Energy Proceedings
Vol 32, 2023
ISSN 2004-2965

Deep reinforcement learning based adaptive energy management for islanded microgrids considering multi-objective optimization

Jiawen Li[1*] , Tao Zhou[1]

1 Shanghai University of Electric Power, Shanghai, 201306; (corresponding author, e-mail: lijiawen@shiep.edu.cn)

ABSTRACT

In order to reduce the frequency deviation and unit generation cost of an isolated microgrid, an adaptive load frequency control (ALFC) method is proposed in this paper. The method employs an adaptive proportional-integral-derivative (PID) controller to achieve adaptive control by adaptively adjusting the parameters of the controller to output the regulation command. In addition, to achieve adaptive regulation of the control parameters, a deep actor-critic (DAC) algorithm is proposed in this paper, which introduces multiple critics and Gaussian noise exploration techniques to enhance the quality of the ALFC strategy. The performance of the proposed method is tested in the Zhuzhou isolated microgrid of the China Southern Grid(CSG), which can effectively reduce frequency deviation and generation cost.

Keywords: energy management, load frequency control, deep reinforcement learning, frequency deviation, generation cost

1. INTRODUCTION

The high penetration rate of intermittent energy sources such as wind and solar power in microgrid, and the increasing number of home temperature-controlled loads such as air conditioners, refrigerators, and electric water heaters [1] cause power imbalance between the generation power and the consumer power in the microgrid system, resulting in power fluctuations in the microgrid and large changes in the frequency shift of the microgrid. A large or small deviation of the frequency measurement from the frequency rating in the microgrid can have a serious impact on the safety and efficiency of the demand-side power equipment and the generation-side equipment in the microgrid system [2]. Therefore, frequency is an important factor in measuring the power quality of a microgrid system and an important indicator to maintain the stability of the system operation.

Loadfrequency control (LFC) is the core of grid frequency regulation technology.The basic task of LFC is to regulate the power system supply frequency to keep it at the reference value [3], which is an important means to ensure the quality of power supply and the safe, reliable and economic operation of the power system.

The increasing use of renewable energy sources, taking into account environmental and economic factors, poses many new challenges for the control of power systems due to the uncertainty and intermittency of renewable energy output. In microgrids with high penetration of new energy sources, the operating conditions change frequently and require higher responsiveness of existing economic optimization solutions. The distributed control strategy does not require a central controller and uses a point-to-point sparse communication topology to reduce the computational complexity by decentralizing the global optimization objectives to the local controller. Compared with the centralized control strategy, the distributed control enhances the system reliability and avoids the single point of failure [4].

Xu et al. [5] proposed a fully distributed load restoration algorithm based on MAS for solving the load restoration problem after fault clearing in microgrids. In the designed algorithm, each agent only interacts with its neighbors for information, and the global information is obtained through the consistency algorithm, and each agent makes load recovery decisions synchronously. The method of setting coefficients in the algorithm is also given. Zhang et al. [6] used the average consistency algorithm to achieve the global sharing of load demand in an isolated microgrid with high wind power penetration, and controlled the output of each wind turbine to meet the power supply and demand balance. And the PI controller is used to compensate the possible control deviation.Xu et al. [7] proposed a distributed subgradient-based algorithm to coordinate the operation of different kinds of distributed microsources in the microgrid for the drawback of traditional sag control with large oscillation in dynamic response and steady-state deviation. The effect of system frequency deviation is attributed to the utilization level control of

This is a paper for Applied Energy Symposium 2023: Clean Energy towards Carbon Neutrality (CEN2023), April 23-25, 2023, Ningbo, China.

each unit, which can well maintain the power balance and significantly improve the system dynamic performance. The above control strategy only ensures the balance of power supply and demand in the microgrid, without considering the economic operation.

The above distributed LFC strategy can indeed achieve fast LFC, but the immaturity of the technology and the large amount of communication required between units and units lead to this type of method instead to cause improper regulation of the units, which in turn leads to system frequency fluctuations.

Conventional centralized LFC strategies are widely used in islanded microgrids because of their centralized control, low communication and simple management. This control method is simple, reliable, and easy to implement, and can achieve ideal steady-state results. However, due to the presence of nonlinear constraints in the power system such as generation rate constraint, valve position constraint, and the constant change of parameters in the system, the effect of PID control method is not satisfactory [8]. Sahu et al. [10] proposed to use PSO algorithm for global and individual optimization of the system to obtain the optimal target value.Saikia et al. [11] used a two-degree-of-freedom PID algorithm with two weight factors for LFC control. However, these PID controller parameters usually remain constant once they are set, and the control effect of fixed parameter PIDs is greatly reduced when the system parameters change dynamically.

To cope with the shortcomings of PID algorithms, some scholars have proposed intelligent LFC methods. liu et al. [12] proposed a neural network predictive fuzzy control. kocaarslan et al. [13] proposed a fuzzy gain schedule proportional-integral controller. Li et al. [14] proposed a robust sliding mode control.

However, the conventional centralized LFC still lacks adaptive control capability to adapt to complex islanded microgrids [15].

With the development of artificial intelligence, deep reinforcement learning is applied to islanded microgrid control [16-18]. One of the actor-critic algorithms is widely used because of its strong adaptive decision-making capability. However, the lack of robustness of such methods leads to their insufficient performance. To cope with this problem, it is necessary to combine the pervasiveness of PID control algorithms with the strong adaptive decision making capability of actor critic algorithms and adopt techniques to improve the algorithm robustness to adapt to complex islanded microgrids and thus achieve efficient LFC.

In this paper, an adaptive load frequency control (ALFC) method is proposed. The method uses an adaptive proportional-integral-derivative controller to achieve adaptive control by adaptively adjusting the parameters of the controller to output the regulation commands of individual units. In addition, to realize the adaptive regulation of control parameters, a deep actor-critic (DAC) algorithm is proposed in this paper, which introduces multiple critics and Gaussian noise exploration techniques to enhance the quality of the ALFC strategy. The performance of the proposed method is tested in the Zhuzhou isolated microgrid of the Southern Power Grid, which can effectively reduce frequency deviation and generation cost.

The contribution of this paper is shown below.

1) An adaptive load frequency control (ALFC) method This day proposes to use an adaptive proportional-integral-derivative controller to achieve adaptive control by adaptively adjusting the parameters of the controller.

2) A highly robust DAC algorithm is proposed to introduce multiple critics and Gaussian noise exploration techniques to improve the quality of ALFC strategies.

3) ALFC can effectively reduce frequency deviation and generation cost to achieve multi-objective optimization.

The organization of the paper is shown as follows: Section 2 describes the islanded microgrid model, Section 3 describes the DAC algorithm, Section 4 describes the case study, and Section 5 gives the conclusion.

2. ISLAND MICROGRID STRUCTURE

2.1 Basic composition of island microgrid

Microgrids are composed of various types of micro power sources. In this paper, the principles of photovoltaic cells, wind turbines, micro gas turbines, and diesel generators are introduced.

1) Micro gas turbine

In this paper, we study the islanding control of microgrid. In islanding operation, micro gas turbines play a crucial role in maintaining the safe and stable operation of the microgrid. In terms of adjustable capacity range, fast response characteristics, safety and economy, this type of turbine can be used as the main frequency regulation unit. The LFC model of a typical micro gas turbine and a small single-cycle turbine is shown in the literature [3].

2) Diesel generator

Diesel generator is a kind of small power generation equipment, refers to the diesel fuel, etc. as the fuel, the diesel engine as the engine to drive the generator power generation power machinery. It is a standby power supply to compensate for power shortage, whether it is the main control or standby, can play a vital role. Its frequency regulation is directly related to the fuel regulation, so that the fuel/generating power is in a balanced state. If the frequency changes with the system load, the speed fluctuation can be avoided by adjusting the oil port opening or the stepper motor rack position to reasonably increase the fuel supply. Since the speed is proportional to the frequency, the frequency control method means that the governor self-regulates the fuel cycle supply according to the generator speed, so that the speed is always within the required range [3]. The characteristics are shown as follows.

$$\Delta\dot{f} = -\frac{1}{T_{\mathrm{p}}}\Delta f + \frac{K_{\mathrm{p}}}{T_{\mathrm{p}}}\left(\Delta P_{\mathrm{d}} + \Delta P_{\mathrm{PV}} + \Delta P_{\mathrm{B}} - \Delta P_{\mathrm{L}}\right) \tag{1}$$

$$\Delta\dot{P}_{\mathrm{d}} = -\frac{1}{T_{\mathrm{t}}}\Delta P_{\mathrm{d}} + \frac{1}{T_{\mathrm{t}}}\Delta X_{\mathrm{d}} \tag{2}$$

$$\Delta\dot{X}_{\mathrm{d}} = -\frac{1}{T_{\mathrm{g}}R_{\mathrm{I}}}\Delta f - \frac{1}{T_{\mathrm{g}}}\Delta X_{\mathrm{d}} - \frac{1}{T_{\mathrm{g}}}\Delta E + \frac{1}{T_{\mathrm{g}}}u \tag{3}$$

$$\Delta\dot{E} = K_{\varepsilon}\Delta f \tag{4}$$

where *Δf, ΔX_d , ΔE is* the microgrid frequency variation, governor valve position variation and integral control variation, respectively. k_{p} is the power system gain, T is the power system time constant K_{p} is the power system gain, T is the power system time constant;

3) Wind turbine

The wind turbine principle is to convert the wind energy into mechanical energy by rotating the turbine blades driven by the natural wind, and the turbine blades drive the generator system through the transmission system, thus converting the mechanical energy into electrical energy. The object of this paper is a variable speed and constant frequency wind turbine, which is the mainstream type of wind turbine nowadays. Compared with other types of wind turbines, the rotor speed can be adjusted and changed so as to fully utilize the wind energy and participate in frequency regulation. Wind turbines mainly include cage asynchronous generators, double-fed asynchronous generators and permanent magnet synchronous generators, and the specific selection stage must consider the wind turbine and speed control method to ensure that the selection results really meet the actual needs. Generally speaking, for fixed pitch constant speed and variable pitch constant frequency wind turbines, the first category is selected for the former, while the latter can be selected as double-fed or permanent magnet type. After a comprehensive comparison and analysis, the direct-drive multi-stage synchronous generator is chosen in this case, taking into account the actual situation. The wind energy capture factor is an important parameter to determine the wind turbine power, and the Bates theory is an important basis for the conversion of wind energy into mechanical energy by the turbine:

$$P = 0.5\times C_p \times \pi \times R^2 \times \rho \times V_w^3 \tag{5}$$

where ρ is the air density; C_p is the maximum wind energy capture factor; R is the blade radius; V_w is the wind speed.

$$P_{\mathrm{wp}}(V) = \begin{cases} 0 & V < V_{\mathrm{cutin}}, V > V_{\mathrm{cutout}} \\ P_{\mathrm{rated}}\dfrac{V - V_{\mathrm{cutin}}}{V_{\mathrm{rated}} - V_{\mathrm{cutin}}} & V_{\mathrm{cutin}} \le V < V_{\mathrm{rated}} \\ P_{\mathrm{rated}} & V_{\mathrm{rated}} \le V \le V_{\mathrm{cutout}} \end{cases} \tag{6}$$

where V_{cutin} and V_{cutout} are the cut-in and cut-out wind speeds for fan operation, V_{rated} is the rated wind speed, and P_{rated} is the rated output of the fan.

4) Photovoltaic power generation

In this paper, when studying load frequency control, PV power is also considered as an uncontrollable power source, i.e., like wind power, it is treated as a random negative load, which requires determining the output power of PV power. The formula for calculating the output power of PV power plant P_{el} is as follows.

$$P_{el} = \eta_q \eta_T \eta_i \eta_n \eta_l P_{AZ} = \eta_z P_{AZ} \tag{7}$$

Where $\eta_q \eta_T \eta_i \eta_n \eta_l$ is a constant and P_{AZ} is the installed capacity.

5) Fuel Cell

In this paper, a SOFC power generation system is studied to generate electrical energy through the electrochemical reaction between hydrogen and oxygen with the following output voltage.

$$V_{\mathrm{out}} = E_0 + \frac{R_s T}{2F}\ln\left\{\frac{\left(P_{H_2}P_{o_2}^{0.5}P_{co_2}\right)}{P_{H_2o}P_{co_2}}\right\} \tag{8}$$

$$P_{\mathrm{out}}(V) = I_{st}V_{out} \tag{9}$$

where P_{H2} , P_{O2} and P_{CO2} the pressure of the gas.

2.2 ALFC Architecture

The secondary frequency regulation is performed by the generator's frequency regulator, which can eliminate the fixed error by changing the set value of the generator unit output power, and can restore the frequency to the steady-state value. Although the secondary non-differential frequency regulation has been able to directly realize the non-differential regulation, but due to the quantization of frequency regulation and conversion frequency regulation process of specific time conditions, the response speed of the secondary non-differential frequency regulation may be much lower than the primary frequency regulation, so in the frequency regulation process for the unintelligible disturbance situation a secondary frequency regulation to cooperate with each other so as to eliminate the frequency deviation.

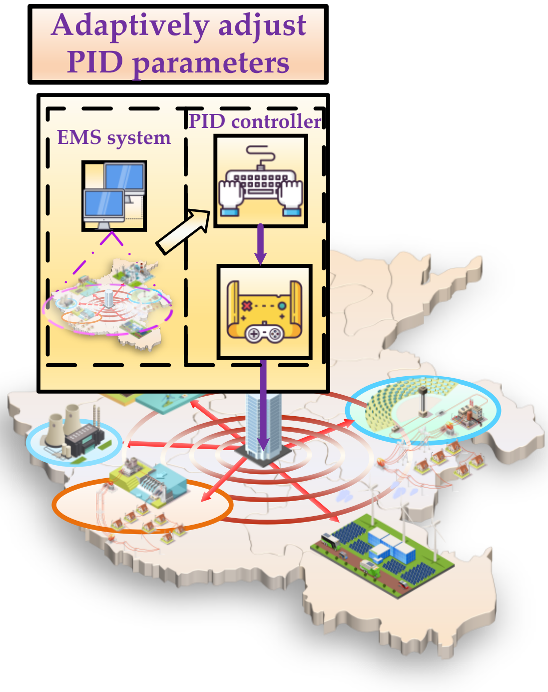

Fig. 1 ALFC model

Layered control technology refers to the entire control system is divided into two layers, or multi-layer control, generally in the uppermost layer are equipped with a central controller, used to send action instructions to the lower controller, the lower controller can consider the instructions to optimize, or the coordination of control between units. This top-down control is realized through the communication network, so a perfect and reliable communication network is an important guarantee for the normal operation of the control system, and once the communication fails, the micro network will not be able to continue to operate. In this paper, we study the topmost central controller and propose an adaptive load frequency control (ALFC) method. The method uses an adaptive proportional-integral-derivative controller to achieve adaptive control by adaptively adjusting the parameters of the controller to output a total regulation command, which is then fixed by the distributor.

2.3 Objective functions and constraints

ALFC controller in the load frequency control strategy needs to consider not only reducing the frequency deviation but also reducing the generation cost The ALFC controller in the load frequency control strategy needs to consider not only reducing the frequency deviation but also reducing the generation cost of the unit which are shown as follows.

$$\min \sum_{t=1}^{T}|\Delta f| + \sum_{t=1}^{T}\sum_{i=1}^{n}\left(\alpha_i \Delta P_{Gi}^2 + \beta_i \Delta P_{Gi} + \gamma_i\right) \tag{10}$$

3. DEEP ACTOR CRITIC

The agent is the decision maker that keeps learning (i.e., reinforcement learning algorithm), and all the contents other than the Intelligent Agent are called the environment. In contrast to the process control process, the Agent will be the controller, and the control system other than this will form the environment. The above decision process is formulated as Markov Decision Process (MDP) [1].

The mathematical description of the Markov decision process is shown in the five-tuple in equation (11):

$$M = \left(s, a, p\left(s', r \mid s, a\right), \gamma, R\right) \tag{11}$$

where s is the state space of the intelligence in reinforcement learning, which is discrete in some problems, here is the grid state [1];

a is the action space of the intelligence in reinforcement learning, in this case the parameters of the PID controller;

R is the direct reward (or desired direct reward) received after state transfer in reinforcement learning, which usually corresponds to the objective function;

γ is discount rate. It is used to express the uncertainty of the future, i.e., the discount rate reflects the importance of long-term returns, thus avoiding the situation of infinite long-term returns.

In the DAC algorithm, the Actor network, i.e., the policy network, is used to generate policies, and the Critic network, i.e., the value network, is used to fit the value function and evaluate the policies generated by the Actor network. Obviously DAC algorithm belongs to the off-policy type, because the Critic network iterates continuously while optimizing the Actor network, but the action of policy generation generated by the Actor network does not completely depend on the Critic network [1].

The DAC algorithm introduces critic target network to calculate the objective function, where the expression of the objective value function is as follows [1]:

$$J = r + \gamma Q'\left(s', a', w'\right) \tag{12}$$

DAC algorithm is a deterministic strategy as the name implies, deterministic strategy is efficient, but the disadvantage is that it is not exploratory enough, for this reason, external noise interference is often added in the action selection [1]:

$$a = \mu\left(s \mid \theta_{\mu}\right) + N \tag{13}$$

It choose Gaussian noise as the interference, Gaussian noise is a class of noise whose probability density function obeys Gaussian distribution, also known as normal distribution, also known as normal distribution.

$$N\left(x; \mu, \sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{14}$$

The introduction of Gaussian noise makes the original action selection produce a random fluctuation around the mean value, which increases the generalization ability of the model, improves the exploration efficiency of the control task, and brings us a more continuous exploration and a richer set of actions. In fact, we can add Gaussian noise to DAC algorithm because DAC algorithm is an off-policy algorithm, we have different strategies in updating the gradient and action selection, we update the gradient normally, but apply noise in action selection, so that it does not affect the iteration of the algorithm, but the action after receiving noise enriches our In practice, we often take the approach of gradually reducing the value of Gaussian noise, in order to improve the exploration and generalization ability of the algorithm in the early stage, and to ensure the stability of the algorithm in the later stage of training [1].

4. CASE STUDIES

To verify the function of the algorithm, it simulate and test the ALFC based on DAC algorithm and introduce the ALFC ' based on DDPG algorithm [3], PSO-fuzzy -FOPID [2], GA-fuzzy-PID [3], TS-fuzzy-PID [7], and PSO-PID [13] as comparisons. The performance of the proposed method is tested in the Zhuzhou islanded microgrid of the China Southern Grid, the total simulation time is 6 hours. includes step disturbance, wind power and photovoltaic disturbance [3].

Tab. 2 Statistical results

Control algorithms	Average frequency deviation (HZ)	Generation cost ($)
	$\lvert\Delta f\rvert_{avg}$	C^{total}
DAC	0.004277	2000.8
DDPG	0.005514	2001.3
PSO-fuzzy-FOPID	0.005243	2000.9
GA-fuzzy-PID	0.005181	2000.9
TS-fuzzy-PID	0.005520	2001.1
PSO-PID	0.004332	2001.3

As shown by Table 1, the Average frequency deviation of DAC decreases by 1.28%- 29.05% and the total generation cost decreases by 0.0049%- 0.024% compared to other methods. This is because DAC uses Gaussian noise and multi-critic technique to increase the performance of the algorithm, which results in improved ALFC performance and thus adaptive control. the DDPG algorithm does not use Gaussian noise for performance improvement and does not use multi-critic technique to prevent the algorithm from falling into local optimum. Therefore, its obtained ALFC performance is lower. The adaptive control method based on fuzzy rules is too simple due to the rules and thus difficult to adapt to the complex islanded microgrid environment. pso-pi optimizes the parameters of the PID control algorithm, but is unable to control the parameters of the PID algorithm in real time, which in turn leads to low performance of the algorithm.

5. CONCLUSIONS

In this paper, an adaptive load frequency control (ALFC) method is proposed. The method uses an adaptive proportional-integral-derivative controller to achieve adaptive control by adaptively adjusting the parameters of the controller to output the regulation command of the unit. In addition, to achieve adaptive regulation of the control parameters, a deep actor-critic (DAC) algorithm is proposed in this paper, which introduces multiple critics and Gaussian noise exploration techniques to enhance the quality of the ALFC strategy. The performance of the proposed method is tested in Zhuzhou isolated microgrid of Southern Power Grid, and the Average frequency deviation of DAC is reduced by 1.28%- 29.05% compared with PSO-fuzzy-FOPID, GA-fuzzy-PID, TS-fuzzy-PID, and PSO-PID algorithms The average frequency deviation of DAC decreases by 1.28%- 29.05% and the total generation cost decreases by 0.0049%- 0.024%.

REFERENCE

[1] Li J, Yu T, Zhang X, Li F, Lin D, Zhu H. Efficient experience replay based deep deterministic policy gradient for AGC dispatch in integrated energy system. Appl Energy 2021;285:116386.

[2] Polimeni S, Meraldi L, Moretti L, Leva S, Manzolini G. Development and experimental validation of hierarchical energy management system based on stochastic model predictive control for Off-grid Microgrids. Adv Appl Energy 2021;2:100028.

[3] Li J, Yang S, Yu T. Data-driven cooperative load frequency control method for microgrids using effective exploration - distributed multi - agent deep

reinforcement learning. IET Renew Power Gener 2022;16:655-70.
[4] Tang F, Niu B, Zong G, Zhao X, Xu N. Periodic event-triggered adaptive tracking control design for nonlinear discrete-time systems via reinforcement learning. Neural Netw 2022;154:43-55.
[5] Xu Y, Liu W. Novel multiagent based load restoration algorithm for microgrids. IEEE Trans Smart Grid 2011;2:152-61.
[6] Zhang W, Xu Y, Liu W, Ferrese F, Liu L. Fully distributed coordination of multiple DFIGs in a microgrid for load sharing. IEEE Trans Smart Grid 2013;4:806-15.
[7] Xu Y, Zhang W, Liu W, Wang X, Ferrese F, Zang C, et al. Distributed subgradient-based coordination of multiple renewable generators in a microgrid. IEEE Trans Power Syst 2013;29:23-33.
[8] Kundur P, Balu NJ, Lauby MG. Power system stability and control. New York: Mcgraw-Hill; 1994.
[9] Hain Y, Kulessky R, Nudelman G. Identification-based power unit model for load-frequency control purposes. IEEE Trans Power Syst 2000;15:1313-21.
[10] Sahu RK, Panda S, Sekhar GTC. A novel hybrid PSO-PS optimized fuzzy PI controller for AGC in multi area interconnected power systems. Int J Electr Power Energy Syst 2015;64:880-93.
[11] Saikia LC, Nanda J, Mishra S. Performance comparison of several classical controllers in AGC for multi-area interconnected thermal system. Int J Electr Power Energy Syst 2011;33:394-401.
[12] Liu XJ, Zhang JW. CPS compliant fuzzy neural network load frequency control. 2009 American Control Conference. St. Louis, MO, USA: IEEE; 2009. p. 2755-60.
[13] Kocaarslan I, Çam E. Fuzzy logic controller in interconnected electrical power systems for load-frequency control. Int J Electr Power Energy Syst 2005;27:542-9.
[14] Li H, Wang X, Xiao J. Adaptive event-triggered load frequency control for interconnected microgrids by observer-based sliding mode control. IEEE Access 2019;7:68271-80.
[15] Li J, Yu T, Yang B. Coordinated control of gas supply system in PEMFC based on multi-agent deep reinforcement learning. Int J Hydrog Energy 2021;46:33899-914.
[16] Li J, Yu T. A novel data-driven controller for solid oxide fuel cell via deep reinforcement learning. J Clean Prod 2021;321:128929.
[17] Li J, Yu T, Yang B. A data-driven output voltage control of solid oxide fuel cell using multi-agent deep reinforcement learning. Appl Energy 2021;304:117541.
[18] Wu H, Pratt A, Munankarmi P, Lunacek M, Balamurugan SP, Liu X, et al. Impact of model predictive control-enabled home energy management on large-scale distribution systems with photovoltaics. Adv Appl Energy 2022;6:100094..