An Adaptive Dynamic Programming Based Method for Optimization of Electricity Consumption in Office Buildings

Guang Shi, Qinglai Wei
The State Key Laboratory of Management
and Control for Complex Systems
Institute of Automation
Chinese Academy of Sciences
Beijing 100190, China
Email: shiguang2012@ia.ac.cn, qinglai.wei@ia.ac.cn

Derong Liu
School of Automation and Electrical Engineering
University of Science and Technology Beijing
Beijing 100083, China
Email: derong@ustb.edu.cn

***Abstract*—In this paper, an adaptive dynamic programming (ADP) based method is developed to optimize electricity consumption of rooms in office buildings through optimal battery management. Rooms in office buildings are generally divided into office room, computer room, storage room, meeting room, etc., each of which has different characteristics of electricity consumption, as divided into electricity consumption from sockets, lights and air-conditioners in this paper. The developed method based on ADP is elaborated, 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 show the practical effect of the developed method.**

I. 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 of energy consumption. Recently extensive research has been conducted in both theory and practice. In [1], a distributed energy control algorithm was proposed for energy consumers to control their energy consumption. Cau et al. [2] presented a novel energy management strategy to control an isolated microgrid powered by a photovoltaic array and a wind turbine and equipped with two different energy storage systems. In [3], a system involving renewable energy for smart home energy management was developed to optimize home energy consumption. With in-depth development of smart grid, increasing intelligence is required in the design of efficient energy management systems. Therefore, optimal battery management becomes an important approach to saving expense on power in smart grid.

Proposed by Werbos [4], [5], adaptive dynamic programming (ADP) [6], [7] has been verified with strong ability to solve the optimization problem of complex non-linear systems by means of its strong self-learning capacity. The method of ADP circumvents the "curse of dimensionality" in dynamic programming (DP) by using the forward-in-time approach to solve the Hamilton-Jacobi-Bellman equation [5]. Recent years witnessed extensive research on ADP [9]–[11]. In [5], 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 [8], 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 [12], [13], 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 has been focused on residential energy systems [12]–[16], rather than energy consumption in office buildings. Nevertheless, as a significant component of urban structure, office buildings take up 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 [17], [18]. Therefore, it is of great importance to improve the electricity consumption of office buildings based on electricity storage.

In our previous work [19], a data-driven classification method based on echo state network (ESN) is developed to classify rooms in office buildings into different categories, including office room, computer room, storage room and meeting room. 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 Section II. The

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

978-1-5090-0620-5/16/$31.00 ©2016 IEEE

developed optimization algorithm of electricity consumption based on ADP is elaborated in Section III. In Section IV, a detailed case study is presented to show the effectiveness and superiority of the developed algorithm. Finally, the conclusion is drawn in Section V.

II. 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.

A. Electricity Consumption Management System of a Room in an Office Building

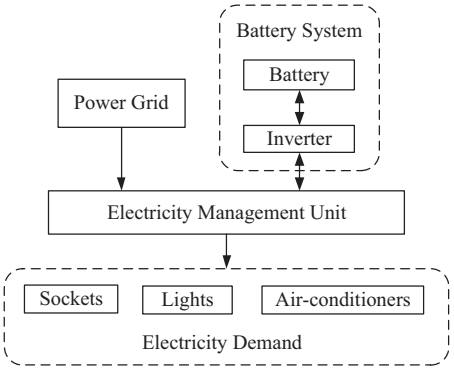

Fig. 1. Electricity consumption management system of a room in an office building.

As shown in Fig. 1, the electricity consumption management system of a room in an office building 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 [12], 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 hour, $P_b(t)$ 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.

B. 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-hour 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) shall 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 price, 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 equation (3) and other relevant conditions. 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 $\gamma_a(t) = P_{La}(t)/P_L(t)$.

III. 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.

A. Adaptive Dynamic Programming

In accordance with Bellman's principle of optimality [20], the method of dynamic programming 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 depends on the initial state $x(l)$ and the initial time l. Dynamic programming aims to obtain a series of control actions $u(l)$, $l=t,t+1,\cdots$, which minimize the performance index function in (6). Based on Bellman's principle of optimality [20], the optimal performance meets the Hamilton-Jacobi-Bellman (HJB) equation as follows

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

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

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

ADP is a method based on the iteration between policy improvement and value approximation of dynamic programming solutions. In general, four major schemes of ADP are available, i.e., heuristic dynamic programming (HDP), dual heuristic programming (DHP) and their action-dependent versions. The scheme of ADP concerned in this paper is named action-dependent heuristic dynamic programming (ADHDP), where the model network is not explicitly required in the design. Next, the design of ADHDP will be elaborated.

B. Action-Dependent Heuristic Dynamic Programming

To solve the optimal control problem concerned in this paper, the method of ADHDP is adopted, because an explicit model network is not required. Fig. 2 shows a typical scheme of ADHDP, in which the critic network is trained to minimize the following error:

$$E_q=\sum_t E_q(t)=\sum_t[Q(t-1)-U(t)-\gamma Q(t)]^2, \tag{9}$$

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

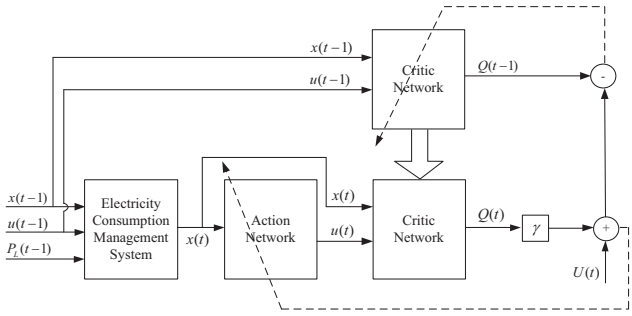

Fig. 2. A typical scheme of ADHDP.

$$Q(t)=Q[x(t),u(t)], \tag{10}$$

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

If $E_q(t)=0$ at all time t, it is impled by (9) that

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

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

$$\left\{\begin{array}{c}x(t-1)\\u(t-1)\end{array}\right\}\rightarrow\{Q(t-1)\}, \tag{12}$$

where $x(t-1)$ and $u(t-1)$ denote the input 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

$$Q(t-1)\approx U(t)+\gamma Q(t), \tag{13}$$

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

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

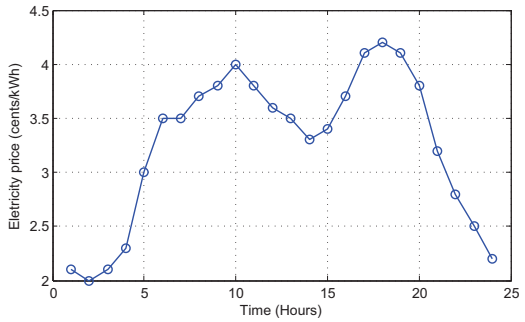

Fig. 3. Typical electricity price in non-summer seasons.

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

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.

IV. 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. 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.

In our previous work [19], a data-driven classification method based on echo state network (ESN) is developed to classify rooms in office buildings into different categories, including office room, computer room, storage room and meeting room. Based on the results in [19], we apply the method developed in this paper 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 United States from ComEd Company in [22]. Combined with the real-time electricity price shown in Fig. 3, results of different categories of rooms are respectively presented as follows.

A. Office Room

As given in [19], Room 3 on the 4th floor is an office room. Based on the electricity demand and real-time electricity price in 5 working days, the electricity consumption of the office room is optimized with the developed method based on ADHDP, and optimal control strategies of the battery are shown in Fig. 4. We can see that all the three types of electricity consumption present a typical "double-peak" characteristic, namely, all the three types of electricity consumption reach their peak values in mid-morning around 11:00 and mid-afternoon around 16:00 on a working day, and achieve a low point at noon since part of personnel working in the office room usually go out for lunch then, who may switch off some of their electrical appliances using sockets, turn off some of the lights or adjust the temperature set for the air-conditioners, while some others who have their lunch inside the rooms may still consume some electricity. After optimization by the developed method, 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. In addition, the total expense on electricity from the grid in the office room for 5 days is originally 262.78 cents and reduced to 209.07 cents after optimization with a total saving of 20.44%.

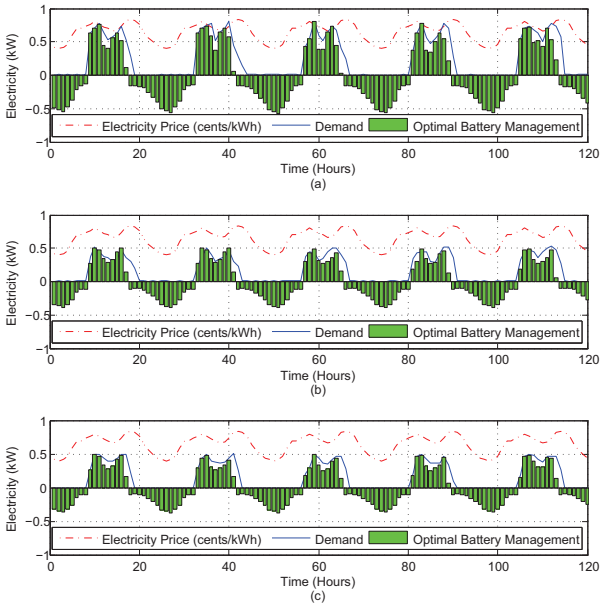

Fig. 4. Electricity management of the office room. (a) Electricity management of sockets in 5 working days. (b) Electricity management of lights in 5 working days. (c) Electricity management of air-conditioners in 5 working days.

B. Computer Room

For the computer room of Room 4 on the 6th floor, which contains some hosts, servers, switches and other computer equipments, the electricity consumption of the room is optimized with the developed method based on the electricity demand and real-time electricity price in 5 working days, and optimal control strategies of the battery are shown in Fig. 5. We can see that the most remarkable difference of the curves from those of the office room above is indicated 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 hours. However, both of the curves of electricity

consumption from lights and air-conditioners are almost in the same form as those in the office room, with "double-peak" characteristic specifically, due to the similar working schedules of personnel in the computer room. After optimization by the developed method, 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 for 5 days is originally 362.71 cents and reduced to 285.14 cents after optimization with a total saving of 21.39%.

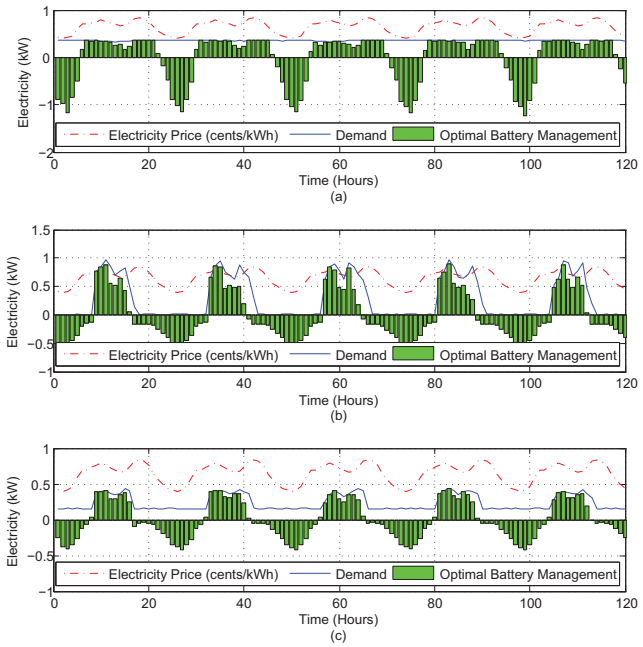

Fig. 5. Electricity management of the computer room. (a) Electricity management of sockets in 5 working days. (b) Electricity management of lights in 5 working days. (c) Electricity management of air-conditioners in 5 working days.

C. Storage Room

Room 3 on the 13th floor is a storage room, where articles requiring a constant temperature for storage, are stored. Similarly, with a same battery installed, the optimization method based on ADHDP is implemented to improve the electricity consumption in the storage room, and optimal control strategies of the battery in 5 working days are shown in Fig. 6. It can be seen that all the three curves display entirely different characteristics, none of which still takes on the "double-peak" characteristic, but the electricity consumption from air-conditioners keeps at a constant level 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. 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 for 5 days is originally 315.17 cents and reduced to 243.69 cents after optimization with a total saving of 22.68%.

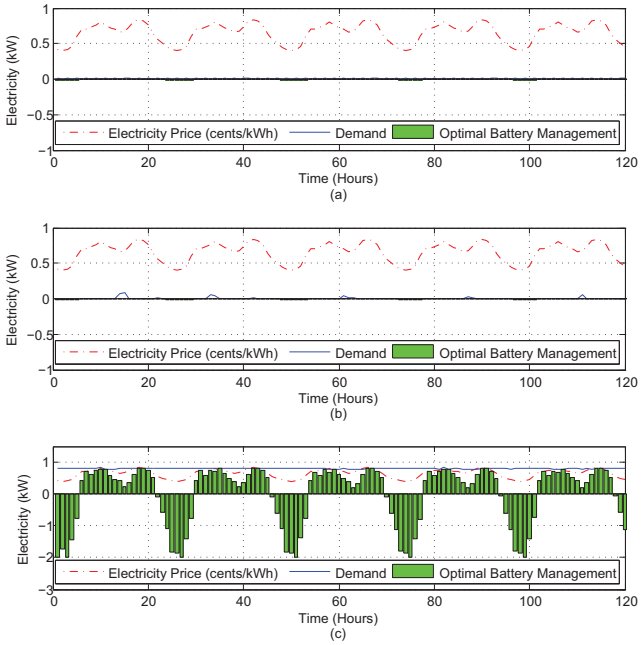

Fig. 6. Electricity management of the storage room. (a) Electricity management of sockets in 5 working days. (b) Electricity management of lights in 5 working days. (c) Electricity management of air-conditioners in 5 working days.

D. Meeting Room

Finally, the meeting room of Room 5 on the 8th floor is chosen. Since the meeting room is occasionally used without a fixed pattern, all the three demands of electricity consumption are close to zero except when meetings are held inside the room. Therefore, it is unnecessary to install a battery in the room and therefore the 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.

E. Expense Comparison

To evaluate the superiority of the developed method, we compare it with the particle swarm optimization (PSO) algorithm [14] 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

$$\begin{aligned} x_\ell(t) &= x_\ell(t-1) + \nu_\ell(t), \\ \nu_\ell(t) &= \omega\nu_\ell(t-1) + \phi_1\rho_1^T(p_\ell - x_\ell(t-1)) \\ &\quad + \phi_2\rho_2^T(p_g - x_\ell(t-1)), \end{aligned} \tag{15}$$

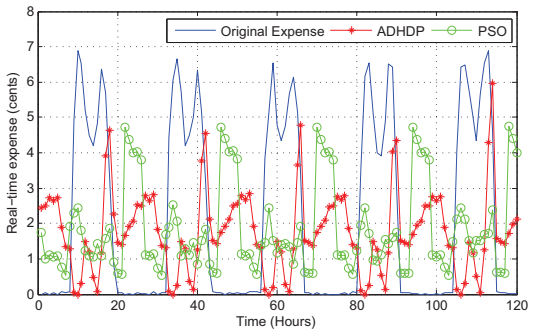

Fig. 7. Real-time expense comparison between ADHDP and PSO algorithms.

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. The comparison of real-time expense between ADHDP and PSO in 5 working days is shown in Fig. 7, and the comparison of total expense within the same period is shown in Table I, which demonstrates the superiority of the ADHDP algorithm concerned in this paper.

TABLE I. TOTAL EXPENSE COMPARISON

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

V. CONCLUSION

Based on a practical office building, where rooms are classified into office room, computer room, storage room and meeting room [19], and electricity consumption in each room is divided into electricity consumption from sockets, lights and air-conditioners, and combined with real-time electricity price, an optimization method based on action-dependent heuristic dynamic programming (ADHDP) is developed to improve the electricity consumption of each category of room by means of optimal battery management, thus saving the total expense on electricity from the power grid. The developed method is elaborated, and neural networks are employed to implement the method. Finally, practical effect of the developed method is presented with a case study on an office building.

REFERENCES

[1] K. Ma, G. Hu, and C. J. Spanos, "Distributed energy consumption control via real-time pricing feedback in smart grid," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 5, pp. 1907–1914, Sep. 2014.

[2] G. Cau, D. Cocco, M. Petrollese, S. K. Kaer, and C. Milan, "Energy management strategy based on short-term generation scheduling for a renewable microgrid using a hydrogen storage system," *Energy Conversion and Management*, vol. 87, pp. 820–831, Nov. 2014.

[3] J. Han, C. Choi, W. Park, I. Lee, and S. Kim, "Smart home energy management system including renewable energy based on ZigBee and PLC," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 2, pp. 198–202, May 2014.

[4] P. J. Werbos, "Advanced forecasting methods for global crisis warning and models of intelligence," *General Systems Yearbook*, vol. 22, pp. 25–38, 1977.

[5] P. J. Werbos, "A menu of designs for reinforcement learning over time," in *Neural Networks for Control*, W. T. Miller, R. S. Sutton and P. J. Werbos, Eds. Cambridge: MIT Press, 1991, pp. 67–95.

[6] Q. Wei and D. Liu, "Adaptive dynamic programming for optimal tracking control of unknown nonlinear systems with application to coal gasification," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 4, pp. 1020–1036, Oct. 2014.

[7] M. Fairbank, E. Alonso, and D. Prokhorov, "An equivalence between adaptive dynamic programming with a critic and backpropagation through time," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 12, pp. 2088–2100, Dec. 2013.

[8] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs," *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 997–1007, Sep. 1997.

[9] F. Wang, N. Jin, D. Liu, and Q. Wei, "Adaptive dynamic programming for finite-horizon optimal control of discrete-time nonlinear systems with ε-error bound," *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 24–36, Jan. 2011.

[10] Q. Wei, H. Zhang, and J. Dai, "Model-free multiobjective approximate dynamic programming for discrete-time nonlinear systems with general performance index functions," *Neurocomputing*, vol. 72, no. 7-9, pp. 1839–1848, Mar. 2009.

[11] D. Liu and Q. Wei, "An iterative ε-optimal control scheme for a class of discrete-time nonlinear systems with unfixed initial state," *Neural Networks*, vol. 32, pp. 236–244, 2012.

[12] T. Huang and D. Liu, "A self-learning scheme for residential energy system control and management," *Neural Computing and Applications*, vol. 22, no. 2, pp. 259–269, Feb. 2013.

[13] M. Boaro, D. Fuselli, F. D. Angelis, D. Liu, Q. Wei, and F. Piazza, "Adaptive dynamic programming algorithm for renewable energy scheduling and battery management," *Cognitive Computation*, vol. 5, no. 2, pp. 264–277, Jun. 2013.

[14] D. Fuselli, F. D. Angelis, M. Boaro, D. Liu, Q. Wei, S. Squartini, and F. Piazza, "Action dependent heuristic dynamic programming for home energy resource scheduling," *International Journal of Electrical Power and Energy Systems*, vol. 48, pp. 148–160, Jun. 2013.

[15] Q. Wei, D. Liu, and G. Shi, "A novel dual iterative Q-learning method for optimal battery management in smart residential environments," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2509–2518, Apr. 2015.

[16] Q. Wei, D. Liu, G. Shi, and Y. Liu, "Multibattery optimal coordination control for home energy management systems via distributed iterative adaptive dynamic programming," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, pp. 4203–4214, Jul. 2015.

[17] Z. Amjadi and S. S. Williamson, "Power-electronics-based solutions for plug-in hybrid electric vehicle energy storage and management systems," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 2, pp. 608–616, Feb. 2010.

[18] J. M. Guerrero, P. C. Loh, T. L. Lee, and M. Chandorkar, "Advanced control architectures for intelligent microgrids-part II: power quality, energy storage, and AC/DC microgrids," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 4, pp. 1263–1270, Apr. 2013.

[19] G. Shi, Q. Wei, Y. Liu, Q. Guan, and D. Liu, "Data-driven room classification for office buildings based on echo state network," in *Proceedings of 27th Chinese Control and Decision Conference*, Qingdao, China, May 2015.

[20] R. E. Bellman, *Dynamic Programming*. Princeton, New Jersey: Princeton University Press, 1957.

[21] J. Si and Y.-T. Wang, "On-line learning control by association and reinforcement," *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 264–276, Mar. 2001.

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