swarm optimization (PSO) algorithm (Fuselli et al. 2013) with respect to expense on electricity from the grid in the above-mentioned office room. In the PSO algorithm, each particle naturally moves to an optimal or near-optimal position. Initialized by the swarm size of $\mathcal{G}$, the position of each particle denoted by $x_\ell(t)$, $\ell = 1, 2, \ldots, \mathcal{G}$ and the movement denoted by the velocity vector $v_\ell(t)$, the update rule of the PSO algorithm is expressed as

$$\begin{aligned} x_\ell(t) &= x_\ell(t-1) + v_\ell(t), \\ v_\ell(t) &= \omega v_\ell(t-1) + \phi_1 \rho_1^T (p_\ell - x_\ell(t-1)) \\ &\quad + \phi_2 \rho_2^T (p_g - x_\ell(t-1)), \end{aligned} \tag{17}$$

where the inertia factor $\omega = 0.7$, the correction factors $\rho_1 = \rho_2 = [1, 1]^T$, ϕ_1 and ϕ_2 are randomly initialized in $[0, 1]$, p_ℓ denotes the best position of particles, and p_g denotes the global best position. After both the ADHDP and PSO algorithms are implemented for 50 iterations in the same computer hardware conditions, the comparison of real-time expense between ADHDP and PSO in 5 working days is shown in Fig. 12, and the comparison of total expense within the same period is shown in Table 1, which demonstrates the superiority of the ADHDP algorithm concerned in this paper.

5 Conclusion and future work

Based on a practical office building with rooms classified into office rooms, computer rooms, storage rooms and meeting rooms (Shi et al. 2015), an optimization method based on action-dependent heuristic dynamic programming (ADHDP) is developed to improve the electricity consumption in each category of rooms through optimal battery management. Finally, the total expense on electricity from the power grid can be saved. The developed method is elaborated, and neural networks are employed to implement the method. Practical effect of the developed method is presented with a case study on an office building. In the case study, the total expenses on electricity from the power grid in three selected rooms, i.e., an office room, a computer room and a storage room, are saved by 20.44, 21.39 and 22.68 %, respectively.

However, energy losses during charging/discharging of batteries and the lifetime of batteries are not considered in this paper, but the two issues may reduce the expense saved by the developed optimization algorithm to a certain extent. In future work, we will investigate how to reduce the energy losses by adding thermal insulation materials outside batteries, controlling the charging/discharging power, limiting the maximum and minimum storage energies, adjusting the electrolyte density inside batteries, avoiding pollutants outside batteries, etc., which will also extend the lifetime of batteries.

On the other hand, renewable sources including solar and wind energies may be introduced into the management system to further improve the electricity consumption of rooms in office buildings and reduce expense on electricity from the power grid. Moreover, with more data obtained, we may extend our study to optimizing electricity consumption of the entire building besides the rooms.

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

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Amjadi Z, Williamson SS (2010) Power-electronics-based solutions for plug-in hybrid electric vehicle energy storage and management systems. IEEE Trans Industr Electron 57(2):608–616

Anvari-Moghaddam A, Monsef H, Rahimi-Kian A (2015) Cost-effective and comfort-aware residential energy management under different pricing schemes and weather conditions. Energy Build 86:782–793

Arsuaga-Rios M, Vega-Rodriguez MA (2015) Multi-objective energy optimization in grid systems from a brain storming strategy. Soft Comput 19(11):3159–3172

Bellman RE (1957) Dynamic programming. Princeton University Press, Princeton

Bertsekas DP, Tsitsiklis JN (1996) Neuro-dynamic programming. Athena Scientific, Belmont

Boaro M, Fuselli D, Angelis FD, Liu D, Wei Q, Piazza F (2013) Adaptive dynamic programming algorithm for renewable energy scheduling and battery management. Cogn Comput 5(2):264–277

Data of real-time electricity price from ComEd Company, the United States. [Online]. https://rrtp.comed.com/live-prices/?

Enns R, Si J (2003) Helicopter trimming and tracking control using direct neural dynamic programming. IEEE Trans Neural Netw 14(8):929–939

Fuselli D, Angelis FD, Boaro M, Liu D, Wei Q, Squartini S, Piazza F (2013) Action dependent heuristic dynamic programming for home energy resource scheduling. Int J Electr Power Energy Syst 48(6):148–160

Guerrero JM, Loh PC, Lee TL, Chandorkar M (2013) Advanced control architectures for intelligent microgrids-part II: power quality, energy storage, and AC/DC microgrids. IEEE Trans Industr Electron 60(4):1263–1270

Huang T, Liu D (2013) A self-learning scheme for residential energy system control and management. Neural Comput Appl 22(2):259–269

Jaeger H (2001) The 'echo state' approach to analysing and training recurrent neural networks. German National Research Center for Information Technology, St. Augustin, Germany, Technical Report 148

Jaeger H, Haas H (2004) Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science 304:78–80

Lewis FL, Vrabie D, Vamvoudakis KG (2012) Reinforcement learning and feedback control: using natural decision methods to design optimal adaptive controllers. IEEE Control Syst 32(6):76–105

Li C, Zhang G, Wang M, Yi J (2013) Data-driven modeling and optimization of thermal comfort and energy consumption using type-2 fuzzy method. Soft Comput 17(11):2075–2088

Liu D, Wei Q (2012) An iterative ε-optimal control scheme for a class of discrete-time nonlinear systems with unfixed initial state. Neural Netw 32:236–244

Ma K, Hu G, Spanos CJ (2014) Distributed energy consumption control via real-time pricing feedback in smart grid. IEEE Trans Control Syst Technol 22(5):1907–1914

Molina D, Venayagamoorthy GK, Liang J, Harley RG (2013) Intelligent local area signals based damping of power system oscillations using virtual generators and approximate dynamic programming. IEEE Trans Smart Grid 4(1):498–508

Na J, Herrmann G (2014) Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous-time unknown nonlinear systems. IEEE/CAA J Autom Sin 1(4):412–422

Ni Z, He H (2013) Heuristic dynamic programming with internal goal representation. Soft Comput 17(11):2101–2108

Ni Z, He H, Wen J (2013) Adaptive learning in tracking control based on the dual critic network design. IEEE Trans Neural Netw Learn Syst 24(6):913–928

Prokhorov DV, Wunsch DC (1997) Adaptive critic designs. IEEE Trans Neural Netw 8(5):997–1007

Severini M, Squartini S, Piazza F (2013) Hybrid soft computing algorithmic framework for smart home energy management. Soft Comput 17(11):1983–2005

Shi G, Wei Q, Liu Y, Guan Q, Liu D (2015) Data-driven room classification for office buildings based on echo state network. The 2015 27th Chinese control and decision conference (CCDC), IEEE, pp 2602–2607

Si J, Wang YT (2001) On-line learning control by association and reinforcement. IEEE Trans Neural Netw 12(2):264–276

Song R, Xiao W, Wei Q (2013) Multi-objective optimal control for a class of nonlinear time-delay systems via adaptive dynamic programming. Soft Comput 17(11):2109–2115

Wang F, Zhang H, Liu D (2009) Adaptive dynamic programming: an introduction. Comput Intell 4(2):39–47

Wei Q, Liu D (2014) Adaptive dynamic programming for optimal tracking control of unknown nonlinear systems with application to coal gasification. IEEE Trans Autom Sci Eng 11(4):1020–1036

Wei Q, Wang F, Liu D, Yang X (2014) Finite-approximation-error based discrete-time iterative adaptive dynamic programming. IEEE Trans Cybern 44(12):2820–2833

Wei Q, Liu D, Shi G (2015a) A novel dual iterative Q-learning method for optimal battery management in smart residential environments. IEEE Trans Industr Electron 62(4):2509–2518

Wei Q, Liu D, Shi G, Liu Y (2015b) Multibattery optimal coordination control for home energy management systems via distributed iterative adaptive dynamic programming. IEEE Trans Industr Electron 62(7):4203–4214

Wei Q, Liu D, Xu Y (2016a) Neuro-optimal tracking control for a class of discrete-time nonlinear systems via generalized value iteration adaptive dynamic programming approach. Soft Comput 20(2):697–706

Wei Q, Liu D, Liu Y, Song R (2016b) Optimal constrained self-learning battery sequential management in microgrids via adaptive dynamic programming. IEEE/CAA J Autom Sin (accepted)

Wei Q, Lewis FL, Sun Q, Yan P, Song R (2016c) Discrete-time deterministic Q-learning: a novel convergence analysis. IEEE Trans Cybern. doi:10.1109/TCYB.2016.2542923
Werbos PJ (1977) Advanced forecasting methods for global crisis warning and models of intelligence. General Syst Yearb 22:25–38
Werbos PJ (1991) A menu of designs for reinforcement learning over time. In: Miller WT, Sutton RS, Werbos PJ (eds) Neural networks for control. MIT Press, Cambridge
Xu H, Jagannathan S (2013) Stochastic optimal controller design for uncertain nonlinear networked control system via neuro dynamic programming. IEEE Trans Neural Netw Learn Syst 24(3):471–484
Xu B, Yang C, Shi Z (2014a) Reinforcement learning output feedback NN control using deterministic learning technique. IEEE Trans Neural Netw Learn Syst 25(3):635–641
Xu X, Lian C, Zuo L, He H (2014b) Kernel-based approximate dynamic programming for real-time online learning control: an experimental study. IEEE Trans Control Syst Technol 22(1):146–156
Zhao Q, Xu H, Jagannathan S (2014) Near optimal output feedback control of nonlinear discrete-time systems based on reinforcement neural network learning. IEEE/CAA J Autom Sin 1(4):372–384