Journal of Cleaner Production 204 (2018) 1169–1177

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Deep learning adaptive dynamic programming for real time energy management and control strategy of micro-grid

Nan Wu [a, b], Honglei Wang [a, *]

[a] School of Management, Guizhou University, Guiyang, 550025, China
[b] School of Mechanical and Electrical Engineering, Guizhou Normal University, Guiyang, 550025, China

ARTICLE INFO

Article history:
Received 30 September 2017
Received in revised form
7 August 2018
Accepted 7 September 2018
Available online 7 September 2018

Keywords:
Micro-grid
Deep learning
Adaptive dynamic programming
Real-time energy management
Distributed energy

ABSTRACT

With the widespread application of distributed renewable energy in power systems, energy management problems in micro-grids have become increasingly significant. Offline or static optimization methods are frequently employed to solve the problems which are typically discrete and nonlinear. There are few online optimization methods and they are not only complex but also normally consider the distributed generations and loads in micro-grid as a whole. The outcome is that the online methods fail to reflect the composing characteristic of distributed multi-energy and the contribution made by developing renewable energy to reduce the consumption of traditional fossil energy. Moreover, results obtained by either the offline or the on-line methods can deviate to certain extents from the true values. This paper treats the management of distributed energy in micro-grids as an optimal control problem. Using the system control theory, a framework of real-time management of distributed energy in micro-grids is proposed. A deep learning adaptive dynamic programming is proposed for this framework. Due to the introduction of the concept of closed-loop feedback, the proposed management and control strategy is a real-time algorithm. Furthermore, the accuracy of managing and controlling the objective function can be improved. The gap of the flexible load after the optimization can be helpful in terms of guiding the flexible load consumers to change their habits of energy consumption, thereby reducing the coal-fired power generation and providing room for reducing carbon emissions. Because it is real-time, micro-grid operators can also realize intra-day scheduling, which can be done through combining this optimized management and control strategy with the mechanism of energy operating and managing. Moreover, some data sets of the real-time optimal control strategy are obtained in this paper. These intuitive and accurate data sets can be used to further optimize energy operating and managing of micro-grids, thereby realizing effective energy management for micro-grids. Finally, the real-time and effectiveness of the proposed management and control strategy are proved by the simulations. Some positive conclusions are drawn.

1. Introduction

With the rapid rising global consumption of fossil energy, carbon emissions and environmental pollution problems become more and more serious. Therefore, being clean energy, renewable energy has been increasingly used all over the world, and has become a focus in the future development of global energy. As an important area of fossil energy consumption, the power system has also begun to gradually increase the share of electricity generated by clean energy such as wind power and solar energy, which has reduced the consumption of traditional fossil energy effectively (DM, 2012). The information of consuming these energy sources is implied in a large number of data sets of electricity production and consumption. Among the modes of using renewable energy, the regional distributed multi-energy supply mode integrated with smart grid is one of the mainstream developments. One region can form a micro-grid. The micro-grid consists of distributed generations, batteries, and various loads and is a power supply system that can either operate independently or integrate with a power distribution network. By connecting with the power distribution network, the micro-grid can sell redundant clean power to the network or purchase power from the network if more power is

* Corresponding author.
E-mail addresses: eenwu@gznu.edu.cn (N. Wu), gzdxhlwang@163.com (H. Wang).

https://doi.org/10.1016/j.jclepro.2018.09.052
0959-6526/

needed (Katiraei et al., 2008). Micro-grid can not only reduce effectively long-distance power transmission losses, but also reduce fossil energy power generation, thereby reducing carbon emissions. At present, micro-grid has been researched for energy services, network expansion, and improving efficiency of using energy and there have been some practical applications (Amin and Wollenberg, 2005; Vojdani, 2008; Lpakchi and Albuyeh, 2009). The energy management strategy (EMS) for the micro-grid is to optimize its operational objectives, such as optimizing the long-term operating cost by purchasing and selling electricity to the distribution network while fully utilizing new energy. In addition, because of the large number of distributed power supplies in the micro-grid, it is necessary to carry out statistical analysis and mining of a large number of multi-source power data. These data are collected with the power sensors in order to formulate optimized strategies of energy management for the micro-grid.

Since the micro-grid contains a variety of distributed power sources, energy storage devices, etc., it has more complex structure and more data sources. Therefore, most of its energy management problems are considered to be an offline optimization problem of typical ahead-day scheduling. (Chaouachi et al., 2012; Palma-Behnke et al., 2013; Khodaei, 2014; Shi et al., 2015), which requires early predicting before optimizing. However, distributed energy is intermittent, and most of the loads can change instantaneously, making it difficult to obtain predictions close to real values. There is still a gap between the optimal control strategy and the actual operation of the system. At present, offline optimization algorithms are mostly particle swarm optimization, mixed integer programming, and sequence quadratic programming, etc. (Choi et al., 2011; Cecati et al., 2011).

In addition, the offline optimizing strategy is not real-time, it can't make real-time adjustments to the strategy of energy management and control according to the changing environment. If the environment doesn't change in accordance with the trend of the prior forecast, this method will be invalid, which reduces greatly the efficiency of energy management. To solve the offline optimization problems mentioned above, Siano (Siano et al., 2012) and Pourmousavi (Pourmousavi et al., 2010) proposed real-time energy optimizing strategy based on dividing the timeline into small intervals to improve the efficiency of energy management. Recently, some attempts were made to solve the problems of real-time micro-grid energy management arising from intermittent energy, loads, and demands in micro-grid (Salinas et al., 2013; Huang et al., 2013; Sun et al., 2015). These methods took environmental changes into account and can optimize the long-term cost management for micro-grid with renewable energy. However, all the existing methods assumed that the supply and demand of the system is balanced, ignoring the operating constrains of distributed power supplies in micro-grid. Furthermore, there is no optimal correction target, which makes it difficult for the application of the existing methods to practical control decision-making.

To make the strategy of energy management for micro-grid more applicable, real-time precise, based on the principle of closed-loop control with approximation and correction, this paper solves the problem of management of distributed energy in micro-grid as an optimal control problem. By giving optimal control targets and critic targets and using the approximation structure of functions, a deep learning adaptive dynamic programming based real-time optimization strategy for managing distributed energy in micro-grid is proposed. The proposed strategy can achieve the following goals: (1) optimal control of the long-term operational cost of micro-grid so that renewable energy can be fully used in micro-grid, reduced carbon emissions and the level of environment pollution caused by the consumption of fossil energy, and moreover the optimized strategy of energy management; (2) improved power supply services for users in micro-grid and optimized energy management for micro-grid; (3) real-time management and control decision making for the energy management of micro-grid.

This paper is organized as follows: the system model and objective function are introduced in section 2; a deep learning adaptive dynamic programming based strategy of on-line energy management for micro-grid is proposed in section 3. The effectiveness of the proposed strategy is verified in section 4; the outcome of the paper is summarized in section 5.

2. System model

The model of the micro-grid system is described in this section. Assume that the entire micro-grid is run by an operator whose goal is to minimize the cost of power in the micro-grid whilst the power quality of the supply service is ensured. The model of distributed energy resources (DERs) is given firstly, including adjustable and non-adjustable parts, and then the cost model of storage device and the shedding model of flexible loads are given, and finally the purchase and sale model of the micro-grid and distribution network and the cost model of the overall system optimization are given.

2.1. Overview of the micro-grid system

Various kinds of distributed generation such as wind, solar, gas, water, and so on, exist in the micro-grid. The distributed generation units are represented as $G = \{g_1, g_2, g_3, \ldots, g_G\}$. In order to use intermittent energy effectively and without causing impacts to the grid, the energy is usually stored before transmitting to customers. Storage devices in the grid are expressed as $S = \{s_1, s_2, s_3, \ldots, s_S\}$. The loads in the grid can be divided into rigid and flexible loads. Operators of the micro-grid can divide or combine the flexible loads in order to achieve the lowest cost of the power supply, whilst ensuring that the quality of the power supply service of the entire system will not be reduced significantly. The load in grid can be expressed as $L = \{l_1, l_2, l_3, \ldots, l_L\}$. All the distributed power supplies, energy storage devices, and controllable loads in the micro-grid are connected to the control center through a two-way communication network. The micro-grid energy control center (MGECC) can acquire real-time information of the distributed power supplies and flexible loads and then based on the optimization strategy of energy management, make adjustments to these devices through local controllers (LCs). The diagram of the system is shown in Fig. 1.

2.2. Distributed generation model

In the micro-grid, the distributed power supplies can be divided into two categories. One is the intermittent power sources, such as wind and solar power. The generated power is constrained by the natural environment (i.e., the strength of the wind and the intensity of the light), rather than manually controlled. The other is the adjustable powers, such as gas and water power.

2.2.1. Intermittent distributed renewable generations

Intermittent distributed renewable generation units consist of solar and wind power, which can be expressed as $g_r \in G$. These renewable generations can be considered as only incurring the cost of maintenance and no cost of generation. Therefore, the cost of generating these energies in this paper is considered to be zero. To overcome the instability caused by the intermittent power supplies, generated energy is often stored in energy storing devices before being transmitted to loads.

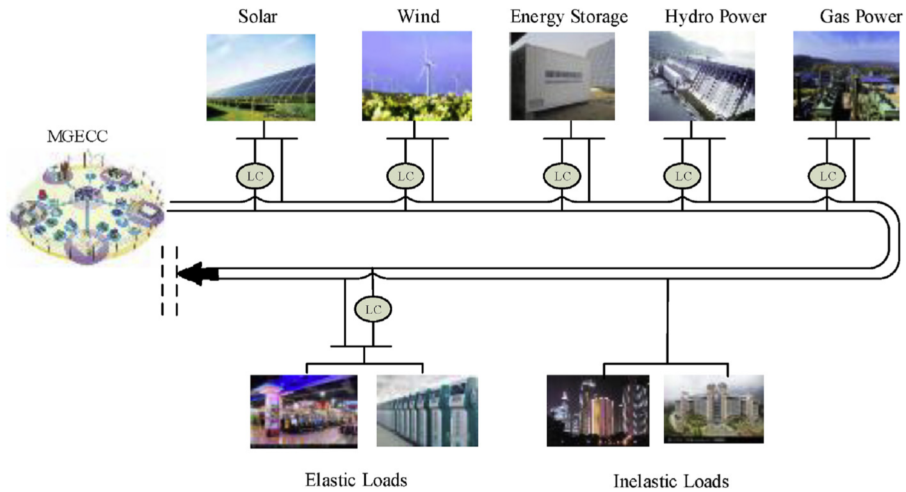


Fig. 1. Diagram of the micro-grid system.

2.2.2. Traditional power generation

Traditional power generation can be adjusted according to changing power consumption in the micro-grid. The adjustable power supplies in the micro-grid are mainly gas and water power. The generation unit of water power can be expressed as $w \in G_w$ whose constraints of operation are:

$$0 \le p_w(t) \le p_{w\max} \quad (1)$$

$$|p_w(t) - p_w(t-1)| \le p_{w\max} \quad (2)$$

where $p_{w\max}$ is the maximum generation of water power. The model of the cost of generation at time t is (Yun et al., 2007):

$$C_w(p_w(t)) = a_w p_w(\mathrm{t})^2 + \beta_w p_w(t) + \kappa_w \quad (3)$$

where a_w, β_w, κ_w are constants. The other adjustable power supply is gas whose generation unit is $g \in G_g$, and at time t the constrains and the model of the cost of generation are:

$$0 \le p_g(t) \le p_{g\max} \quad (4)$$

$$\left|p_g(t) - p_g(t-1)\right| \le p_{g\max} \quad (5)$$

$$C_g\left(p_g(t)\right) = a_g p_g(\mathrm{t}) \quad (6)$$

where $p_{g\ \max}$ is the maximum output of the gas station and a_g is a constant.

2.3. Storage device model

In this paper, the distributed energy storage unit of the micro-grid is a storage battery and all storage units form a battery energy storage system (BESS). The intermittent distributed power generation in the micro-grid is stored in the BESS firstly and then the stable output power is provided by the BESS. For a storage battery in a micro-grid $s \in S$, when it is charged at time t, the active power $p_s(t)$ is positive, i.e., the intermittent distributed power supply charges the storage battery. If the storage battery is discharged, the active power $p_s(t)$ is negative, i.e., both the intermittent distributed power supply and battery are supplying power to loads in the micro-grid. $E_s(t)$ is the storage capacity of the storage battery at time t. Operating constraints of $E_s(t)$ are as follows:

$$p_{d\max} \le p_s(t) \le p_{c\max} \quad (7)$$

$$E_s(t+1) = E_s(t) + p_s(t) \quad (8)$$

$$E_{s\min} \le E_s(t) \le E_{s\max} \quad (9)$$

where -$p_{d\max}$ is the maximum allowable discharge rate, $p_{c\max}$ is the maximum allowable charge rate, $E_{s\min}$ is the minimum capacity that storage battery needs to retain, and $E_{s\max}$ is the maximum capacity that the storage battery can store.

To punish the impact of fast charging and fast discharging of energy storage batteries on the service life of storage batteries, a cost function is introduced in this paper. For a given storage battery s, its running cost model at time t is as follows (Sun et al., 2015):

$$C_s(p_s(t)) = a_s p_s(t)^2 + \kappa_s \quad (10)$$

where a_s is a fast charging and discharging penalty factor and κ_s is a basic cost factor. Both are positive constants. The cost function $C_s(p_s(t))$ is convex. Moreover, the capacity of the battery energy storage system is set to fully meet the storage and scheduling requirements of intermittent distributed power supplies.

2.4. Flexible load model

Loads in micro-grid mainly consist of rigid and flexible loads. Under normal conditions, the power supply must be ensured for rigid loads. For flexible loads, such as central air conditioning, charging stations for electric vehicles, etc., power can be supplied according to the condition of the power supply or appropriate strategies of supplying power can be made. For each load unit $l \in L$ in the micro-grid, the supplied power at time t $p_l(t)$ must satisfy:

$$p_{l\min}(t) \le p_l(t) \le p_{l\max}(t) \quad (11)$$

where $p_{l\min}(t)$ is the rigid load that must be satisfied, and $p_{l\max}(t)$ is the maximum demand of power in the grid. Micro-grid operators should try to meet the needs of customers, and ideally the left side of (11) is equal to the right side. However, the power demand of customers is random and unpredictable. Therefore, a shortage or a surplus of the distributed power supplies can be expected in the micro-grid. Especially when the shortage of the power supplies is big, a lot of flexible loads in the grid will be cut off to ensure the lowest cost of power supply services, which can cause a sharp decline in the customer satisfaction. To avoid this situation, power supplies in the micro-grid must be managed optimally.

In this paper, in order to allow the micro-grid operators provide a high quality service without cutting flexible loads to achieve the minimum cost of power supplies, a cost function representing cutting flexible loads is used to punish the act of sacrificing customer satisfaction for the lowest cost of power supplies.

$$C_l(p_l(t)) = \gamma_l(p_{l\max}(t) - p_l(t))^{1/2} \tag{12}$$

where γ_l is the positive penalty factor.

2.5. Optimization of the objective function

In the micro-grid, the aim of energy management includes firstly, retaining the lowest operational cost of power generation, secondly, providing stable and high quality power supplies, and thirdly, real-time control decision making according to the changes of energy. However, the generation of distributed energy is intermittent, which is a challenge to the real-time balance of the supply and demand in the micro-grid. This means that the micro-grid operators cannot cut off too many flexible loads when the power supplies are not sufficient and furthermore, cannot purchase a large quantity of thermal power to satisfy all customers since that will increase the operating cost. Therefore, a strategy of real-time energy management is needed.

Distributed energy resources are subject to the significant impact of the environment and season, resulting in the frequent insufficiency of power supplies in the micro-grid. In this situation, micro-grid operators need to purchase power from elsewhere for supplying customers, the cost is:

$$C_m(p_m(t)) = \rho \cdot p_m(t) \tag{13}$$

where ρ is the price of power, $p_m(t)$ is the amount purchased. In order to ensure the minimum cost of power supplies, the operators will usually not buy all the needed power, but buy a portion of the needed power and cut off some flexible loads at the same time. To guarantee the quality of the supply services, a function is used to limit the operation of cutting off flexible loads:

$$\mu(t) = \frac{1}{T}\Sigma \frac{p_{l\max}(t) - E(p_l(t))}{p_{l\max}(t) - p_{l\min}(t)} \le \sigma \tag{14}$$

where $p_{l\max}(t) - p_{l\min}(t)$ is the total amount of flexible loads that can be cut off, $p_{l\max}(t) - E(p_l(t))$ is the amount of flexible loads being cut off, and σ is the removal ratio of flexible loads.

One of the goals of operating the micro-grid is to retain the lowest long-term cost of power generation and supply. The total cost includes generating power, energy storage and conversion, purchasing power, and cutting-off flexible load. Therefore, the objective function of optimization can be derived as:

$$\begin{cases} \min\limits_{\mu(t)} \left(\phi_w \underset{w\in G_w}{\Sigma} C_w(p_w(t)) + \phi_g \underset{g\in G_g}{\Sigma} C_g\left(p_g(t)\right) + \phi_s \underset{s\in S}{\Sigma} C_s(p_s(t)) + \phi_l \underset{l\in L}{\Sigma} C_l(p_l(t)) \right) + \phi_m C_{\mathrm{m}}(P_m(t)) \\ \text{s.t. } (1)\ (2)\ (4)\ (5)\ (7)-(9)\ (11) \end{cases} \tag{15}$$

Solving the objective function is constrained by (14), i.e., the micro-grid operators cannot cut off too many flexible loads in pursuit of the lowest cost and the removal ratio of flexible loads cannot exceed σ. This is a typical closed-loop control for a control system. If the objective function is regarded as the control object and the removal ratio of flexible loads is considered as the error of the feedback control, the problem becomes a typical optimization problem. Meanwhile, the real-time performance of a control problem is one of the basic properties of the problem, which can guarantee the dynamic and real-time properties of the optimization target. To solve typical optimal control problems, this paper proposes a deep learning adaptive dynamic programming based approach. The proposed approach can not only meet the requirements mentioned above but also improve the accuracy of optimization. Therefore, the deep learning adaptive dynamic programming is used to solve the optimization problem in this paper to determine the optimal strategy of energy management for the micro-grid.

3. Deep learning adaptive dynamic programming based real-time energy management strategy

3.1. Principle of dynamic programming and adaptive dynamic programming

Dynamic systems are ubiquitous in nature. An important branch of dynamic systems theory is optimal control. Optimal control has been widely used in many fields, such as system engineering, economic management, decision making, etc. In 1957, Bellman proposed an effective tool for solving optimal control problems: dynamic programming (Bellman, 1966). The core of this method is Bellman's principle of optimality which says that if there is a nonlinear system whose dynamic equation is:

$$x(k+1) = F(x(k), u(k), k) \tag{16}$$

where $x(k)$ is state of the system whose initial state $x(k) = x_k$ is a given value, $u(k)$ is the control input of the system, and $F(.)$ is the utility function of the system. The performance indicator function of the system can be defined as:

$$J(x(i), i) = \sum_{k=i}^{\infty} \lambda^{k-i} F(x(k), u(k), k) \tag{17}$$

The target of control is to solve the sequence of admission control (or decision) $u(k)$, $k = 1, 2, 3, \ldots$, making the cost function (17) reach its minimum. According to Bellman's principle, starting from time k the minimum cost of any state includes two parts. One is the minimum cost required for time k, and the other is the sum of the minimum cost from time k+1 to infinite, i.e.:

$$J^*(x(k)) = \min_{u(k)} \{F(x(k), u(k), k) + \lambda J^*(x(k+1))\} \tag{18}$$

At this point, the corresponding control strategy $u(t)$ at time k is also optimal, i.e.:

$$u^*(k) = \arg\min_{u(k)}\{F(x(k), u(k), k) + \lambda J^*(x(k+1))\} \tag{19}$$

Therefore, the dynamic programming method is a powerful tool for solving optimal control problems (Zhang et al., 2012). However, it is difficult to use dynamic programming directly in practical applications. The reason is that optimal control needs to affect a system in a time sequencing manner and give out optimal control indicators of a control sequence. Nevertheless, the performance indicator function of the entire system is completely unknown before the sequence is completed, i.e., the problem of "Curse of dimensionality". Webros in 1977 first proposed adaptive dynamic programming (ADP) to solve the problem. The ADP method is essentially using the structure of function approximation to fit the cost function and control strategy of dynamic programming, deriving solutions for optimal control of nonlinear systems (Powell, 2009). A typical ADP structure is shown in Fig. 2.

In Fig. 2, the performance indicator function can be expressed as:

$$J(x(k)) = l(x(k), u(x(k))) + J(x(k+1)) \tag{20}$$

where $u(x(k))$ is the variable of feedback control and the performance indicator function $J(x(k))$ and $J(x(k+1))$ are the output of the critic network. If the weight of the critic network is w, the right side of (20) can be expressed as:

$$d(x(k), w) = l(x(k), u(x(k))) + J(x(k+1), w) \tag{21}$$

The left side of (20) can be written as $J(x(k), w)$. By adjusting the weight w of the critic network the mean square error function (22) can be minimized to obtain the optimal performance indicator function.

$$w^* = \underset{w}{\mathrm{argmin}}\left\{|J(x(k), w) - d(x(k), w)|^2\right\} \tag{22}$$

According to the principle of optimality, the optimal control has to satisfy the first order differential necessary condition which is:

$$\begin{aligned}\frac{\partial J^*(x(k))}{\partial u(k)} &= \frac{\partial l(x(k), u(k))}{\partial u(k)} + \frac{\partial J^*(x(k+1))}{\partial u(k)} \\ &= \frac{\partial l(x(k), u(k))}{\partial u(k)} + \frac{\partial J^*(x(k+1))}{\partial x(k+1)}\frac{\partial f(x(k), u(k))}{\partial u(k)}\end{aligned} \tag{23}$$

The obtained optimal control is:

$$\begin{aligned}u* = \underset{u}{\mathrm{argmin}}\Bigg(\Bigg|&\frac{\partial J(x(k))}{\partial u(k)} - \frac{\partial l(x(k), u(k))}{\partial u(k)} \\ &- \frac{\partial J*(x(k+1))}{\partial x(k+1)}\frac{\partial f(x(k), u(k))}{\partial u(k)}\Bigg|\Bigg)\end{aligned} \tag{24}$$

In recent years, the theory of adaptive dynamic programming has been developed. Online adaptive methods have been proposed to solve optimal control problems of nonlinear analog systems and optimal tracking and stabilizing problems of nonlinear discrete systems (Vamvoudakis et al., 2009; Vamvoudakis and Lewis, 2010, 2012; Dierks and Jagannathan, 2010). Therefore, adaptive dynamic programming has become an extremely important method for solving both scientific and engineering problems of modern complex systems (Lewis and Liu, 2009). The method has already been used in many practical engineering systems, such as population control, temperature control, tracking control of UAV helicopters, etc. (Nodland et al., 2013; Zhang et al., 2012; Yadav et al., 2007; Tang et al., 2007).

3.2. Deep learning adaptive dynamic programming based real-time energy management strategy

With the inspiration of the above-mentioned basic form of adaptive dynamic programming, by combining the simplified ADP structure with the characteristics of self-learning feature extraction of deep learning, this paper proposes a deep learning adaptive dynamic optimization algorithm to solve the problem of energy management for micro-grids. The proposed algorithm does not require the model to predict the system state at the next moment and only includes the critic network and the action network, which can reduce the dependence of the controller on the system model. The structure of the proposed algorithm has the closed-loop feedback typical of control theories. The basic structure is shown in Fig. 3. Deep learning origins from neural network. Deep learning can obtain representations of characteristics on deep levels and avoid the complexity of selecting the characteristics manually (Lecun et al., 2015). Therefore, with the real-time energy management strategy, the proposed deep learning adaptive dynamic programming algorithm in this paper also makes the real-time control algorithm have online self-learning ability. This means that the traditional optimization algorithm has been enhanced with real-time performance. The proposed algorithm can overcome effectively the non-real time problem in the static decision making of traditional optimization algorithms. Furthermore, the proposed algorithm has the advantages of simpler structure, more precise closed-loop control and better online characteristics, which can meet the needs of real-time energy management for micro-grids.

In this system, the input of the action network is the state of the controlled object, i.e., the amount of distributed energy in the micro-grid, $p_w(t)$, $p_g(t)$ and $p_s(t)$. The output is the control strategy, i.e., the amount of power $p_{\mathrm{m}}(t)$ that has to be purchased in order to satisfy the objective function of the minimum cost. The power $p_{\mathrm{m}}(t)$ is equal to zero if the constrain given in (14) does not exist. In that case, the operational cost of the system will be the minimum if no power is purchased from the distribution network. However, subject to the constraint in (14), $p_{\mathrm{m}}(t)$ has to be optimized so that the minimum cost of the system can be achieved. Since the output of the action network $u(t) = p_m(t)$ needs to satisfy the constraints, the system has to be trained in order to get the output $p_{\mathrm{m}}(t)$.

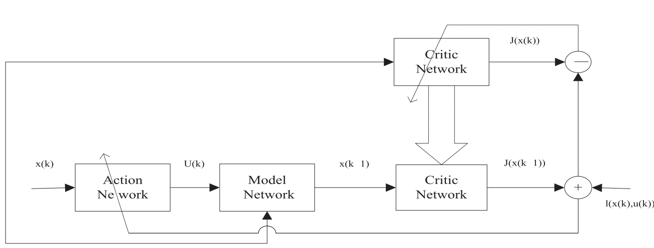


Fig. 2. Structure of adaptive dynamic programming.

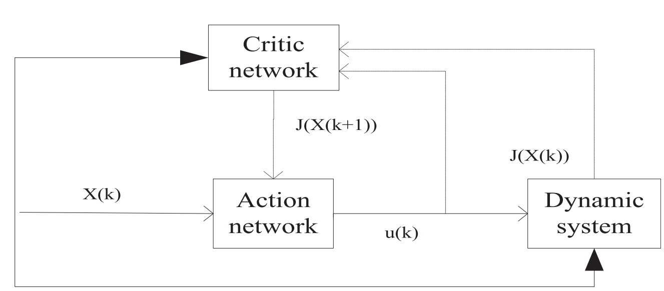


Fig. 3. Basic structure of deep learning adaptive dynamic programming.

The input of the critic network is the state of the controlled object and the control strategy. The output is the cost function. The utility function can be defined according to the target of control. The objective function in this paper is defined to realize the lowest cost of generating and supplying power in the micro-grid. In order to improve the accuracy of using the system to approximate the objective function, it is necessary to control the error in the output of the action network, i.e., reducing the error to the lowest acceptable value (the minimum error principle of control system). Therefore, for both action network and critic network in Fig. 3, multilayer neural network of deep learning is used. The topology is shown in Fig. 4.

In the micro-grid environmental changes will cause fluctuating output of distributed power supplies. This calls for real-time decision making to form a basis of the energy management for micro-grid. In order to retain the minimum long-term cost of generating and supplying power in the micro-grid, appropriate measures must be adopted for flexible loads. The measures include purchasing a portion of the demanded power and removing some of the flexible loads (or energy saving measures). These measures have been considered in the objective function of system optimization detailed earlier in this paper. Based on the structure of the model described above, to allow the real-time optimization of the objective function, a flow chart of implementing the deep learning adaptive dynamic programming based optimizing strategy is shown in Fig. 5.

The steps are as follows:

1) Set parameters for the structure of the system, action network, and critic network, respectively. Both action network and critic network use multilayer neural network of deep learning. The settings of the neural network include the learning rate of the network, the number of nodes on the input layer, the hidden layer and the output layer, and the maximum number of iteration;
2) Collect data and set input parameters of the controlled state. The collected data include real time output power of distributed power supplies;
3) Initialize training of the action network and critic network, the input variables and the learning rate;
4) Use equation (14) to determine whether the constraint relating to customer satisfaction is met. If not, the next network training will be carried out. If yes, no changes will be made to the current values;
5) Use the state of the controlled objects as the input of action network, carry out training of action network, update the weight of action network, and obtain the output, i.e., the control strategy;
6) Use the state of the controlled object and the control strategy as the input of critic network, carry out training of critic network, update the weight of critic network, and obtain the output, i.e., the cost function;
7) Keep the control strategy and calculate the corrected objective function;
8) Repeat steps (4)–(7) during the observation period until the end of the optimization process, and obtain the output, i.e., the control strategy.

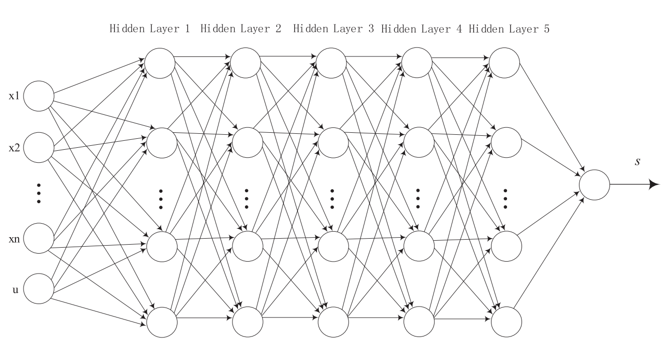


Fig. 4. Multilayer neural network topology of deep learning.

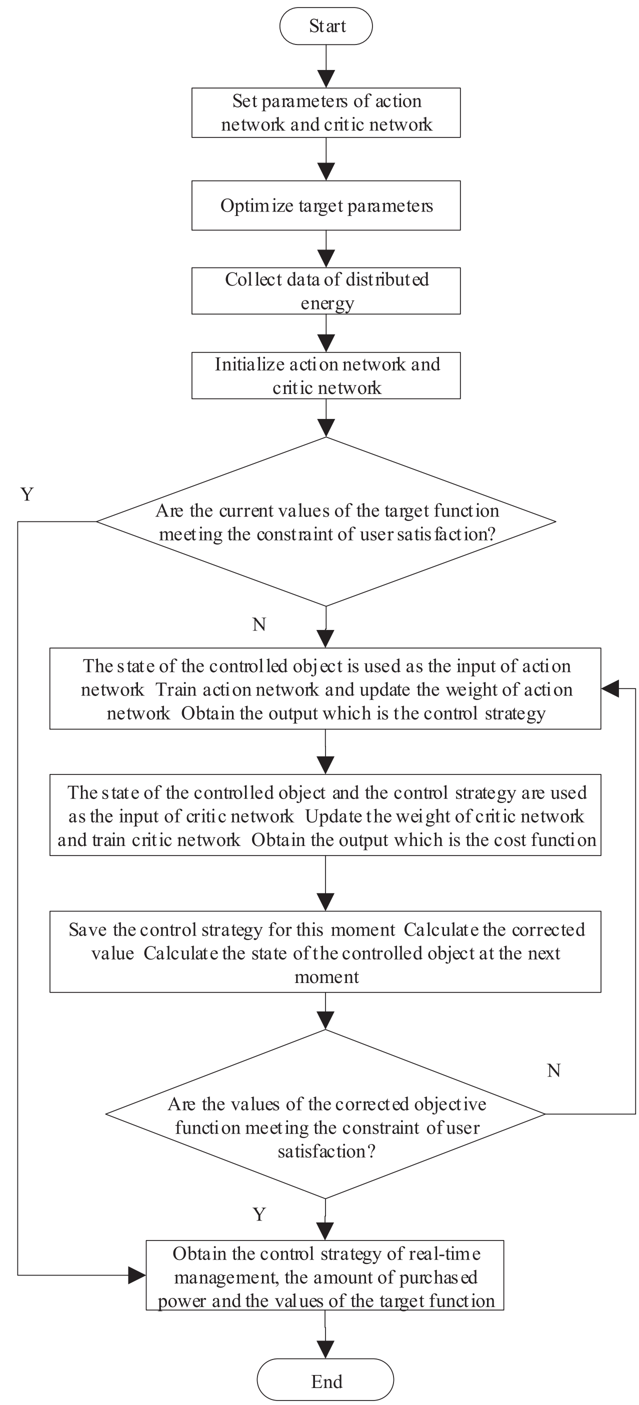


Fig. 5. Flowchart of micro-grid energy management strategy based on deep learning adaptive dynamic programming.

After training action network, the amount of the purchased

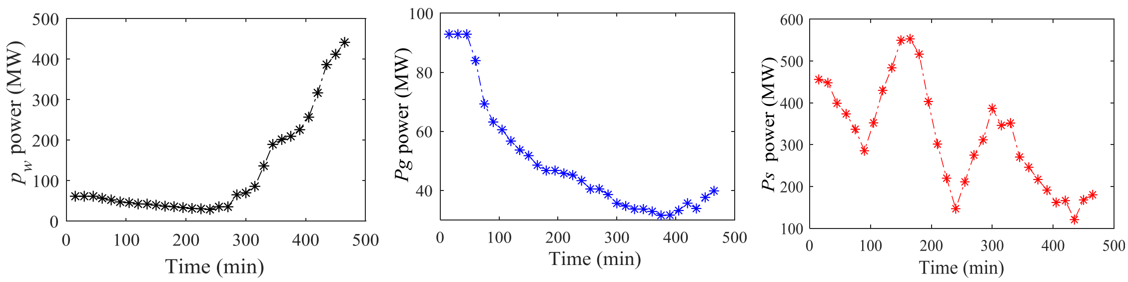


Fig. 6. Amount of power from distributed power supplies.

power $p_m(t)$ is obtained. If both conditions of the objective function and customer satisfaction are met, the training will be stopped. Otherwise, the training of action network and critic network will be repeated until an optimizing strategy meeting the conditions is found.

4. Simulation and analysis

In order to assess the effectiveness of the proposed optimal control strategy, distributed power supplies in the micro-grid are sampled referring to 15 min specified in the standard of Guizhou Power Grid Corp. The total length of acquired data is 465 min, as shown in Fig. 6. The horizontal axis is the time of data acquisition and the vertical axis is the distributed power generation. The parameters of the cost model for distributed energy are set as follows: the cost coefficient of water power generation $a_w = 5$, $\beta_w = 12$, $\kappa_w = 0$; the cost coefficient of gas power generation $a_g = 0.2$; the penalty factor of power satisfaction $\gamma_1 = 0.3$; the price of purchasing $\rho = 0.4$ CNY/KWh. Parameters of the objective optimization function are set as $\phi_w = \phi_g = \phi_s = \phi_l = \phi_m = 1$. The coefficients of the storage battery cost function are set as $a_s = 1$, $\kappa_s = 0$. The battery energy storage system (BESS) $E_{smax}(t)$ has a maximum storage capacity of 1200 MW and a minimum storage capacity $E_{smin}(t)$ of 240 MW. The initial capacity $E_s(0)$ of the battery energy storage system is set to 600 MW. The minimum rigid load is set to 50% of the maximum demand of loads. In order to ensure the quality of power supply services, the cut-off ratio cannot be higher than 15% of the flexible load. In addition, the distributed generations of wind power and photovoltaic power in the micro-grid all are stored in the BESS firstly and then through the storage device power of standard frequency and voltage is supplied to users. All the examples in this paper are calculated using a PC with a CPU of 3.4 GHz, RAM of 8 GB, an operating system of Windows 10, and a MATLAB release of R2017b V9.0.3.

The collected real-time power data and the above parameters are substituted into the deep learning adaptive dynamic programming algorithm proposed in this paper. A total of 4 hidden layers are applied in deep learning. A neural network with 30 neurons on each layer is employed to train and optimize the critic network and action network. The time required for each epoch is 0.099s. After running the program, the relation between the defined target and the output and the relation between the number of training epochs and the mean squared error are obtained as Fig. 7. Each epoch is a process in which a complete data set passes through the neural network once and returns once, that is, an

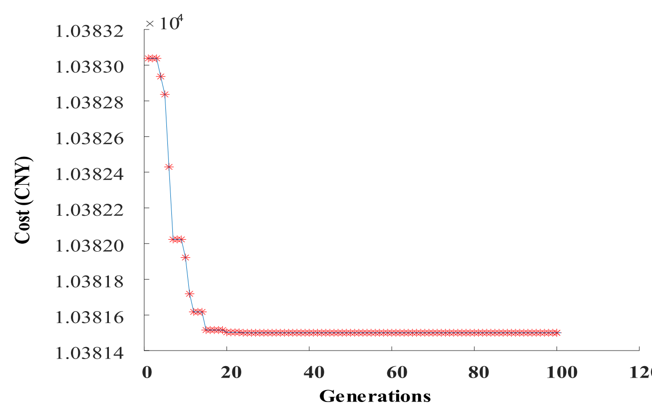
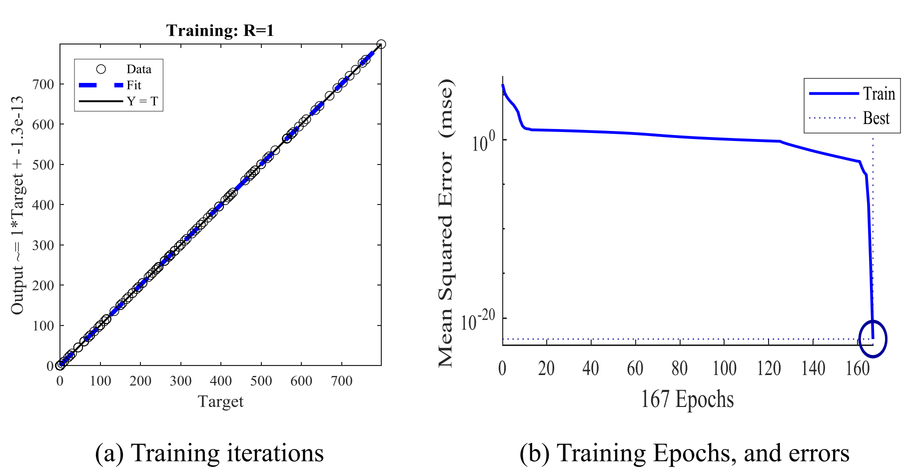
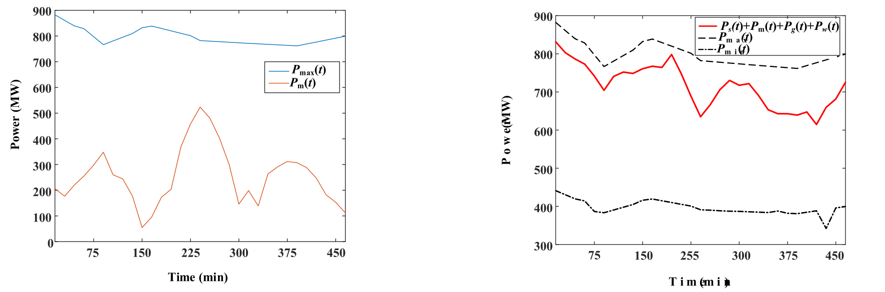


Fig. 8. Approximating the objective function of the micro-grid through optimal cost control.



(a) Training iterations (b) Training Epochs, and errors

Fig. 7. Training process of deep learning.



(a) Maximum demanded power and purchased power

(b) Maximum demanded power, supplied power and minimum demanded power

Fig. 9. Strategy of real time power supply in the micro-grid.

epoch.

Fig. 7 shows, that deep learning achieves a better training effect. It can also be seen that the target values of the training process are generally consistent with the output values. After 167 training epochs of the deep learning neural network, the best approximation is achieved, taking 16.533 s and with a mean squared error of 4.3967×10^{-23}. The results suggest that the proposed method can converge quickly to actual values of the objective function, and moreover real-time and accurate control strategies can be obtained. Fig. 8 shows, with the control strategy of energy management proposed in this paper, the fast optimizing process of the objective single cost function for the power delivery in the micro-grid. “Generations” represents the number of iterations of the objective cost function iterations in it.

Meanwhile, Fig. 9 shows the maximum power demand and purchased power under the condition of the minimum cost of power supplies and the real-time control strategy of power supplies for the maximum power demand and the amount of power supplied by the micro-grid. If distributed power supplies are insufficient, the micro-grid operators can follow the principle of the lowest cost and buy a portion of the demanded power from the distribution network to meet the power requirement of most flexible loads.

Fig. 9 also shows that the micro-grid operators cut off some flexible loads from the network in order to achieve the lowest cost.

On the other hand, some gaps in the flexible load after optimization can be used to guide flexible consumers to change their habits of energy consumption and plan their own energy use more rationally or seek other clean energy sources as alternatives of coal power. This is beneficial to the implementation of environment protection policies aiming to reduce the generation of coal electricity and carbon emissions advocated by various countries. Moreover, according to the precise large data sets obtained from this strategy, for users who do not want to change their habits of power consumption, micro-grid companies can also make some ladder prices or incentive policies to guide flexible consumers to consume more rationally. This is not only beneficial to the long-term development of micro-grid companies, but also helpful for promoting consumers to change the traditional concept of consuming energy so that the utilization of clean energy can be more effective.

The simulation results also indicate that the cost make-up of power supplies influences greatly the optimization strategy of power supplies in the micro-grid. This can provide room for engineers to further improve and upgrade the relevant technologies and equipment, e.g., improving the efficiency of power generation using renewable energy, developing storage batteries with larger capacities and lowering costs of charging and discharging. Achieving these goals will not only improve the efficiency of energy management for micro-grid greatly but also be helpful for achieving the global energy saving and emission reducing targets.

5. Conclusion

Based on the optimal control theory of closed-loop feedback, a deep learning adaptive dynamic programming based strategy of real-time energy management for micro-grids is proposed in this paper. Using the deep learning adaptive dynamic programming, the cost function and control strategy of dynamic programming can be approximated. Bellman's principle of optimality is used to solve the optimal control problem in a real-time manner. The obtained performance indicator function is subsequently used for solving the defined optimization problem. Compared with offline management strategies and some segmented real-time strategies, the proposed strategy has the characteristics of real-time, reliability and accuracy typical of control systems. Furthermore, the proposed strategy allows the decision-making data of energy management for micro-grids to be real-time and self-learning. Therefore, intra-day scheduling can be implemented according to the energy management and operating policies of micro-grid operators. At the same time, due to the real-time correction of control errors, the control strategy can get closer to the actual state and follow the changing trend. The obtained real-time data of the control strategy can be regarded as a basis for guiding the change of the energy consumption concept of flexible loads. Moreover, these intuitive and accurate large data sets are useful for further optimizing of energy operating and managing for the micro-grid.

The simulation analysis proves the effectiveness of the proposed management strategy. Given the optimal real-time data, the strategy can find the best result of the objective function. The obtained result is helpful for online real-time decision making, optimizing the operational cost and improving the efficiency of utilizing renewable energy. The benefit will be that carbon emissions and environmental pollution caused by the extensive use of fossil fuels can be reduced effectively.

Acknowledgements

This work was supported by the key science and technology projects of China Southern Power Grid Co., Ltd [grant number GZKJQQ00000417].

References

Amin, S.M., Wollenberg, B.F., 2005. Toward a smart grid: power delivery for the 21st century. IEEE Power Energy Mag. 3, 34–41.
Bellman, R., 1966. Dynamic programming. Science 153, 34–37.
Cecati, C., Citro, C., Siano, P., 2011. Combined operations of renewable energy systems and responsive demand in a smart grid. IEEE Trans. Sustain. Energy 2, 468–476.
Chaouachi, A., Kamel, R.M., Andoulsi, R., Nagasaka, K., 2012. Multiobjective intelligent energy management for a microgrid. IEEE Trans. Ind. Electron. 60, 1688–1699.
Choi, S., Park, S., Kang, D.J., Han, S.J., Kim, H.M., 2011. A microgrid energy management system for inducing optimal demand response. In: IEEE International Conference on Smart Grid Communications, pp. 19–24.
Dierks, T., Jagannathan, S., 2010. Optimal tracking control of affine nonlinear discrete-time systems with unknown internal dynamics. In: Decision and Control, 2009 Held Jointly with the 2009 Chinese Control Conference. Cdc/ccc 2009. Proceedings of the IEEE Conference on, pp. 6750–6755.
DM, N., 2012. Reforming competitive electricity markets to meet environmental targets. Econ. Energy Environ. Pol. 1, 69–82.
Huang, Y., Mao, S., Nelms, R.M., 2013. Adaptive electricity scheduling in microgrids. IEEE Trans. Smart Grid 5, 270–281.
Katiraei, F., Iravani, R., Hatziargyriou, N., Dimeas, A., 2008. Microgrids management. Power Energy Mag. IEEE 6, 54–65.
Khodaei, A., 2014. Microgrid optimal scheduling with multi-period islanding constraints. IEEE Trans. Power Syst. 29, 1383–1392.
Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436.
Lewis, F., Liu, D., 2009. Reinforcement Learning and Approximate Dynamic Programming for Feedback Control. IEEE Press.
Lpakchi, A., Albuyeh, F., 2009. Grid of the future: are we ready to transition to a smart grid. IEEE Power Energy Mag. 7, 52–56.
Nodland, D., Zargarzadeh, H., Jagannathan, S., 2013. Neural network-based optimal adaptive output feedback control of a helicopter UAV. IEEE Trans. Neural Networks Learn. Syst. 24, 1061–1073.
Palma-Behnke, R., Benavides, C., Lanas, F., Severino, B., Reyes, L., Llanos, J., Sáez, D., 2013. A microgrid energy management system based on the rolling horizon strategy. IEEE Trans. Smart Grid 4, 996–1006.
Pourmousavi, S.A., Nehrir, M.H., Colson, C.M., Wang, C., 2010. Real-time energy management of a stand-alone hybrid wind-microturbine energy system using particle swarm optimization. IEEE Trans. Sustain. Energy 1, 193–201.
Powell, W.B., 2009. Approximate Dynamic Programming: Solving the Curses of Dimensionality. Wiley.
Salinas, S., Li, M., Li, P., Fu, Y., 2013. Dynamic energy management for the smart grid with distributed energy resources. IEEE Trans. Smart Grid 4, 2139–2151.
Shi, W., Xie, X., Chu, C.C., Gadh, R., 2015. Distributed optimal energy management in microgrids. IEEE Trans. Smart Grid 6, 1137–1146.
Siano, P., Cecati, C., Yu, H., Kolbusz, J., 2012. Real time operation of smart grids via FCN networks and optimal power flow. IEEE Trans. Ind. Inform. 8, 944–952.
Sun, S., Dong, M., Liang, B., 2015. Joint supply, demand, and energy storage management towards microgrid cost minimization. In: IEEE International Conference on Smart Grid Communications, pp. 109–114.
Tang, G.Y., Zhao, Y.D., Zhang, B.L., 2007. Optimal output tracking control for nonlinear systems via successive approximation approach ☆. Nonlinear Anal. Theor. Meth. Appl. 66, 1365–1377.
Vamvoudakis, K., Vrabie, D., Lewis, F., 2009. Online Policy Iteration Based Algorithms to Solve the Continuous-time Infinite Horizon Optimal Control Problem, pp. 36–41.
Vamvoudakis, K.G., Lewis, F.L., 2010. Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem ☆. Automatica 46, 878–888.
Vamvoudakis, K.G., Lewis, F.L., 2012. Online solution of nonlinear two-player zero-sum games using synchronous policy iteration. Int. J. Robust Nonlinear Control 22, 1460–1483.
Vojdani, A., 2008. Smart integration. IEEE Power Energy Mag. 6, 71–79.
Yadav, V., Padhi, R., Balakrishnan, S.N., 2007. Robust/optimal temperature profile control of a high-speed aerospace vehicle using neural networks. IEEE Trans. Neural Network. 18, 1115.
Yun, L., Tan, C.S., Lo, K.L., 2007. The effects of locational marginal price on electricity market with distributed generation. Proc. Chin. Soc. Electr. Eng.
Zhang, T., Lu, Y., Zhu, Q., Shen, Q., 2012. Adaptive dynamic surface control with unmodeled dynamics. In: Control Conference, pp. 2859–2864.