

A self-learning scheme for residential energy system control and management

Ting Huang · Derong Liu

Received: 19 February 2011 / Accepted: 12 July 2011
© Springer-Verlag London Limited 2011

Abstract In this paper, we apply intelligent optimization method to the challenge of intelligent price-responsive management of residential energy use, with an emphasis on home battery use connected to the power grid. For this purpose, a self-learning scheme that can learn from the user demand and the environment is developed for the residential energy system control and management. The idea is built upon a self-learning architecture with only a single critic neural network instead of the action-critic dual network architecture of typical adaptive dynamic programming. The single critic design eliminates the iterative training loops between the action and the critic networks and greatly simplifies the training process. The advantage of the proposed control scheme is its ability to effectively improve the performance as it learns and gains more experience in real-time operations under uncertain changes of the environment. Therefore, the scheme has the adaptability to obtain the optimal control strategy for different users based on the demand and system configuration. Simulation results demonstrate that the proposed scheme can financially benefit the residential customers with the minimum electricity cost.

Keywords Adaptive dynamic programming (ADP) · Residential energy management · Battery storage · Scheduling · Price-responsive management · Real-time pricing · Smart grid

1 Introduction

Over the last decade, the human beings have become more and more dependent on the electricity for their daily life. The rising cost, the environmental concerns, and the reliability issues all underlie the needs and the opportunities for developing new intelligent control and management system of residential hybrid energy usage. There has been considerable discussion of the importance of distributed energy storage, including batteries in the home, as a way to create more price-responsive demand and as a way to integrate more renewable energy resources more effectively into power grids. It is envisioned that distributed energy storage technologies could reduce the combustion of fossil fuel, supply reliable energy in concert with other energy sources and financially benefit residential customers.

The development of an intelligent power grid, i.e., the smart grid, has attracted significant amount of attention recently. Considerable research and development activities have been carried out in both industry and academia [1, 17, 20, 23, 31, 33, 39, 41]. Along with the development of smart grid, more and more intelligence has been required in the design of the residential energy management system. Smart residential energy management system provides end users the optimal management of energy usage by means of robust communication capability, smart metering and advanced optimization technology.

There are extensive research efforts in adaptive dynamic programming (ADP) in the past three decades [3, 4, 7, 14,

This work was supported by the National Science Foundation under Grant ECCS-1027602.

T. Huang · D. Liu (✉)
Department of Electrical and Computer Engineering,
University of Illinois at Chicago, Chicago, IL 60607, USA
e-mail: dliu@ece.uic.edu

T. Huang
e-mail: thuang20@uic.edu

15, 24, 25, 35, 37, 42, 43, 46, 47]. ADP is defined as designs that approximates dynamic programming solutions in the general case, i.e., approximates optimal control over time in nonlinear environments. Many practical problems in real world can be formulated as cost minimization problems, such as energy optimization, error minimization and minimum time controls. Dynamic programming which provides truly optimal solutions to these problems is very useful. However, due to the “curse of dimensionality” [5], for real-world multidimensional problems it is often computationally intractable to run the backward numerical process required to obtain the dynamic programming solutions. Over the years, progress has been made to circumvent the “curse of dimensionality” by building a system, called “critic,” to approximate the cost function in dynamic programming. The idea is to approximate dynamic programming solutions by using a function approximation structure such as neural networks to approximate the cost function. Interest in ADP has grown in the power sector, and a few applications have appeared for generators and grid management [29, 32, 36, 38], which focus mainly on the industrial customers. Very few research address the issues of residential energy system control and management. This would be the first application to price-responsive residential demand of any kind.

The main focus of this paper is on proposing a computationally feasible and self-learning optimization-based optimal operating control scheme for the residential energy system with batteries. We aim to minimize the total operating cost over the scheduling period in a residential household by optimally scheduling the operation of batteries, while satisfying a set of constraints imposed by the requirements on the system and the capacities of individual components of the system. Operational scheduling of storage resources in the power system has been the subject of many studies. The simplest and most straightforward strategies are predefined rule-based [6, 11, 19, 34]. A set of IF–THEN rules are created according to the corresponding scenarios. When a specific scenario happens, the operating strategy employs some predetermined rules. Rule-based strategies are relatively simple and can be adapted to a lot of scenarios. However, limitation is obvious that every scenario has to be considered in advance, which is not practical, especially for large complex systems. A large number of more complex optimization techniques have been applied to solve this problem, such as dynamic programming [2, 27, 30, 45], linear programming [10, 13], Lagrange relaxation [28] and nonlinear programming [40]. These techniques aim at reducing either computation time or memory requirements. Recently, computational intelligence methodologies including fuzzy optimization, genetic algorithm, simulated annealing method and particle swarm optimization approach have been employed to deal with

the operation cost of hybrid energy systems with storage systems [8, 9, 16, 21, 44]. Generally, these heuristic approaches can provide a reasonable solution. However, these approaches are not able to adapt to frequent and swift load changes and real-time pricing due to their static nature. Therefore, we develop in the present paper an operational scheme with self-learning ability and adaptability to optimize residential energy systems according to system configurations and user demand. The self-learning scheme based on ADP has the capability to learn from the environment and the residential demand so that the performance of the algorithm will be improved through further learning.

This paper is organized as follows. In Sect. 2, the residential energy system used in the paper is briefly described. The control and management problem of residential energy system is formulated. In Sect. 3, the ADP scheme that is suitable for the application to the residential energy system control and management problem is introduced. In Sect. 4, our self-learning control algorithm for grid-connected energy system in residential households is developed. The present work will assume the use of artificial neural networks as a means for function approximation in the implementation of ADP. In particular, multilayer feedforward neural networks are considered, even though other types of neural networks are also applicable in this case. In Sect. 5, the performance of our algorithm is studied through simulations. The simulation results indicate that the proposed self-learning algorithm is effective in achieving the optimal cost. Finally, in Sect. 6, the paper will be concluded with a few remarks.

2 Description of the residential energy system

The objective of this paper is to apply ADP intelligent optimization method to the challenge of intelligent price-responsive management of residential energy use. Specifically, it is to minimize the sum of system operational cost over the scheduling period, subject to technological and operational constraints of grids and storage resource generators and subject to the system constraints such as power balance and reliability. For this purpose, we focus our research on finding the optimal battery charge/discharge strategy of the residential energy system with batteries and power grids configuration.

2.1 Residential energy system

The residential energy system uses AC utility grid as the primary source of electricity and is intended to operate in parallel with the battery storage system. Figure 1 depicts the schematic diagram of a residential energy system. The system consists of power grids, a sinewave inverter,

a battery system and a power management unit. The battery storage system is connected to power management system through an inverter. The inverter functions as both charger and discharger for the battery. The construction of the inverter is based upon power MOSFET technology and pulse-width modulation technique [16]. The quality of the inverter output is comparable to that delivered from the power grids. The battery storage system consists of lead acid batteries, which are the most commonly used rechargeable battery type. The optimum battery size for a particular residential household can be obtained by performing various test scenarios, which is beyond the scope of the present paper. Generally, the battery is sized to enable it to supply power to the residential load for a period of 12 h.

There are three operational modes for the residential energy system under consideration.

1. Charging mode: when system load is low and the electricity price is inexpensive, the power grids will supply the residential load directly and, at the same time, charge the batteries.
2. Idle mode: the power grids will directly supply the residential load at certain hours when, from the economical point of view, it is more cost-effective to use the fully charged batteries in the evening peak hours.
3. Discharging mode: by taking the subsequent load demands and time-varying electricity rate into account, batteries alone supplies the residential load at hours when the cost of grid power is high.

This system can easily be expanded; i.e., other power sources along with the power grid and batteries like PV panels or wind generators can be integrated into the system when they are available.

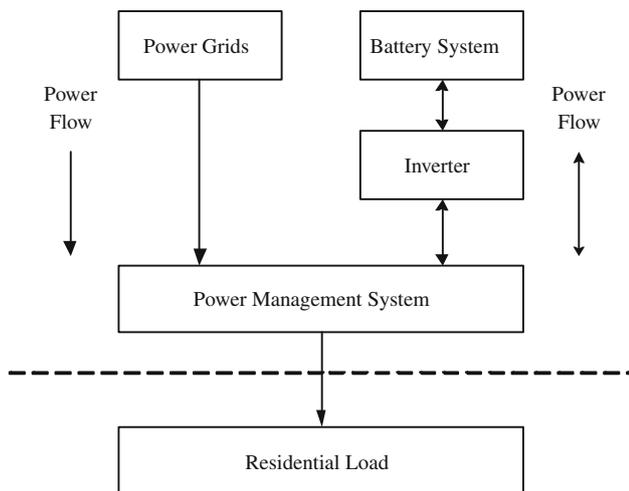


Fig. 1 Grid-connected residential energy system with battery storage

2.2 Load profile

For this study, the optimal scheduling problem is treated as a discrete time problem with the time step as 1-h and it is assumed that the residential load over each hourly time step is varying with noise. Thus, the daily load profile is divided into 24 h periods to represent each hour of the day. Each day can be divided into a greater number of periods to have higher resolution. However, for simplicity and agreement with existing literature [2, 9, 13, 28], we use a 24 h period each day in this work. A typical weekday load profile is shown in Fig. 2. The load factor P_L is expressed as $P_L(t)$ during hour t ($t = 1, 2, \dots, 24$). For instance, at time $t = 19$, the load is 7.8 kW which would require 7.8 kWh of energy. Since the load profile is divided into 1 h steps, the units of the power of energy sources can be represented equally by kW or kWh.

2.3 Real-time pricing

Residential real-time pricing is one of the load management policies used to shift electricity usage from peak load hours to light load hours in order to improve power system efficiency and allow new power system construction projects [21]. With real-time pricing, the electricity rate varies from hour to hour based on wholesale market prices. Hourly, market-based electricity prices typically change as the demand for electricity changes; higher demand usually means higher hourly prices. In general, there tends to be a small price spike in the morning and another slightly larger spike in the evening when the corresponding demand is high. Figure 3 demonstrates a typical daily real-time pricing from [12]. The varying electricity rate is expressed as

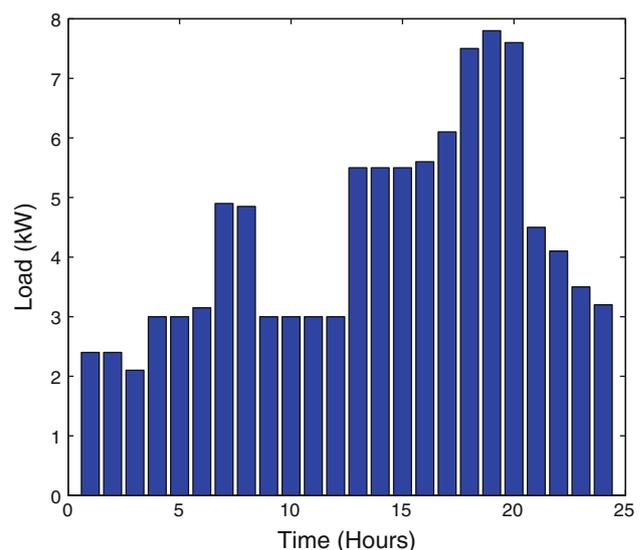


Fig. 2 A typical residential load profile

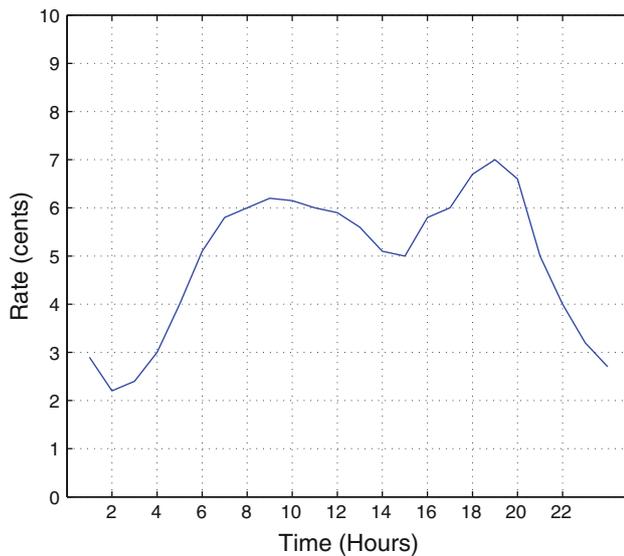


Fig. 3 A typical daily real-time pricing

$C(t)$, the energy cost during the hour t in cents. For the residential customer with real-time pricing, energy charges are functions of the time of electricity use. Therefore, for the situation where batteries are charged during the low rate hours and discharged during high rate hours, one may expect, from an economical point of view, the profits will be made by storing energy during low rate hours and releasing it during the high rate hours. In this way, the battery storage system can be used to reduce the total electricity cost for residential household.

2.4 Battery model

The energy stored in a battery can be expressed as [21, 48]:

$$E_b(t) = E_{b0} - \sum_{i=0}^t P_b(i) \quad (1)$$

$$P_b(i) = V_0 I \alpha(i) \quad (2)$$

$$\alpha(i) = \begin{cases} 1 & (i \leq i_0) \\ K_1(I) dV_0/di & (i > i_0) \end{cases} \quad (3)$$

$$V_0 = V_s - (K_c(Q/(Q - Ji) + N)J + A \exp(-BQ^{-1}Ji)) \quad (4)$$

where $E_b(t)$ is the battery energy at time t , E_{b0} is the peak energy level when the battery is fully charged (capacity of the battery), $P_b(i)$ is the battery power output at time i , V_0 is the terminal voltage of the battery, I is the battery discharge current, $\alpha(i)$ is the current weight factor as a function of discharge time, i_0 is the battery manufacturer specified length of time for constant power output under constant discharge current rate, $K_1(I)$ is the weight factor as a function of the magnitude of the current, V_s is the battery

internal voltage, K_c is the polarization coefficient (ohm cm^2), Q is the available amount of active material (coulombs per cm^2), J is the apparent current density (amperes per cm^2), N is the internal resistance per cm^2 , and A and B are constants.

Apart from the battery itself, the loss of other equipments such as inverters, transformers and transmission lines should also be considered in the battery model. The efficiency of these devices was derived in [48] as:

$$\eta(P_b(t)) = 0.898 - 0.173 |P_b(t)|/P_{\text{rate}}, \quad P_{\text{rate}} > 0 \quad (5)$$

where P_{rate} is the rated power output of the battery, $\eta(P_b(t))$ is the total efficiency of all the auxiliary equipments in the battery system.

Assume that all the losses caused by these equipments occur during the charging period. The battery model used in this work is expressed as follows: when the battery is charged

$$\begin{aligned} E_b(t+1) &= E_b(t) - P_b(t+1) \times \eta(P_b(t+1)), \\ P_b(t+1) &< 0 \end{aligned} \quad (6)$$

and when the battery discharges

$$E_b(t+1) = E_b(t) - P_b(t+1), \quad P_b(t+1) > 0. \quad (7)$$

In general, to improve battery efficiency and extend the battery's lifetime as far as possible, two constraints need to be considered:

1. Battery has storage limit. A battery lifetime may be reduced if it operated at lower amount of charge. In order to avoid damage, the energy stored in the battery must always meet constraint as follows:

$$E_b^{\min} \leq E_b(t) \leq E_b^{\max}. \quad (8)$$

2. For safety, battery cannot be charged or discharged at rate exceeding the maximum and minimum values to prevent damage. This constraint represents the upper and lower limit for the hourly charging and discharging power. A negative $P_b(t)$ means that the battery is being charged, while a positive $P_b(t)$ means the battery is discharging,

$$P_b^{\min} \leq P_b(t) \leq P_b^{\max}. \quad (9)$$

2.5 Load balance

At any time, the sum of the power from the power grids and the batteries must be equal to the demand of residential user

$$P_L(t) = P_b(t) + P_g(t), \quad (10)$$

where $P_g(t)$ is the power from the power grids, $P_b(t)$ can be positive (in the case of batteries discharging) or negative (batteries charging) or zero (idle). It explains the fact that the power generation (power grids and batteries) must

balance the load demand for each hour in the scheduling period. We assume here that the supply from power grids is enough for the residential demand.

2.6 Optimization objectives

The objective of the optimization policy is, given the residential load profile and real-time pricing, to find the optimal battery charge/discharge/idle schedule at each time step which minimize the total cost

$$C_T = \sum_{t=1}^T C(t) \times P_g(t) \tag{11}$$

while satisfying the load balance equation (10) and the operational constraints (5)–(9). C_T represents the operational cost to the residential customer in a period of T hours. To make the best possible use of batteries for the benefit of residential customers, with time of day pricing signals, it is a complex multistage stochastic optimization problem. Adaptive dynamic programming (ADP) which provides approximate optimal solutions to dynamic programming is applicable to this problem. Using ADP, we will develop a self-learning optimization strategy for residential energy system control and management. During real-time operations under uncertain changes in the environment, the performance of the optimal strategy can be further refined and improved through continuous learning and adaptation.

3 Adaptive dynamic programming

In this section, a brief introduction to ADP is presented [25]. Based on Bellman’s principle of optimality [5], dynamic programming is an approach to find an optimal sequence of actions for solving complex optimization problems. Suppose that the following discrete time non-linear system is given

$$x(t + 1) = F[x(t), u(t), t] \tag{12}$$

where $x \in R^n$ denotes the state vector of the system, $u \in R^m$ represents the control action, and F is a transition from the current state $x(t)$ to the next state $x(t + 1)$ under given control action $u(t)$ at time t . Suppose that this system is associated with the performance cost

$$J[x(i), i] = \sum_{k=i}^{\infty} \gamma^{k-i} U[x(k), u(k), k] \tag{13}$$

where U is called the utility function and γ is the discount factor with $0 < \gamma \leq 1$. It is important to realize that J depends on the initial time i and the initial state $x(i)$. The performance cost J is also referred to as the cost-to-go of

state $x(i)$. The objective of dynamic programming problem is to choose a sequence of control actions $u(k)$, $k = i, i + 1, \dots$, so that the performance cost J in (13) is minimized. According to Bellman, the optimal cost from the initial time i on is equal to $J^*[x(i), i] = \min_{u(i)} (U[x(i), u(i), i] + \gamma J^*[x(i + 1), i + 1])$. The optimal control $u^*(i)$ at time i is the $u(i)$ that achieves this minimum, i.e.,

$$u^*(i) = \arg \min_{u(i)} (U[x(i), u(i), i] + \gamma J^*[x(i + 1), i + 1]). \tag{14}$$

ADP is the design based on the algorithm that iterates between a policy improvement routine and a value determination operation to approximate dynamic programming solutions. Generally speaking, there are three design families of ADP: heuristic dynamic programming (HDP), dual heuristic programming (DHP) and globalized dual heuristic dynamic programming (GDHP). The design of ADP we consider in the present paper is called action-dependent heuristic dynamic programming (ADHDP) that does not require the explicit use of a model network in the design. Consider the ADHDP shown in Fig. 4 [24], the critic network in this case will be trained by minimizing the following error measure over time,

$$\begin{aligned} \|E_q\| &= \sum_t E_q(t) \\ &= \sum_t [Q(t - 1) - U(t) - \gamma Q(t)]^2 \end{aligned} \tag{15}$$

where $Q(t)$ represents the critic network output. When $E_q(t) = 0$ for all time t , (15) implies that

$$\begin{aligned} Q(t - 1) &= U(t) + \gamma Q(t) \\ &= U(t) + \gamma[U(t + 1) + \gamma Q(t + 1)] \\ &= \dots \\ &= \sum_{k=t}^{\infty} \gamma^{k-t} U(k). \end{aligned} \tag{16}$$

Clearly, comparing (13) and (16), we have $Q(t - 1) = J[x(t), t]$. Therefore, after the minimization of error function in (15), the output of neural network trained becomes an estimate of the performance cost defined in dynamic programming for $i = t + 1$, i.e., the value of the performance cost in the immediate future.

The input-output relationship of the critic network in Fig. 4 is given by

$$Q(t) = Q[x(t), u(t), t, W_C]$$

where W_C represents the weight vector of the critic network. According to the error function (15), there are two approaches to train the critic network in the present case [24]. We will use the so-called forward-in-time approach.

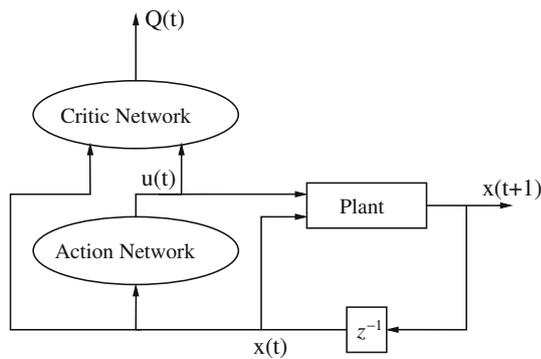


Fig. 4 A typical scheme of an ADHDP

The critic network is trained at time $t - 1$, with the output target given by $U(t) + \gamma Q(t)$. The training of the critic network is to realize the mapping given by

$$C_f: \begin{Bmatrix} x(t-1) \\ u(t-1) \end{Bmatrix} \rightarrow \{U(t) + \gamma Q(t)\}. \quad (17)$$

In this case, we consider $Q(t - 1)$ as the output from the network to be trained and $x(t - 1)$ and $u(t - 1)$ as the input to the network to be trained. We calculate the target output value for the training of the critic network by using its output at time t as indicated in (17). The goal of learning the function given by (17) is to have the critic network output satisfy

$$Q(t - 1) \approx U(t) + \gamma Q(t) \quad \text{for all } t$$

which is required by (16) for approximating dynamic programming solutions.

Using the strategy of [22], the training procedure for the critic network is presented in the following steps:

- Step 1 initialize two critic networks: $cnet1 = cnet2$;
- Step 2 collect data as in (17) including states and action for training;
- Step 3 use $cnet2$ to get $Q(t)$, and then train $cnet1$ for five epochs using the Levenberg–Marquardt algorithm [18];
- Step 4 copy $cnet1$ to $cnet2$, i.e., let $cnet2 = cnet1$;
- Step 5 repeat Steps 3 and 4, e.g., five times;
- Step 6 repeat Steps 2–5, e.g., fifty times;
- Step 7 pick the best $cnet1$ as the trained critic network.

After the training of critic network is completed, we start the action network's training with the objective of minimizing the critic network output $Q(t)$. In this case, the target of the action network training can be chosen as zero, i.e., the action networks weights will be updated so that the critic network output becomes as small as possible. In general, if $U(t)$ is nonnegative, the output of a good critic network should not be negative. The training of the action network in the present ADP is to realize the desired mapping given by

$$A: \{x(t)\} \rightarrow \{0(t)\} \quad (18)$$

where $0(t)$ represents the target values of zero for the critic network output. It is important to realize that the action network will be connected to the critic network during the training as shown in Fig. 4. The target $0(t)$ in (18) is for the output of the whole ADP network, i.e., the output of the critic network after it is connected to the action network as shown in Fig. 4 [25].

After the action network's training is completed, one may check the system's performance, then stop or continue the training procedure by going back to the critic network's training cycle again, if the performance is not acceptable yet.

4 Self-learning scheme for residential energy system

The learning control architecture for residential energy system control and management is based on ADP. However, only a single module will be used instead of two or three modules in the original scheme. The single critic module technique retains all the powerful features of the original ADP, while eliminating the action module completely. There is no need for the iterative training loops between the action and the critic networks and, thus, greatly simplify the training process. There exists a class of problems in realistic applications that have a finite dimensional control action space. Typical examples include inverted pendulum or the cart-pole problem, where the control action only takes a few finite values. When there is only a finite control action space in the application, the decisions that can be made are constrained to a limited number of choices, e.g., a ternary choice in the case of residential energy control and management problem. When there is a power demand from the residential household, the decisions can be made are constrained to three choices, i.e., to discharge batteries, to charge batteries, or to do nothing to batteries. Let us denote the three options by using $u(t) = 1$ for “discharge”, $u(t) = -1$ for “charge”, and $u(t) = 0$ for “idle”. In the present case, we note that the control actions are limited to a ternary choice, or to only three possible options. Therefore, we can further simplify the ADP introduced in Fig. 4 so that only the critic network is needed in the ADP design. Figure 5 illustrates our self-learning control scheme for residential energy system control and management using ADP. The control scheme works in this way: when there is a power demand from the residential household, we will first ask the critic network to see which action (discharge, charge and idle) generates the smallest output value of the critic network; then, the control action from $u(t) = 1, -1, 0$ that generates the smallest critic network output will be chosen. As in the case of Fig. 4,

the critic network in our ADP design will also need the system states as input variables. It is important to realize that Fig. 5 is only a diagrammatic layout that illustrates how the computation takes place while making battery control and management decisions. In Fig. 5, the three blocks for the critic network stand for the same critic network or computer program. From the block diagram in Fig. 5, it is clear that the critic network will be utilized three times in calculations with different values of $u(t)$ to make a decision about whether to discharge or charge batteries or keep it idle. The previous description is based on the assumption that the critic network has been successfully trained. Once the critic network is learned and obtained (offline or online), it will be applied to perform the task of residential energy system control and management as in Fig. 5. The performance of the overall system can be further refined and improved through continuous learning as it learns more experience in real-time operations when needed. In this way, the overall residential energy system will achieve optimal individual performance now and in the future environments under uncertain changes.

In stationary environment, where residential energy system configuration remains unchanged, a set of simple static if-then rules will be able to achieve the optimal scheduling as described previously. However, system configuration including user power demand, capacity of the battery, power rate, etc., may be significantly different from time to time. To cope with uncertain changes of environments, static energy control and management algorithm

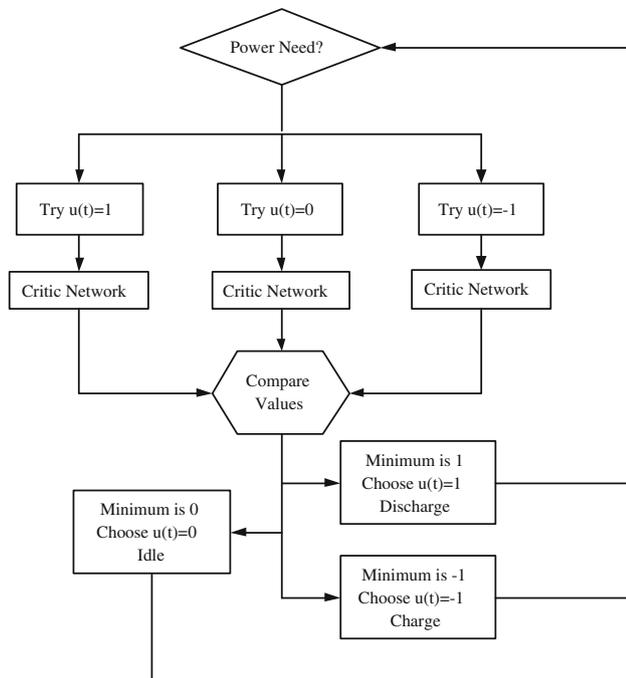


Fig. 5 Block diagram of the single critic approach

would not be proper. The present control and management scheme based on ADP will be capable of coping with uncertain changes of the environment through continuous learning. Another advantage of the present self-learning scheme is that, through further learning as it gains more and more experience in real-time operations, the algorithm has the capability to adapt itself and improve performance. We note that continuous learning and adaptation over the entire operating regime and system conditions to improve the performance of the overall system is one of the key promising attributes of the present method.

The development of the present self-learning scheme for residential energy system control and management involves the following four steps.

Step 1 Collecting data: During this stage, whenever there is a power demand from residential household, we can take any of the following actions: discharge batteries, charge batteries or keep batteries idle and calculate the utility function for the system. The utility function in the present work is chosen as:

$$U(t) = \frac{\text{the electricity charge at time } t}{\text{the possible maximum cost}} \quad (19)$$

During the data collection step, we simply choose actions 1, -1, 0 randomly with the same probability of 1/3. In the meanwhile, the states corresponding to each action are collected. The environmental states we collect for each action are the electricity rate, the residential load, and the energy level of the battery.

Step 2 Training the critic network: We use the data collected to train the critic network as presented in the previous section. The input variables chosen for the critic network are states including the electricity rate, the residential load, the energy level of the battery and the action.

Step 3 Applying the critic network: We apply the trained critic network as illustrated in Fig. 5. Three values of action $u(t)$ will be provided to the critic network at each time step. The action with the smallest output of the critic network is the one the system is going to take.

Step 4 Further updating critic network: We will update the critic network as needed while it is applied in the residential energy system to cope with environmental changes, for example, user demand changes or new requirements for the system. We note that the data has to be collected again and the training of critic network has to be performed as well. In such a case, the previous three steps will be repeated.

Once the training data is collected, we use the forward-in-time method described in the previous section, to train

the critic network. Note that the training for the critic network we describe here can be applied to both the initial training of the critic network and further training of the critic network when needed in the future.

5 Simulation studies

The performance of the proposed algorithm is demonstrated by simulation studies for a typical residential family. The objective is to minimize the electricity cost from power grids over one week horizon by finding the optimal battery operational strategy of the energy system while satisfying load conditions and system constraints. The focus of the present paper is on residential energy system with home batteries connected to the power grids. For the residential energy system, the cost to be minimized is a function of real-time pricing and residential power demands. The optimal battery operation strategy refers to the strategy of when to charge batteries, when to discharge batteries and when to keep batteries idle to achieve minimum electricity cost for the residential user.

The residential energy system consists of power grids, an inverter, batteries and a power management unit as shown in Fig. 1. We assume that the supply from power grid is guaranteed for the residential user demand at any time. The capacity of batteries used in the simulations is 100 kWh and a minimum of 20% of the charge is to be retained. The rated power output of batteries and the maximum charge/discharge rate is 16 kWh. The initial charge of batteries is at 80% of batteries' full-charge. We assume that the batteries and the power grids will not simultaneously provide power to the residential user. At any time, residential power demand is supplied by either batteries or power grids. The power grids would provide the supply to the residential user and, at the same time, charge batteries. It is expected that batteries are charged during the low-rate hours, idle in some mid-rate hours, discharged during high rate hours. In this way, energy and cost savings are both achieved.

The critic network in the present application is a multilayer feedforward neural network with 4–9–1 structure, i.e., four neurons at the input layer, nine neurons at the hidden layer, and one linear neuron at the output layer. The hidden layer uses the hyperbolic tangent function as the activation function. The critic network outputs function Q , which is an approximation to the function $J(t)$ defined as in (13). The four inputs to the critic network are: energy level of batteries, residential power demand, real-time pricing and the action of operation (1 for discharging batteries, -1 for charging batteries, 0 for keeping batteries idle). The local utility function defined in (13) is

$$U(t) = \frac{C(t) \times P_g(t)}{U_{\max}}$$

where $C(t)$ is real-time pricing rate, $P_g(t)$ is the supply from power grids for residential power demand and U_{\max} is the possible maximum cost for all time. The utility function chosen in this way will lead to a control objective of minimizing the overall cost for the residential user.

The typical residential load profile in one week is shown in Fig. 6 [12]. We add up to $\pm 10\%$ random noise in the load curve. From the load curve, we can see that, during weekdays, there are two load peaks occurring in the period of 7:00–8:00 and 18:00–20:00, while during weekend, the residential demand gradually increases until the peak appears at 19:00. Thus, the residential demand pattern during weekdays and during weekend is different. Figure 7 shows the change of the electrical energy level in batteries during a typical one week residential load. From Fig. 7, it can be seen that batteries are fully charged during the midnight when the price of electricity is cheap. After that, batteries discharge during peak load hours or medium load hours, and are charged again during the midnight light load hours. This cycle repeats, which means that the scheme is optimized with evenly charging and discharging. Therefore, the peak of the load curve is shaved by the output of batteries, which results in less consumption of power from the power grids. Figure 8 illustrates the optimal scheduling of home batteries. The bars in Fig. 8 represent the power output of batteries, while the dotted line denotes the electricity rate in real time. From Fig. 8, we can see that batteries are charged during hours from 23:00 to 5:00 next day when the electricity rate is in the lowest range and discharge when the price of electricity is expensive. It is

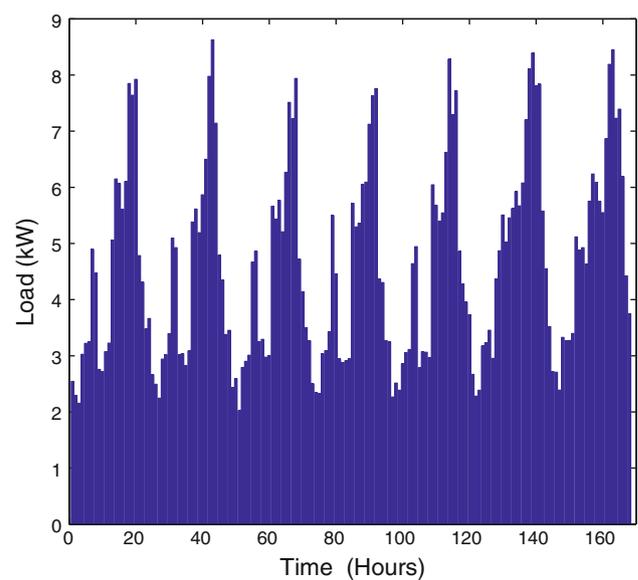


Fig. 6 A typical residential load profile in one week

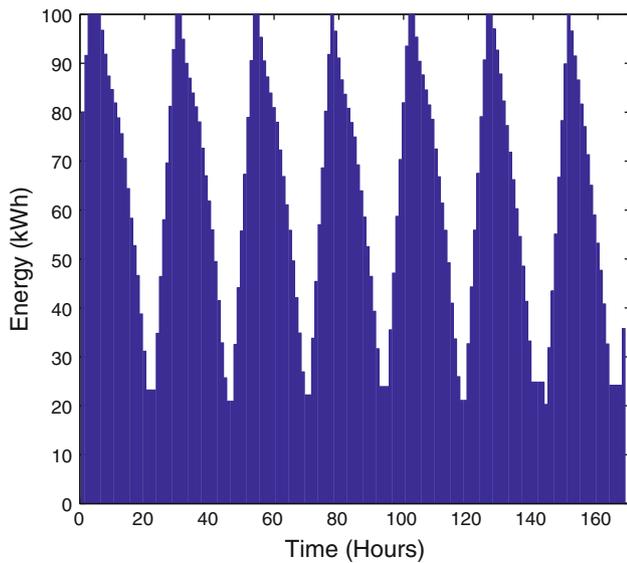


Fig. 7 Energy changes in batteries

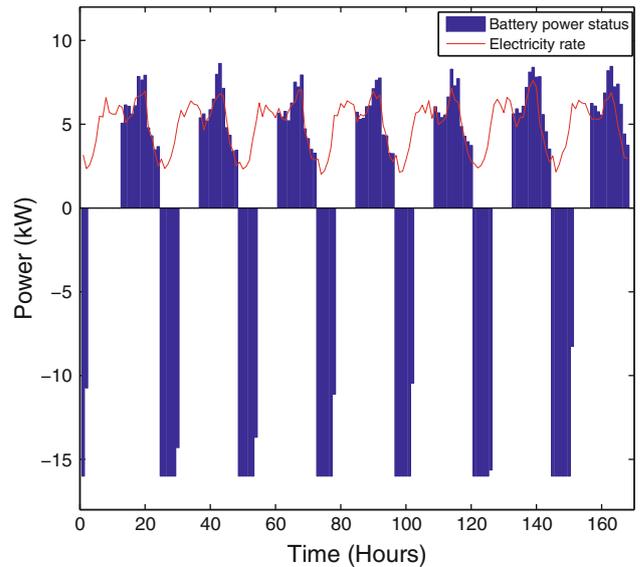


Fig. 9 Scheduling of batteries of fixed daily cycle scheme

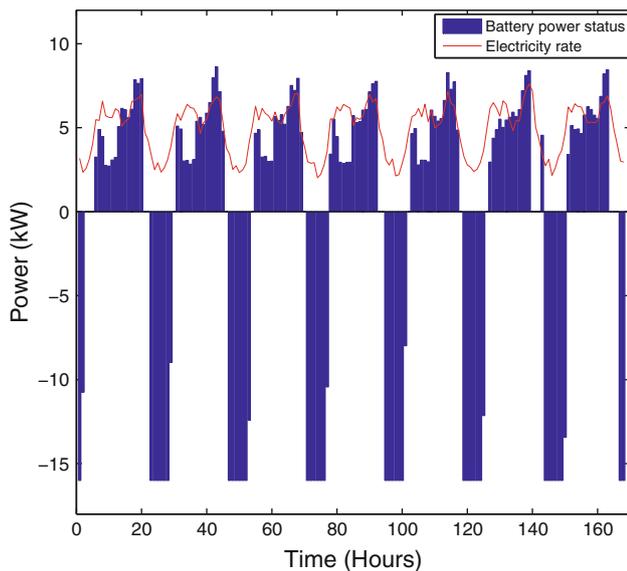


Fig. 8 Optimal scheduling of batteries in one week

observed that batteries discharge from 6:00 to 20:00 during weekdays and from 7:00 to 19:00 during weekend to supply the residential power demand. The difference lies in the fact that the power demand during the weekend is generally bigger than the weekdays’ demand, which demonstrates that the present scheme can adapt to varying load conditions. From Fig. 8, we can also see that there are some hours that the batteries are idle, such as from 3:00 to 5:00 and from 21:00 to 22:00. Obviously, the self-learning algorithm believe that, considering the subsequent load demand and electricity rate, keeping batteries idle during these hours will achieve the most economic return which result in the lowest overall cost to the customer. The cost of

serving this typical residential load in one week is 2866.64 cents. Comparing to the cost using the power grids alone to supply the residential load which is 4124.13 cents, it gives a savings of 1257.49 cents in a week period. This illustrates that a considerable saving on the electricity cost is achieved. In this case, the self-learning scheme has the ability to learn the system characteristics and provide the minimum cost to the residential user.

In order to better evaluate the performance of the self-learning scheme, we conduct comparison studies with a fixed daily cycle scheme. The daily cycle scheme charges batteries during the day time and releases the energy into the residential user load when required during the expensive peak hours at night. Figure 9 shows the scheduling of batteries by the fixed daily cycle scheme. The overall cost is 3284.37 cents. This demonstrates that the present ADP scheme has lower cost. Comparing Fig. 8 with Fig. 9, we can see the self-learning scheme is able to discharge batteries 1 h late from 7:00 to 19:00 during the weekend instead of from 6:00 to 20:00 during weekdays to achieve optimal performance, while the fixed daily cycle scheme ignore the differences of the demand between weekdays and weekend due to the static nature of the algorithm. Therefore, we conclude that the present self-learning algorithm performs better than the fixed algorithm due to the fact that the self-learning scheme can adapt to the varying load consideration and environmental changes.

6 Conclusions

In this paper, we developed a self-learning scheme based on ADP for the new application of residential energy

system control and management. Such a neural network scheme will be obtained after a specially designed learning process that performs approximate dynamic programming. Once the scheme is learned and obtained (offline or online), it will be applied to perform the task of energy cost optimization. The simulation results indicate that the proposed self-learning scheme is effective in achieving minimization of the cost through neural network learning. The key promising feature of the present approach is the ability of the continuous learning and adaptation to improve the performance during real-time operations under uncertain changes in the environment or new system configuration of the residential household. We note that changes in residential demand are inevitable in real-time operations. Therefore, fixed scheme which cannot take demand changes and system characteristics into account is less preferable in practical applications. Another important benefit of the present algorithm is that it can be adapted to different scenarios of different residential customers. Traditional fixed control strategies apply the same control strategy for all system configurations, ignoring the different demands and system configurations. Therefore, this procedure cannot ensure an optimum system design for all customers. With continuous learning and adaptation for residential household energy system, the control scheme based on ADP can obtain the optimal control strategy according to the system configuration and energy utilization of the residential customer. This scheme is customer-centered, unlike the utility-centered, yet effective and simple enough for a real-life use of residential consumers.

References

- Amin SM, Wollenberg BF (2005) Toward a smart grid: power delivery for the 21st century. *IEEE Power Energy Mag* 3(5):34–41
- Bakirtzis AG, Dokopoulos PS (1988) Short term generation scheduling in a small autonomous system with unconventional energy system. *IEEE Trans Power Syst* 3(3):1230–1236
- Balakrishnan SN, Biega V (1996) Adaptive-critic-based neural networks for aircraft optimal control. *J Guid Control Dyn* 19(4):893–898
- Barto AG (1992) Reinforcement learning and adaptive critic methods. In: White DA, Sofge DA (eds) *Handbook of intelligent control: neural, fuzzy and adaptive approaches*. Van Nostrand Reinhold, NY
- Bellman RE (1957) *Dynamic programming*. Princeton University Press, NJ
- Belvedere B, Bianchi M, Borghetti A, Paolone M (2009) A microcontroller-based automatic scheduling system for residential microgrids. In: *Proceedings of the IEEE Power Technology*. pp 1–6
- Bertsekas DP, Tsitsiklis JN (1996) *Neuro-dynamic programming*. Athena Scientific, MA
- Cau TDH, Kaye RJ (2001) Multiple distributed energy storage scheduling using constructive evolutionary programming. In: *Proceedings of the IEEE international conference on power industry computer applications*. pp 402–407
- Chacra FA, Bastard P, Fleury G, Claveul R (2005) Impact of energy storage costs on economical performance in a distribution substation. *IEEE Trans Power Syst* 20(2):684–691
- Chakraborty S, Weiss MD, Simoes MG (2007) Distributed intelligent energy management system for a single-phase high-frequency AC microgrid. *IEEE Trans Ind Electron* 54(1):97–109
- Chiang SJ, Chang KT, Yen CY (1998) Residential photovoltaic energy storage system. *IEEE Trans Ind Electron* 45(3):385–394
- ComEd, USA, <http://www.thewattspot.com>. Accessed 16 May 2010
- Corrigan PM, Heydt GT (2007) Optimized dispatch of a residential solar energy system. In: *Proceedings of the North American Power Symposium*. pp 4183–4188
- Cox C, Stepniewski S, Jorgensen C, Saeks R, Lewis C (1999) On the design of a neural network autolander. *Int J Robust Nonlinear Control* 9:1071–1096
- Dalton J, Balakrishnan SN (1996) A neighboring optimal adaptive critic for missile guidance. *Math Comput Model* 23(1–2): 175–188
- Fung CC, Ho SCY, Nayar CV (1993) Optimisation of a hybrid energy system using simulated annealing technique. In: *Proceedings of IEEE international conference on computer communication control power engineering*. pp 235–238
- Garrity TF (2008) Getting smart. *IEEE Power Energy Mag* 6(2):38–45
- Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. *IEEE Trans Neural Netw* 5(6):989–993
- James M, Jacob B, Scott SG (2006) Dynamic analyses of regenerative fuel cell power for potential use in renewable residential applications. *Int J Hydrogen Energy* 31(8):994–1009
- Klein KM, Springer PL, Black WZ (2010) Real-time ampacity and ground clearance software for integration into smart grid technology. *IEEE Trans Power Deliv* 25(3):1768–1777
- Lee TY (2007) 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* 22(3):774–782
- Lendaris GG, Paintz C (1997) Training strategies for critic and action neural networks in dual heuristic programming method. In: *Proceedings of International Conference on Neural Network*, pp. 712–717
- Li F, Qiao W, Sun H, Wan H, Wang J, Xia Y, Xu Z, Zhang P (2010) Smart transmission grid: vision and framework. *IEEE Trans Smart Grid* 1(2):168–177
- Liu D, Xiong X, Zhang Y (2001) Action-dependent adaptive critic designs. In: *Proceedings of the INNS-IEEE international joint conference on neural network*. pp 990–995
- Liu D, Zhang Y, Zhang H (2005) A self-learning call admission control scheme for CDMA cellular networks. *IEEE Trans Neural Netw* 16(5):1219–1228
- Liu D, Javaherian H, Kovalenko O, Huang T (2008) Adaptive critic learning techniques for engine torque and air-fuel ratio control. *IEEE Trans Syst Man Cybern* 38(4):988–993
- Lo CH, Anderson MD (1999) Economic dispatch and optimal sizing of battery energy storage systems in utility load-leveling operations. *IEEE Trans Energy Convers* 14(3):824–829
- Lu B, Shahidehpour M (2005) Short-term scheduling of battery in a grid-connected PV/battery system. *IEEE Trans Power Syst* 20(2):1053–1061
- Lu C, Si J, Xie X (2008) Direct heuristic dynamic programming for damping oscillations in a large power system. *IEEE Trans Syst Man Cybern B* 38(4):1008–1013

30. Maly DK, Kwan KS (1995) Optimal battery energy storage system (BESS) charge scheduling with dynamic programming. *IEE Proc Sci Meas Technol* 142(6):454–458
31. Meliopoulos APS, Cokkinides G, Huang RK, Farantatos E, Choi S, Lee Y, Yu X (2011) Smart grid technologies for autonomous operation and control. *IEEE Trans Smart Grid* 2(1):1–10
32. Mohagheghi S, Valle Y, Venayagamoorthy GK, Harley RG (2007) A proportional-integrator type adaptive critic design-based neurocontroller for a static compensator in a multimachine power system. *IEEE Trans Ind Electron* 54(1):86–96
33. Mohsenian-Rad A, Wong VWS, Jatskevich J, Schober R, Leon-Garcia A (2010) Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Trans on Smart Grid* 1(3):320–331
34. Momoh JA, Wang Y, Eddy-Posey F (2004) Optimal power dispatch of photovoltaic system with random load In: *Proceedings of IEEE Power Engineering Society*. pp 1939–1945
35. Murray JJ, Cox CJ, Lendaris GG, Saeks R (2002) Adaptive dynamic programming. *IEEE Trans Syst Man Cybern Part C Appl Rev* 32(2):140–153
36. Park JW, Harley RG, Venayagamoorthy GK (2003) Adaptive-critic-based optimal neurocontrol for synchronous generators in a power system using MLP/RBF neural networks. *IEEE Trans Ind Electron* 39(5):1529–1540
37. Prokhorov DV, Wunsch DC (1997) Adaptive critic designs. *IEEE Trans Neural Netw* 8(2):997–1007
38. Qiao W, Harley RG, Venayagamoorthy GK (2009) Coordinated reactive power control of a large wind farm and a STATCOM using heuristic dynamic programming. *IEEE Trans Energy Convers* 24(2):493–503
39. Roncero JR (2008) Integration is key to smart grid management. In: *Proceedings of the IET-CIRED seminar smartgrids for distribution*. pp 1–4
40. Rupanagunta P, Baughman ML, Jones JW (1995) Scheduling of cool storage using non-linear programming techniques. *IEEE Trans Power Syst* 10(3):1279–1285
41. Sheble GB (2008) Smart grid millionaire. *IEEE Power Energy Mag* 6(1):22–28
42. Si J, Wang YT (2001) On-line learning control by association and reinforcement. *IEEE Trans Neural Netw* 12(2):264–276
43. Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. The MIT Press, MA
44. Urbina M, Li Z (2006) A fuzzy optimization approach to PV/battery scheduling with uncertainty in PV generation. In: *Proceedings of the North American Power Symposium*. pp 561–566
45. Wakao S, Nakao K (2006) Reduction of fuel consumption in PV/diesel hybrid power generation system by dynamic programming combined with genetic algorithm. In: *Proceedings of the 1st IEEE World Conference on PV Energy Conversion*. pp 2335–2338
46. Werbos PJ (1977) Advanced forecasting methods for global crisis warning and models of intelligence. *Gen Syst Yearbook* 22:25–38
47. Werbos PJ (1992) Approximate dynamic programming for real-time control and neural modeling. In: White DA, Sofge DA (eds) *Handbook of intelligent control: neural, fuzzy and adaptive approaches*. Van Nostrand Reinhold, NY
48. Yau T, Walker LN, Graham HL, Raithe R (1981) Effects of battery storage devices on power system dispatch. *IEEE Trans Power Appl Syst PAS-100(1):375–383*