

A Novel Data-Driven Intelligent Computing Method for the Secure Control of a Benchmark Microgrid System

Shunjiang Wang¹, Yan Zhao², He Jiang², Rong Chen¹, and Lina Cao¹

¹ State Grid Liaoning Electric Power Company Limited
110006 Shenyang, China

wangshunjiang@163.com, chenrong516@126.com, 13700007416@126.com

² School of Renewable Energy, Shenyang Institute of Engineering
110136 Shenyang, China
zhaoyan@sie.edu.cn, jianghescholar@163.com

Abstract. Microgrid is a small-scale cyber-physical system, and it generally suffers from various uncertainties. In this paper, we investigate the secure control problem of a benchmark microgrid with system uncertainties by using data-driven edge computing technology. First, the state-space function of the benchmark microgrid system is formulated, and parameter uncertainties are taken into consideration. Second, a novel data-driven intelligent computing method is derived from the model-based reinforcement learning algorithm, which only requires system data instead of system models. By utilizing this computing method, the optimal control policy can be obtained in the model-free environment. Third, the Lyapunov stability theory is employed to prove that the uncertain microgrid can be asymptotically stabilized under the optimal control policy. Finally, simulation results demonstrate the control performance can be improved by tuning the parameters in the performance index function.

Keywords: edge computing, microgrid system, secure control, reinforcement learning.

1. Introduction

Nowadays, a large-scale power system is generally composed of several distributed microgrids. With the system operating, a variety of unexpected uncertainties, which severely affect the system stability, are inevitable. Especially for microgrids, the system security issue deserves much attention. In this paper, we will study the secure control problem of a benchmark microgrid [5,13,14,23]. This benchmark microgrid consists of three main parts: power generation, loads and distributed energy storages. The power generation includes regular generation (microturbine), and supplies energy for the demands of various loads.

However, due to the intermittent power injection from photovoltaic arrays and sudden change of load demands, the unbalance between power supply and demand may occur, which will cause the frequency fluctuations and threaten the security of the entire microgrid. Thus, we incorporate distributed energy storages (electric vehicles) into this microgrid to compensate the unbalance. The system data can be measured by sensors and transmitted to the management center through the communication module. The whole microgrid is controlled by using the edge computing technology [2,3,4,24]. The schematic diagram of edge computing for the benchmark microgrid system is shown in Fig. 1.

Different from the traditional automatic control, edge computing is more like an intelligent control method which is based on computing and information, and it mainly concerns the control strategies for dispatch and optimization. In [2], by means of road networks, the problems of frequently moving vehicles and network connectivity were analyzed, and then a modified greedy algorithm for vehicle wireless communication was proposed for network optimization. In [3], a holistic framework to attack the QoS prediction was developed in the IoT environment, and authors designed a fuzzy clustering algorithm which was capable of clustering contextual information. In order to fully utilize hidden features in edge computing environment, the work [24] presented a new matrix factorization model with deep features learning via a convolutional neural network. Since the edge computing technology has such powerful merits, this paper will utilize a novel data-driven intelligent computing method for the secure control of the benchmark microgrid system.

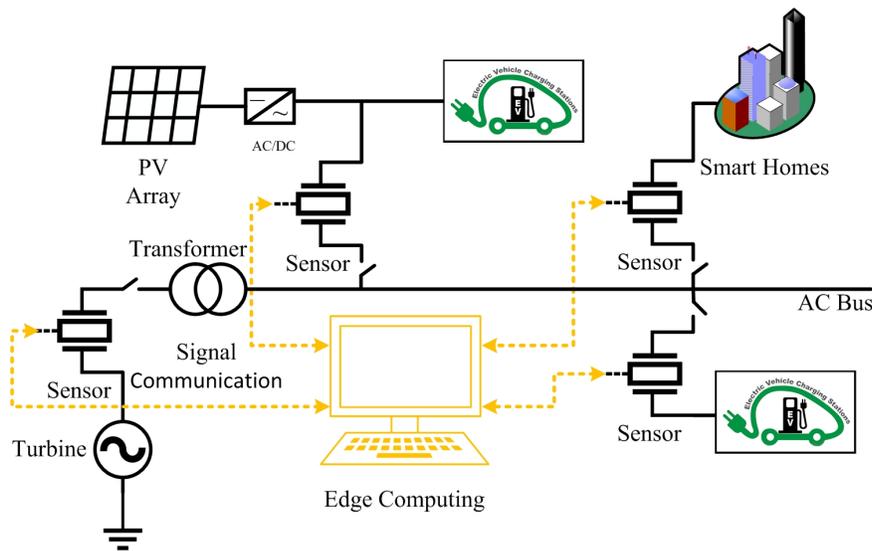


Fig. 1. Schematic diagram of edge computing for the benchmark microgrid system

Previous works regarding the frequency stability issues were mainly based on fuzzy control [5], sliding mode control [13,14], linear matrix inequality (LMI) approach [22] and proportion-integration-differentiation (PID) control [17]. In [12], through modeling the disturbances and parameter uncertainties, an adaptive supplementary control method was proposed for the power system frequency regulation. In [7], a novel second-order sliding mode approach for multi-area power systems was developed by means of an extended disturbance observer. In [19], a new frequency control method was designed for isolated micro-grids via double sliding mode technique. In [16], a second-order sliding mode controller was provided for the power flow control of a hybrid energy storage system. In [21], in order to eliminate the adverse effects of time delays in microgrid, a sliding mode estimation controller was developed to predict time delays and handle the distur-

bance of estimation errors. In [11], to deal with the uncertainties caused by renewable sources, a Takagi-Sugeno fuzzy model was constructed for the microgrid, and a corresponding sliding mode approach was designed. From aforementioned works, we can see the sliding mode control technique is a powerful tool in handling the uncertainties and disturbances.

Unfortunately, these aforementioned works were generally model-based. Due to the existence of system uncertainties, the accurate mathematical models are unavailable, and the model-based control strategies cannot be employed. Therefore, a data-driven secure control method is expected, which motivates the research of this paper.

(1) The proposed data-driven secure control method integrates reinforcement learning, optimal control theory and universal approximator.

(2) This data-driven reinforcement learning method is developed from the model-based policy iteration algorithm. Different from other reinforcement learning methods, it only requires system data instead of system models.

(3) For the secure control issues, the traditional model-based learning approaches will be invalid, because system uncertainties lead to the difficulties in obtaining the accurate mathematical models. As a result, the proposed model-free method becomes the first choice.

In this paper, we investigate the secure control problem of a benchmark microgrid with system uncertainties by using data-driven edge computing technology. The rest of this paper is arranged as follows. First, the problem formulation is given in Section 2. Second, three state-of-the-art reinforcement learning (RL) methods are introduced including policy iteration (PI), value iteration (VI) and a novel data-driven intelligent computing method in Section 3. By utilizing the data-driven computing method, the optimal control policy can be obtained in the model-free environment. Third, in Section 4, the Lyapunov stability theory is employed to prove that the uncertain microgrid can be asymptotically stabilized under the optimal control policy. In Section 5, simulation results demonstrate the control performance can be improved by tuning the parameters in the performance index function. Finally, a brief conclusion is given in Section 6.

2. Problem formulation

The benchmark microgrid [5,13,14,23] investigated in this paper has been introduced. Let us consider the detailed mathematical system model as below

$$\begin{aligned}
 \Delta \dot{\psi}_f &= -\frac{1}{T_p} \Delta \psi_f + \frac{k_p}{T_p} \Delta \psi_t + \frac{k_p}{T_p} \Delta \psi_{v1} + \frac{k_p}{T_p} \Delta \psi_{v2} \\
 \Delta \dot{\psi}_t &= -\frac{1}{T_t} \Delta \psi_t + \frac{1}{T_t} \Delta \psi_g \\
 \Delta \dot{\psi}_g &= -\frac{1}{k_s T_g} \Delta \psi_f - \frac{1}{T_g} \Delta \psi_g + \frac{1}{T_g} u_1 \\
 \Delta \dot{\psi}_{v1} &= -\frac{1}{T_{v1}} \Delta \psi_{v1} + \frac{1}{T_{v1}} u_2 \\
 \Delta \dot{\psi}_{v2} &= -\frac{1}{T_{v2}} \Delta \psi_{v2} + \frac{1}{T_{v2}} u_3
 \end{aligned} \tag{1}$$

where $\Delta\psi_f$ denotes the frequency deviation; $\Delta\psi_t$ is the turbine power; $\Delta\psi_g$ is the governor position value; $\Delta\psi_{v1}$ represents the first electric vehicle power; $\Delta\psi_{v2}$ is the second electric vehicle power; T_t denotes the time constant of turbine; T_g is the time constant of governor; T_p represents the time constant of power system; T_{v1} is the time constant of the first electric vehicle; T_{v2} denotes the time constant of the second electric vehicle; k_p is the gain of power system; k_s represents the speed regulation coefficient; u_1, u_2, u_3 are the control inputs.

Let $x = [\Delta\psi_f, \Delta\psi_t, \Delta\psi_g, \Delta\psi_{v1}, \Delta\psi_{v2}]^T$ and $u = [u_1, u_2, u_3]^T$. The nominal system can be rewritten as

$$\dot{x} = Ax + Bu \tag{2}$$

where $A = \begin{bmatrix} -\frac{1}{T_p} & \frac{k_p}{T_p} & 0 & \frac{k_p}{T_p} & \frac{k_p}{T_p} \\ 0 & -\frac{1}{T_t} & \frac{1}{T_t} & 0 & 0 \\ -\frac{1}{k_s T_g} & 0 & -\frac{1}{T_g} & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{T_{v1}} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_{v2}} \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{T_g} & 0 & 0 \\ 0 & \frac{1}{T_{v1}} & 0 \\ 0 & 0 & \frac{1}{T_{v2}} \end{bmatrix}$.

As is known, a microgrid system is composed of several different units and contains complex structures. With the system operating, internal faults and external disturbances may lead to the change of system structures or the deviation of system parameters. Unfortunately, system uncertainties are generally inevitable, which may affect the control performance. The system (2) with parameter uncertainties can be described by

$$\dot{x} = (A + \Delta A)x + (B + \Delta B)u \tag{3}$$

where u is the secure control policy, which will be designed later. The parameter uncertainties ΔA and ΔB are bounded by $\|\Delta A\| \leq \Delta A_m$ and $\|\Delta B\| \leq \Delta B_m$, respectively.

For the microgrid system, there are two important indexes. One is for the system states, because the frequency deviation should be strictly limited. The other one is for the control inputs, because large control inputs may damage electrical elements and waste unnecessary energies. Therefore, we define the performance index function as

$$J(x(0), u) = \int_0^\infty r(x(\tau), u(\tau))d\tau \tag{4}$$

where $r(x, u) = x^T Qx + u^T Ru$ with positive definite symmetric matrices Q and R . The matrix Q determines the oscillation amplitude of system states, and the matrix R determines the cost of control inputs.

Given the admissible control policy $u(x)$, the value function is expressed as

$$V(x(t)) = \int_t^\infty r(x(\tau), u(x(\tau)))d\tau. \tag{5}$$

In the classical control theory, the optimal control problem is to find out a state feedback control policy which can minimize the value function. Consequently, the optimal value function can be defined as

$$V^*(x(t)) = \min_u \left(\int_t^\infty r(x(\tau), u(x(\tau)))d\tau \right). \tag{6}$$

According to the stationarity condition [18], the optimal control policy is derived by

$$u^*(x) = -\frac{1}{2}R^{-1}(B + \Delta B)^T \nabla V^*(x) \quad (7)$$

where $\nabla V^*(x) = \partial V^*(x)/\partial x$ and $V^*(x)$ should satisfy the following Hamilton-Jacobi-Bellman (HJB) equation

$$0 = r(x, u^*(x)) + \nabla V^{*T}(x)[(A + \Delta A)x + (B + \Delta B)u^*(x)]. \quad (8)$$

From the aspect of engineering, one needs to settle the HJB equation to attain the optimal control strategy. From the aspect of theory, the HJB equation is a complex partial differential equation, and it is difficult or even impossible to obtain its analytical solution. To solve the HJB equation, three algorithms will be introduced in the next section.

3. A brief overview of RL algorithms

In order to achieve the optimal control policy, three iterative RL algorithms including PI, VI and off-policy method will be reviewed in this section.

3.1. PI algorithm

Inspired by previous works [8,10,18,27], a model-based PI algorithm is given in the following Algorithm 1. By using this algorithm, one can obtain $V^*(x)$ and $u^*(x)$ as $i \rightarrow \infty$.

Algorithm 1: PI-based RL method

Step 1: (Initialization)

Let the iteration index $i = 0$.

Select a small enough computation precision ϵ .

Choose an initial admissible control policy $u^{(0)}(x)$.

Step 2: (Policy Evaluation)

With $u^{(i)}(x)$, compute $V^{(i+1)}(x)$ by

$$0 = r(x, u^{(i)}(x)) + (\nabla V^{(i+1)}(x))^T((A + \Delta A)x + (B + \Delta B)u^{(i)}(x)). \quad (9)$$

Step 3: (Policy Improvement)

With $V^{(i+1)}(x)$, update the control policy $u^{(i+1)}(x)$ by

$$u^{(i+1)}(x) = -\frac{1}{2}R^{-1}(B + \Delta B)^T \nabla V^{(i+1)}(x). \quad (10)$$

Step 4:(Termination)

If $\|V^{(i+1)}(x) - V^{(i)}(x)\| \leq \epsilon$ on the given compact set, stop at this step;

Else, let $i = i + 1$ and go back to Step 2.

Due to the easy-to-realize structure and fast convergence, PI method is popular in the field of computer sciences. It starts from an admissible control input, and gradually approaches to the optimal solution by the steps of policy evaluation and policy improvement.

3.2. VI-based integral RL algorithm

The aforementioned Algorithm 1 requires initial admissible control. Although this initial condition simplifies the process of finding optimal solutions and speeds up algorithm convergence [9], it is impractical for some complex systems. Inspired by previous works [6,20], we present the following VI-based integral RL algorithm which is not limited by the initial admissible condition.

Algorithm 2: VI-based integral RL algorithm

Step 1: (Initialization)

Let the iteration index $i = 0$.
 Set a computation precision ϵ .
 Choose an initial value function $V^{(0)}(x) \geq 0$.

Step 2: (Policy Improvement)

With $V^{(i)}(x)$, compute $u^{(i)}(x)$ by

$$u^{(i)}(x) = -\frac{1}{2}R^{-1}(B + \Delta B)^T \nabla V^{(i)}(x). \tag{11}$$

Step 3: (Policy Evaluation)

With $u^{(i)}(x)$, update the value function $V^{(i+1)}(x)$ by

$$V^{(i+1)}(x(t)) = \int_t^{t+\Delta t} r(x(\tau), u^{(i)}(x(\tau)))d\tau + V^{(i)}(x(t + \Delta t)). \tag{12}$$

Step 4:(Termination)

If $\|V^{(i+1)}(x) - V^{(i)}(x)\| \leq \epsilon$ on the given compact set, stop at this step;
 Else, let $i = i + 1$ and go back to Step 2.

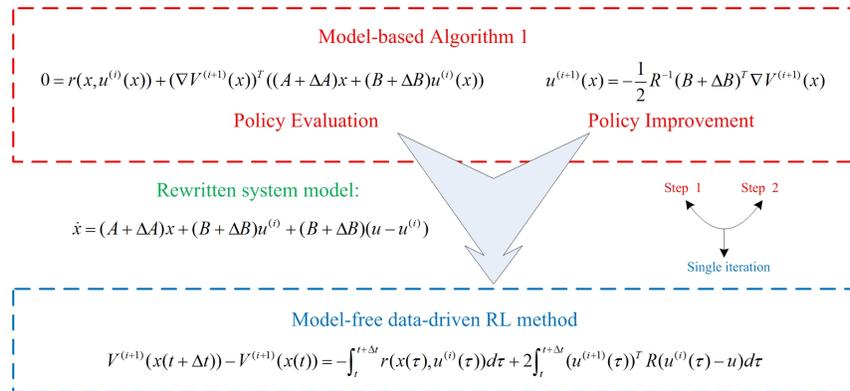


Fig. 2. Derivation diagram of data-driven RL method

Remark 1 Algorithm 1 is completely model-based, and it requires the knowledge of both ΔA and ΔB . Algorithm 2 is partially model-based, and it only needs the knowledge of

ΔB . Unfortunately, the accurate models of system uncertainties ΔA and ΔB are generally unavailable. As a result, both Algorithm 1 and Algorithm 2 are not practical for the real-world engineering. A data-driven method is expected, which only needs system data instead of accurate system models. The derivation diagram of data-driven RL method is shown in Fig. 2. Two iteration steps of Algorithm 1 are combined in a single step in the data-driven RL method. Through this derivation, the system models are avoided.

3.3. Data-driven RL method

In this subsection, we will introduce a model-free method, also called off-policy RL algorithm [10,15]. In order to derive this algorithm, let us rewrite the system (3) as

$$\dot{x} = (A + \Delta A)x + (B + \Delta B)u^{(i)} + (B + \Delta B)(u - u^{(i)}). \tag{13}$$

By means of (13), we have

$$\begin{aligned} \frac{dV^{(i+1)}(x)}{dt} &= (\nabla V^{(i+1)}(x))^T [(A + \Delta A)x + (B + \Delta B)u^{(i)}] \\ &\quad + (\nabla V^{(i+1)}(x))^T (B + \Delta B)(u - u^{(i)}). \end{aligned} \tag{14}$$

In light of (9) and (10) in Algorithm 1, Equation (14) can be rewritten as

$$\frac{dV^{(i+1)}(x)}{dt} = -r(x, u^{(i)}(x)) + 2(u^{(i+1)}(x))^T R(u^{(i)}(x) - u). \tag{15}$$

Integrating both sides of (15) on the interval $[t \ t + \Delta t]$ yields

$$\begin{aligned} &V^{(i+1)}(x(t + \Delta t)) - V^{(i+1)}(x(t)) \\ &= - \int_t^{t+\Delta t} r(x(\tau), u^{(i)}(\tau))d\tau + 2 \int_t^{t+\Delta t} (u^{(i+1)}(\tau))^T R(u^{(i)}(\tau) - u)d\tau. \end{aligned} \tag{16}$$

Remark 2 It can be seen that the policy evaluation step in Algorithm 2 is more easy-to-realize than that in both Algorithm 1 and data-driven RL method. The VI-based integral RL algorithm provides a choice to solve the optimal control issue when the admissible control is unavailable. Although Algorithm 2 contains a more relaxed initialization condition, it also has several drawbacks compared with other two algorithms. By means of initial admissible control, PI-based methods achieve faster convergence [9]. Furthermore, all the iterative control policies in the PI learning procedure are admissible, which cannot be guaranteed in the VI method [8]. The data-driven RL method is completely model-free, while Algorithm 2 is partially model-free. That is because the control input matrix B is still required in Algorithm 2. Actually, when accurate system models and admissible control policies are available, the traditional PI method, i.e., Algorithm 1, is a more convenient choice. In conclusion, RL has given enough choices for different situations. The data-driven RL algorithm can cover most application demands, and the VI-based integral RL algorithm can be used without initial admissible control. The detailed performance comparisons among three iterative RL methods are shown in Fig. 3.

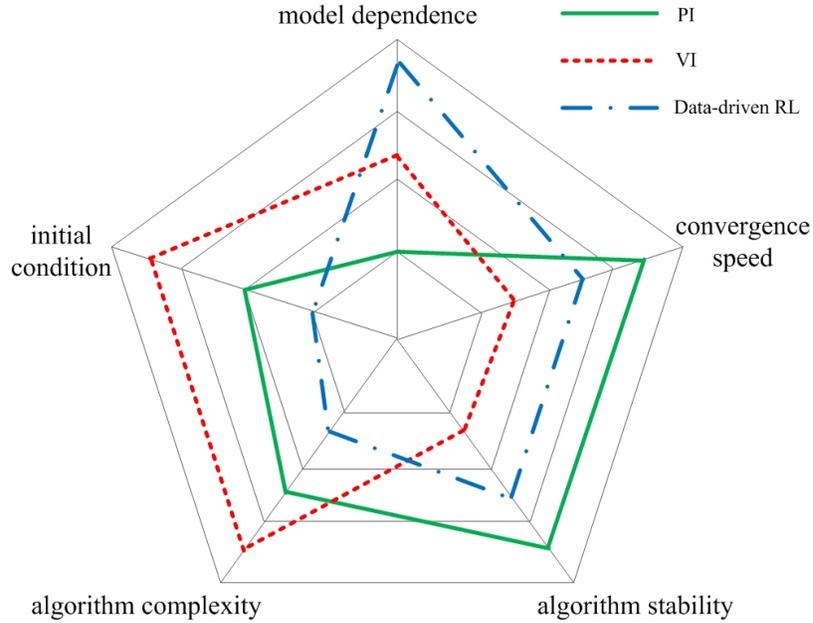


Fig. 3. Performance comparisons among three state-of-the-art RL methods

4. Algorithm implementation and stability analysis

As is known, neural network (NN) has been proved to be a powerful universal approximator. Hence, NNs are generally utilized to implement the aforementioned RL algorithms. A critic NN and an actor NN are constructed to approximate the value function and control policy, respectively. Each NN consists of two parts including tuned NN weights and NN activation functions. In this section, we will provide another implementation tool called generalized fuzzy hyperbolic model (GFHM) [1,25,26]. It also has the property of the universal approximation, and can be finally converted into the similar form as NNs.

Definition 1 [25,26] Let $x = [x_1, x_2, \dots, x_n]^T$ be the model input and y be the single output. $\bar{x} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m]^T$ represents the generalized input, where $\bar{x}_i = x_z - d_{zj}$. w_z denotes the number of transformation about x_z with $z = 1, 2, \dots, n$. d_{zj} is the transformation constant for x_z with $j = 1, 2, \dots, w_z$. $m = \sum_{z=1}^n w_z$ denotes the total number of generalized input variables. The generalized fuzzy hyperbolic rule base should satisfy the following terms:

1. The p th fuzzy rule is expressed as
 IF $(x_1 - d_{11})$ is $F_{x_{11}}$ and \dots and $(x_1 - d_{1w_1})$ is $F_{x_{1w_1}}$ and $(x_2 - d_{21})$ is $F_{x_{21}}$ and \dots
 and $(x_2 - d_{2w_2})$ is $F_{x_{2w_2}}$ and \dots and $(x_n - d_{n1})$ is $F_{x_{n1}}$ and \dots and $(x_n - d_{nw_n})$
 is $F_{x_{nw_n}}$,
 THEN $y^p = c_{F_{11}} + \dots + c_{F_{1w_1}} + c_{F_{21}} + \dots + c_{F_{2w_2}} + \dots + c_{F_{n1}} + \dots + c_{F_{nw_n}}$,

where $F_{x_{z_j}}$ is the fuzzy set corresponding to $x_z - d_{z_j}$ including P_x (Positive) and N_x (Negative) subsets.

2. $c_{F_{z_j}}$ in the “THEN” part should be associated with $F_{x_{z_j}}$ in the “IF” part. That is, $c_{F_{z_j}}$ and $F_{x_{z_j}}$ should exist or disappear at the same time.
3. There should be 2^m fuzzy rules containing all the possible combinations in both P_x and N_x subsets.

The selection of the membership function is important, which should be helpful for the derivation of the following Lemma 1. Hence, the membership functions are given by

$$\begin{aligned} \mu_{P_i}(\bar{x}_i) &= e^{-\frac{1}{2}(\bar{x}_i - q_i)^2}, \\ \mu_{N_i}(\bar{x}_i) &= e^{-\frac{1}{2}(\bar{x}_i + q_i)^2} \end{aligned} \tag{17}$$

where q_i is a positive constant.

Lemma 1 [1,25] *Let the membership functions be given by (17) and the generalized fuzzy hyperbolic rule base be described by Definition 1. Then, the GFHM can be derived by*

$$\begin{aligned} y(x) &= \sum_{i=1}^m \frac{c_{P_i} e^{q_i \bar{x}_i} + c_{N_i} e^{-q_i \bar{x}_i}}{e^{q_i \bar{x}_i} + e^{-q_i \bar{x}_i}} \\ &= \sum_{i=1}^m \sigma_i + \sum_{i=1}^m \varpi_i \frac{e^{q_i \bar{x}_i} - e^{-q_i \bar{x}_i}}{e^{q_i \bar{x}_i} + e^{-q_i \bar{x}_i}} \\ &= \theta + \rho^T \tanh(Y \bar{x}) \end{aligned} \tag{18}$$

where $\sigma_i = \frac{c_{P_i} + c_{N_i}}{2}$, $\theta = \sum_{i=1}^m \sigma_i$, $\rho = [\varpi_1, \varpi_2, \dots, \varpi_m]^T$, $Y = \text{diag}\{q_1, q_2, \dots, q_m\}$, $\varpi_i = \frac{c_{P_i} - c_{N_i}}{2}$ and $\tanh(Y \bar{x}) = [\tanh(q_1 \bar{x}_1), \tanh(q_2 \bar{x}_2), \dots, \tanh(q_m \bar{x}_m)]^T$.

The purpose of presenting Lemma 1 is to construct the similar expression form as NNs. According to (18), the GFHM can be further rewritten as

$$y(x) = \phi^T(x)W \tag{19}$$

where $W = [\theta, \rho^T]^T$ and $\phi(x) = [1, \tanh(q_1 \bar{x}_1), \tanh(q_2 \bar{x}_2), \dots, \tanh(q_m \bar{x}_m)]^T$. Here, GFHM gets the same expression form as NNs, where W can be seen as NN weights and $\phi(x)$ is similar to the NN activation function. Lemma 1 has provided the NN expression form for GFHM. Next, we will present Lemma 2 to demonstrate its property of the universal approximation.

Lemma 2 [1,25] *For arbitrary continuous function $f(x)$ on the compact set Ω and arbitrary constant $\epsilon > 0$, there exists at least one GFHM such that $\sup_{x \in \Omega} |f(x) - y(x)| < \epsilon$.*

Based on Lemma 2, the GFHM has been proved to be a universal approximator. By means of the form of (19), the GFHM provides another choice to implement RL algorithms besides NNs. Especially for dealing with nonlinear systems, the GFHM is a powerful tool in identifying nonlinear continuous functions [25].

Next, we will propose the following theorem for the stability analysis.

Theorem 1 *If the optimal control policy $u^*(x)$ is employed, then the uncertain system (3) can be asymptotically stabilized.*

Proof. Choose the following Lyapunov function candidate:

$$V = V^*(x). \quad (20)$$

According to the HJB equation (8), we can obtain

$$\begin{aligned} \dot{V} &= \dot{V}^*(x) \\ &= \nabla V^{*T}(x)[(A + \Delta A)x + (B + \Delta B)u^*(x)] \\ &= -x^T Qx - u^{*T}(x)Ru^*(x). \end{aligned} \quad (21)$$

From (21), it is obvious that $\dot{V} \leq 0$. Therefore, according to the Lyapunov stability theory, the uncertain system (3) can be asymptotically stabilized under the optimal control policy $u^*(x)$. Furthermore, by tuning the parameters Q and R , the control performance can be improved, which will be demonstrated in the simulation part.

The proof is completed. ■

5. Simulation result

In order to illustrate the effectiveness of our proposed scheme, we present the following simulation result. The values of system parameters for the numerical simulation are shown in Table 1.

Table 1. Values of system parameters

Parameters	T_g	T_i	T_p	T_{v1}	T_{v2}	k_s	k_p
Values	0.5	1	1	0.5	0.5	1	2

Through Table 1, the matrices A and B can be obtained. Let the system uncertainties be $\Delta A = [0.1, 0, 0, 0, -0.2; 0, 0.1, 0, 0, 0; 0, 0, 0.2, 0, 0; 0, 0, 0, 0.2, 0; 0, 0, 0, 0, 0.2]$ and $\Delta B = [0, 0, 0; 0, 0, 0; -0.2, 0, 0; 0, -0.2, 0; 0, 0, -0.2]$.

Set the system initial values $x(0) = [0.2; -0.3; 0.2; -0.1; 0.1]$. First, we present the simulation result without any control inputs in Fig. 4(a), where it is observed that the system without control inputs cannot be stabilized after 8 seconds.

Second, we set $Q = 0.2I_5$ and $R = I_3$, and apply the associated optimal control policy to the uncertain system. The simulation result is shown in Fig. 4(b), where we can see the system with the optimal control policy can be stabilized after 8 seconds. Third, we set $Q = 20I_5$ and $R = I_3$, and apply the associated optimal control policy to the uncertain system. The simulation result is shown in Fig. 4(c), where it can be observed that the system states under the optimal control with $Q = 20I_5$ get convergence within 4 seconds. Through the simulation results of Fig. 4, we can verify the validity of Theorem 1.

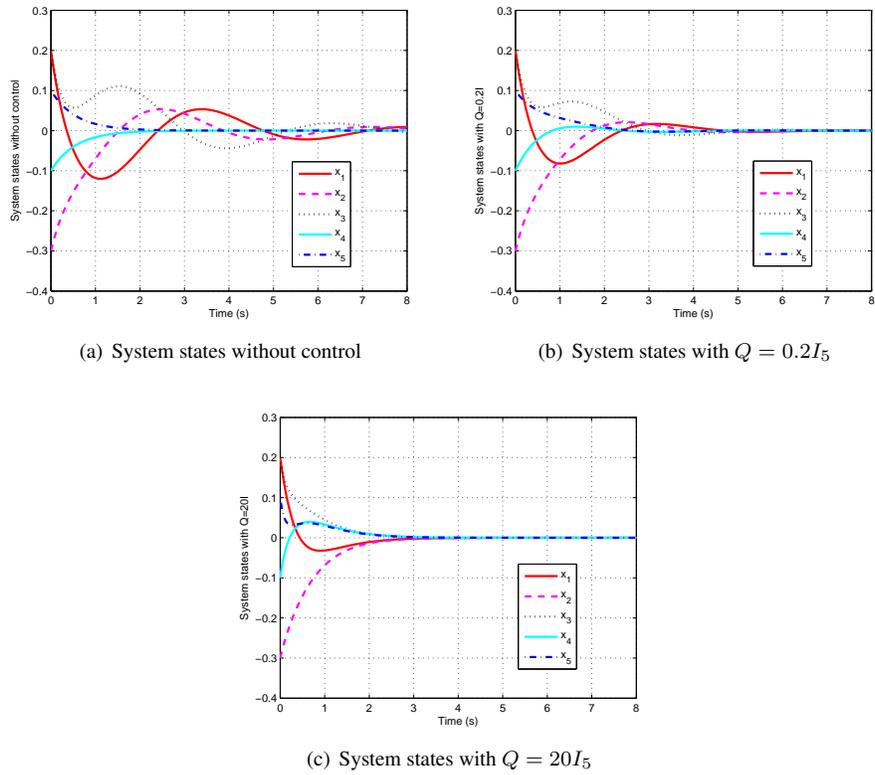


Fig. 4. Comparisons of control performance among different conditions

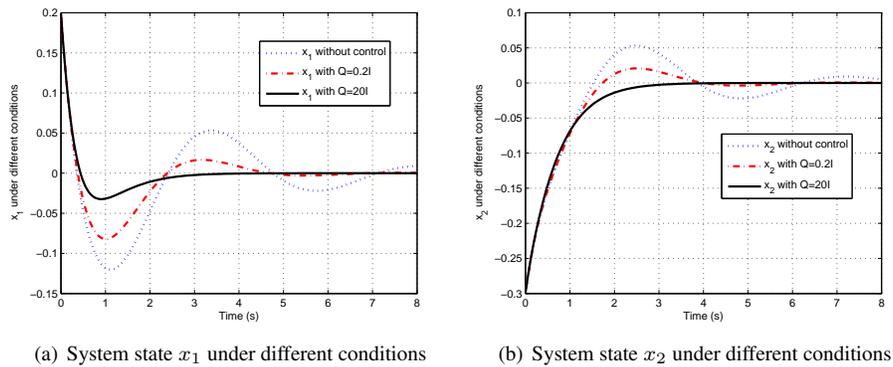


Fig. 5. System states x_1 and x_2 under different conditions

In addition, we present the system states x_1 and x_2 under different conditions in Fig. 5(a) and Fig. 5(b), respectively. From Fig. 5, we can see the system states x_1 and x_2 with $Q = 20I_5$ achieve faster convergence and better control performance than them without control or with $Q = 0.2I_5$, which implies the control performance can be improved by tuning the parameter Q in the performance index function.

From the simulation results, we can see the secure control strategy relies on the optimal control theory, which implies the optimal control policy not only has the optimality but also the robustness. The matrix Q in the performance index function plays an important role in handling the system uncertainties and state deviation. Once the optimal control policy is applied to the system, the security as well as the optimality will be guaranteed by tuning the matrix Q . Different from other secure control methods, the proposed control scheme not only concerns the robustness but also the optimality of the entire control process.

6. Conclusion

In this paper, the secure control problem of a benchmark microgrid with system uncertainties has been investigated by using data-driven edge computing technology. The corresponding mathematical problem formulation has been derived and established. Three state-of-the-art RL methods have been reviewed in details. The stability analysis has been presented through the Lyapunov stability theory, which indicates the uncertain microgrid can be asymptotically stabilized under the optimal control policy. Furthermore, simulation results have demonstrated the control performance can be significantly improved by tuning the parameters in the performance index function. In the future, it is expected that our proposed scheme can be applied to other information systems.

Acknowledgments. This work was supported by Liaoning Revitalization Talents Program (XLYC1907138), the Natural Science Foundation of Liaoning Province (2019-MS-239), the Doctoral Scientific Research Foundation of Liaoning Province (2020-BS-181), the Technology Innovation Talent Fund of Shenyang (RC190360) and Liaoning BaiQianWan Talents Program.

References

1. Cui, Y., Zhang, H., Wang, Y., Gao, W.: Adaptive control for a class of uncertain strict-feedback nonlinear systems based on a generalized fuzzy hyperbolic model. *Fuzzy Sets and Systems* 302, 52–64 (2016)
2. Gao, H., Liu, C., Li, Y., Yang, X.: V2VR: Reliable hybrid-network-oriented V2V data transmission and routing considering RSUs and connectivity probability. *IEEE Transactions on Intelligent Transportation Systems* DOI: 10.1109/TITS.2020.2983835
3. Gao, H., Xu, Y., Yin, Y., Zhang, W., Li, R., Wang, X.: Context-aware QoS prediction with neural collaborative filtering for internet-of-things services. *IEEE Internet of Things Journal* 7(5), 4532–4542 (2020)
4. Gao, H., Duan, Y., Shao, L., Sun, X.: Transformation-based processing of typed resources for multimedia sources in the IoT environment. *Wireless Networks* DOI: 10.1007/s11276-019-02200-6
5. Khooban, M.H., Niknam, T., Blaabjerg, F., Dragicevic, T.: A new load frequency control strategy for micro-grids with considering electrical vehicles. *Electric Power Systems Research* 143, 585–598 (2017)

6. Lewis, F.L., Vrabie, D., Vamvoudakis, K.G.: Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers. *IEEE Control Systems* 32(6), 76–105 (2012)
7. Liao, K., Xu, Y.: A robust load frequency control scheme for power systems based on second-order sliding mode and extended disturbance observer. *IEEE Transactions on Industrial Informatics* 14(7), 3076–3086 (2018)
8. Liu, D., Wei, Q.: Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems* 25(3), 621–634 (2014)
9. Luo, B., Liu, D., Huang, T., Yang, X., Ma, H.: Multi-step heuristic dynamic programming for optimal control of nonlinear discrete-time systems. *Information Sciences* 411, 66–83 (2017)
10. Luo, B., Wu, H.N., Huang, T.: Off-policy reinforcement learning for H_∞ control design. *IEEE Transactions on Cybernetics* 45(1), 65–76 (2015)
11. Mi, Y., Zhang, H., Fu, Y., Wang, C., Loh, P.C., Wang, P.: Intelligent power sharing of DC isolated microgrid based on fuzzy sliding mode droop control. *IEEE Transactions on Smart Grid* 10(3), 2396–2406 (2019)
12. Mu, C., Tang, Y., He, H.: Improved sliding mode design for load frequency control of power system integrated an adaptive learning strategy. *IEEE Transactions on Industrial Electronics* 64(8), 6742–6751 (2017)
13. Mu, C., Liu, W., Xu, W., Rabiul Islam, M.: Observer-based load frequency control for island microgrid with photovoltaic power. *International Journal of Photoenergy* 2017, 1–11 (2017)
14. Mu, C., Tang, Y., He, H.: Observer-based sliding mode frequency control for micro-grid with photovoltaic energy integration. In: 2016 Power and Energy Society General Meeting. pp. 1–5. IEEE
15. Song, R., Lewis, F.L., Wei, Q., Zhang, H.: Off-policy actor-critic structure for optimal control of unknown systems with disturbances. *IEEE Transactions on Cybernetics* 46(5), 1041–1050 (2016)
16. Tabart, Q., Vechiu, I., Etxeberria, A., Bacha, S.: Hybrid energy storage system microgrids integration for power quality improvement using four-leg three-level npc inverter and second-order sliding mode control. *IEEE Transactions on Industrial Electronics* 65(1), 424–435 (2018)
17. Tan, W.: Unified tuning of PID load frequency controller for power systems via IMC. *IEEE Transactions on Power Systems* 25(1), 341–350 (2010)
18. Vamvoudakis, K.G., Lewis, F.L.: Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* 46(5), 878–888 (2010)
19. Wang, C., Mi, Y., Fu, Y., Wang, P.: Frequency control of an isolated micro-grid using double sliding mode controllers and disturbance observer. *IEEE Transactions on Smart Grid* 9(2), 923–930 (2018)
20. Xiao, G., Zhang, H., Qu, Q., Jiang, H.: General value iteration based single network approach for constrained optimal controller design of partially-unknown continuous-time nonlinear systems. *Journal of the Franklin Institute* 355(5), 2610–2630 (2018)
21. Yan, H., Zhou, X., Zhang, H., Yang, F., Wu, Z.: A novel sliding mode estimation for microgrid control with communication time delays. *IEEE Transactions on Smart Grid* 10(2), 1509–1520 (2019)
22. Yang, F., He, J., Wang, D.: New stability criteria of delayed load frequency control systems via infinite-series-based inequality. *IEEE Transactions on Industrial Informatics* 14(1), 231–240 (2018)
23. Yang, J., Zeng, Z., Tang, Y., Yan, J., He, H., Wu, Y.: Load frequency control in isolated microgrids with electrical vehicles based on multivariable generalized predictive theory. *Energies* 8(3), 2145–2164 (2015)
24. Yin, Y., Chen, L., Xu, Y., Wan, J., Zhang, H., Mai, Z.: QoS prediction for service recommendation with deep feature learning in edge computing environment. *Mobile Networks and Applications* DOI: 10.1007/s11036-019-01241-7

25. Zhang, H., Cui, Y., Wang, Y.: Hybrid fuzzy adaptive fault-tolerant control for a class of uncertain nonlinear systems with unmeasured states. *IEEE Transactions on Fuzzy Systems* 25(5), 1041–1050 (2017)
26. Zhang, H., Zhang, J., Yang, G.H., Luo, Y.: Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming. *IEEE Transactions on Fuzzy Systems* 23(1), 152–163 (2015)
27. Zhao, D., Zhang, Q., Wang, D., Zhu, Y.: Experience replay for optimal control of nonzero-sum game systems with unknown dynamics. *IEEE Transactions on Cybernetics* 46(3), 854–865 (2016)

Shunjiang Wang received the Ph.D. degree in computer application technology in 2019 from University of Chinese Academy of Sciences, Shenyang, China. He is now working at Power Dispatching and Control Center, State Grid Liaoning Electric Power Company Limited. His current research interests include power system automation and clean energy accommodation.

Yan Zhao received the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2008. He is currently a professor with the School of Renewable Energy, Shenyang Institute of Engineering, Shenyang. His current research interests include sliding mode control and integrated energy system.

He Jiang received the Ph.D. degree in control theory and control engineering in 2019 from Northeastern University, Shenyang, China. He is now working at School of Renewable Energy, Shenyang Institute of Engineering. His current research interests include adaptive dynamic programming, sliding mode control, multi-agent system control and their industrial applications.

Rong Chen is now working at State Grid Liaoning Electric Power Company Limited. Her current research interests include power system automation and clean energy accommodation.

Lina Cao is now working at State Grid Liaoning Electric Power Company Limited. Her current research interests include economic dispatch and automatic control of power systems.

Received: September 12, 2019; Accepted: June 8, 2020.