Coordinated optimal dispatch of composite energy storage microgrid based on double deep *Q*-network

Zheyong Piao

Baicheng Power Supply Company,
State Grid Jilin Electric Power State,
Baicheng, Beijing, China
Email: pzy8100@163.com

Tianyu Li*

School of Electrical Engineering,
Changchun Institute of Technology,
Changchun, Jilin, China
Email: wy315881@163.com
*Corresponding author

Benfa Zhang

State Grid Songyuan Power Supply Company,
Songyuan, Jilin, China
Email: 525119761@qq.com

Lei Kou

Institute of Oceanographic Instrumentation,
Qilu University of Technology (Shandong Academy of Sciences),
Qingdao, Jinan, Shandong, China
Email: koulei1991@hotmail.com

Abstract: In order to optimise the coordinated control of micro-grid complex energy storage including photovoltaic and wind power, improve the absorption ability of distributed energy generation and reduce the cost, this paper proposes a Double Deep *Q*-Network reinforcement learning algorithm to train agents to interact with the microgrid environment and learn the optimal scheduling control mechanism. The agent trains the input state and outputs an optimal action to drive the agent's behaviour, including environment perception, action perception and task coordination. It then successfully completes the given task in the complex decision environment. This method can realise multi-objective control for different times, weather conditions and seasons and flexibly process energy storage, hydrogen storage and load energy to achieve coordinated distribution. First of all, a composite energy storage microgrid system model connected to the main power grid is constructed, and deep reinforcement learning activities, state space, reward mechanism and other links are designed. Secondly, in the aspect of learning distributed generation data, a combination of training set and test set of data is proposed for model learning and training. Finally, the optimisation scheduling results of reinforcement learning are analysed for different scenarios of composite energy storage microgrid.

Keywords: microgrid; deep reinforcement learning; composite energy storage; optimise scheduling.

Reference to this paper should be made as follows: Piao, Z., Li, T., Zhang, B. and Kou, L. (2024) 'Coordinated optimal dispatch of composite energy storage microgrid based on double deep *Q*-network', *Int. J. Wireless and Mobile Computing*, Vol. 26, No. 1, pp.92–98.

Biographical notes: Zheyong Piao received his MSc degree in Electrical Engineering and Automation from Northeast Electric Power University in 2007. He is a Senior Engineer at Baishan Power Supply Company of State Grid in Baishan City. His current research interests include electricity market and production plan management work.

Tianyu Li received his MSc degree in Computer Science and Technology from Jilin University in 2010. He is an Associate Professor at Changchun Institute of Engineering. His current research interests include smart grids and new energy optimisation control technologies.

Benfa Zhang received his MSc degree in Computer Science and Technology from Northeast Electric Power University in 2016. He is a Research Associate at State Grid Songyuan Power Supply Company, Songyuan, China. His current research interests include artificial intelligence, machine learning, microgrid and wind farms.

Lei Kou received his MSc degree in Computer Science and Technology from Northeast Electric Power University in 2017 and PhD degree in Electrical Engineering from Northeast Electric Power University in 2020. He is a Research Associate at Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences). His current research interests include data encryption, artificial intelligence, machine learning, smart grid and offshore wind farms.

1 Introduction

As an emerging form of energy management, microgrid has developed rapidly in recent years, and it is the key to provide a reasonable energy dispatching strategy for the safe, stable and economical operation of the microgrid system. At the same time, microgrid is a small power generation and distribution system composed of distributed energy, energy storage device, energy conversion device, load and protection device, etc., which is rapidly developed to solve the problem of distributed energy integration such as photovoltaic power generation and wind power generation (Yang et al., 2014; Kou et al., 2023). With the large-scale application of wind power and photovoltaic power generation, microgrids have become a hot spot for research in various countries. Relying on key technologies such as operation control and energy management, microgrids can operate both on-grid and off-grid islands (Sang et al., 2022). In micro-energy grids, the instability of renewable energy, the multi-directional nature of energy flow, and the dispatch demand of multiple energy sources at different time and space levels make multi-energy flow coupling not only a characteristic of integrated energy systems, but also one of the major challenges faced by its energy management (Zhang et al., 2016). Seeking effective optimisation strategies to achieve high-quality and efficient energy management of the system is very important to improve the performance of the system. The research on microgrid includes many aspects, such as the research on microgrid architecture, power electronics technology, investment and operating costs, security and operation and maintenance, energy dispatching control strategy, etc. Among them, energy storage control strategy has been widely concerned as the research focus of microgrid energy dispatch (Liang et al., 2020). The traditional microgrid form is single, which cannot meet the flexible energy scheduling between the complex demand side and the microgrid, so it is necessary to configure an energy storage system. It is difficult for a single energy storage to meet the roles of power and energy at the same time, and the reasonable way is to use composite energy storage; Composite energy storage is a complex problem, difficult to use accurate data models and numerical calculation methods, requiring data-driven artificial intelligence methods (Foruzan et al., 2018; Kou et al., 2022; He et al., 2017; Wang et al., 2022).

Existing research proposes traditional techniques such as mixed integer linear programming and heuristic algorithms to optimise energy management in microgrids. However, there are many shortcomings in the related methods, and linear programming cannot deal with the problem of many decision variables in a limited calculation time. Although the heuristic algorithm can deal with the problem of many decision-making variables, the optimal solution obtained by it is the approximate optimal solution, and it is easier to fall into the local optimal solution and the gap between the approximate optimal solution and the actual optimal solution cannot be measured and estimated. With the rise of artificial intelligence, technologies based on intelligent learning have made significant progress in decision-making. Reinforcement Learning (RL) 'is a machine learning algorithm that makes optimal temporal decisions in an uncertain environment. Reinforcement learning involves the decision maker (agent) learning how to take actions in a specific state through continuous interaction with the environment in order to maximise cumulative returns (Sutton et al., 2018; Mason and Grijalva, 2019; Foruzan et al., 2018).

This paper aims at the problem of complex energy storage scheduling based on clean energy microgrid model. Energy storage system can effectively solve the intermittent problem of distributed energy output. Therefore, this paper adopts the intelligent microgrid group with wind/light/storage/network co-output power generation mode. This paper mainly studies the influence of different seasons and weather on distributed power generation, but considering the complexity of composite power generation model, this paper mainly considers the establishment of photovoltaic power generation model. At the same time, real-time optimisation control of charging and discharging state of battery and hydrogen storage device can improve the utilisation rate of distributed renewable energy. In addition, the microgrid model has both short-term and long-term characteristics: long-term storage involves electrolysis of water to produce hydrogen for energy storage, while short-term storage deals with insufficient or excess energy through discharge or charging of battery banks respectively. Efficient operation of storage devices in microgrids characterised by photovoltaic panels with storage capacity. At the same time, for a large number of original photovoltaic power generation data, the original data is divided into two data sets for model training, which are respectively effective data and test data. The effective data is used for the training of the microgrid model, while the test data is used to verify the training parameters of the model. Meanwhile, the execution of the strategy on the invisible time series is evaluated periodically to ensure that the agent does not over adapt to the limited training data. Firstly, the paper introduces the microgrid composite energy storage model. Secondly, it focuses on the framework and algorithm flow of deep reinforcement learning. Finally, an example analysis proves the effectiveness of the algorithm.

2 Energy storage microgrid model and scenario

2.1 Energy storage microgrid structure

This paper considers the energy dispatch problem of composite energy storage microgrid. Microgrids can exchange energy with the main grid at electricity market prices. At the same time, the time-of-use electricity price is adopted. The microgrid can purchase external power through the energy management system at low power, and when the energy is sufficient, the energy of the battery and hydrogen storage system can be stored, without considering the loss in the process. The structure is shown in Figure 1. The DC bus side includes photovoltaic, wind power, energy storage system, electrolytic cell and fuel cell, while the AC bus side includes AC composite and generator.

2.2 Energy storage microgrid component model

In the microgrid, the AC master side and the DC master side are connected through AC/DC converter. The main components include: This part does not consider the influence of wind power:

1) *Photovoltaic power generation*: The power supply of the microgrid system comes from photovoltaic power generation and energy storage system, and the photovoltaic output model (Kuznetsova et al., 2013) is

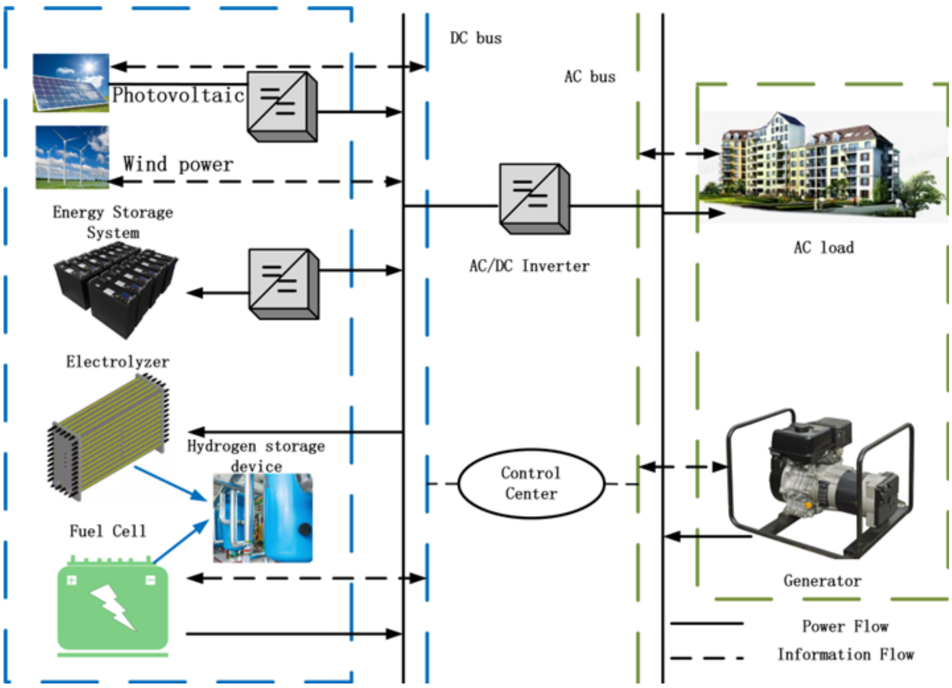
$$P_{PV}(t) = R(t)S\eta_{PV} \tag{1}$$

where: $R(t)$ indicates the light intensity at t time, in W/m^2; S is the area of the target photovoltaic panel. The ride of the above two indicates the radiation power received by the solar panel; η_{PV} is the electric energy conversion efficiency of photovoltaic panels.

2) *Energy storage system*: The energy storage system acts as a short-term reserve, and its charge and discharge response signal is sent by the battery. In a microgrid, the energy storage system can maintain the balance of supply and demand in the microgrid. $S_{bat}(t)$ represents the real-time energy of the battery, and $S_{bat}(t-1)$ is the energy before charging and discharging. P_{bat}^{cha} and P_{bat}^{dis} are charge and discharge power, η_{bat} is charge and discharge efficiency, then:

$$S_{bat}(t) = \begin{cases} S_{bat}(t-1) - \int \frac{P_{bat}^{dis}}{\eta_{bat}} dt \\ S_{bat}(t-1) + \eta_b \int P_{bat}^{cha} dt \end{cases} \tag{2}$$

Figure 1 Physical structure of microgrid (see online version for colours)



3) *Hydrogen storage system*: As a long-term energy storage device, the hydrogen storage system has low charging and discharging efficiency and small peak power, but it can store energy for a long time through electrolysis, which is mainly used to balance the energy imbalance between seasons. Can be expressed (Leo et al., 2014 and Leo et al., 2014) as

$$S_{hydro}(t) = \begin{cases} S_{hydro}(t-1) - \int \frac{P_{hydro}^{\text{dis}}}{\eta_{hydro}} dt \\ S_{hydro}(t-1) + \eta_{hydro} \int P_{hydro}^{cha} dt \end{cases} \quad (3)$$

In the formula: P_{hydro}^{cha} and P_{hydro}^{dis} are the power of electrolysis and release of the hydrogen storage system, $S_{hydro}(t)$ represents the real-time energy of the hydrogen storage equipment, $S_{hydro}(t-1)$ represents the capacity of the equipment before electrolysis or release, and η_{hydro} represents the electrolysis or release efficiency.

4) *Load*: The load as a whole is a general term as an energy consumption part, and for the microgrid system with a small total load in this article, the range of fluctuations by random factors is large and un-adjustable. In this paper, the load curve is fed into the microgrid system as a fixed energy consumption.

2.3 Energy storage microgrid constraints

Microgrid energy optimisation operation constraints include power balance constraints and related equipment operation constraints.

1 *Power balance constraints*: At any time in the microgrid, the real-time balance of power supply and demand must be roughly ensured to avoid the problem of system power abandonment or power loss, which is expressed (Mbuwir et al., 2017) as:

$$\delta_t - d_t - \Delta p_t^{bat} - \Delta p_t^{hydro} = 0 \quad (4)$$

$$d_t = P_{load}^t - P_{pv}^t \quad (5)$$

In the formula: The positive and negative of δ_t indicates that insufficient power supply and overpower supply are the phenomenon of power loss and abandonment, and P_{load}^t, P_{pv}^t represent the load and the real-time power of photovoltaics. $\Delta p_t^{bat}, \Delta p_t^{hydro}$ represent the net charging power of the energy storage system and hydrogen storage system at t time, respectively, which can be expressed (Zhang and Hredzak, 2019) as:

$$\Delta p_t^{bat} = P_{bat}^{cha} - P_{bat}^{dis} \quad (6)$$

$$\Delta p_t^{hydro} = P_{hydro}^{cha} - P_{hydro}^{dis} \quad (7)$$

2 *Capacity constraints*: The energy storage and hydrogen storage systems meet the capacity constraints and state-of-charge constraints, where $B_{soc}(t)$ represents the charged charge, namely:

$$S_{bat,\min} \le S_{bat}(t) \le S_{bat,\max} \quad (8)$$

$$S_{hydro,\min} \le S_{hydro}(t) \le S_{hydro,\max} \quad (9)$$

$$0 \le B_{soc}(t) \le 1 \quad (10)$$

3 *Power constraints*: For the properties of the energy storage system or the hydrogen storage system itself, the charging and discharging of the two must have certain restrictions (Duan et al., 2019):

$$0 \le P_{dis/cha,t}{}^{bat} \le P_{dis/cha,\max}{}^{bat} \quad (11)$$

$$0 \le P_{dis,cha,t}{}^{hydro} \le P_{dis/cha,\max}{}^{hydro} \quad (12)$$

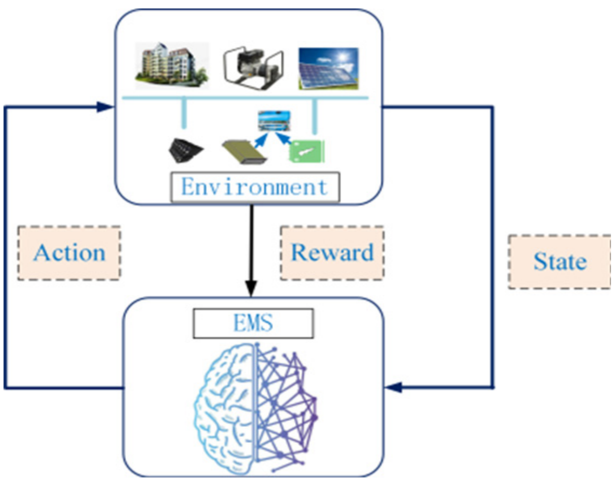
Formula: $P_{dis/cha,t}{}^{bat}$ represents the energy storage system t time discharge and charging power, $P_{dis,cha,t}{}^{hydro}$ represents the hydrogen storage system t time release, electrolysis power.

3 Microgrid learning for deep neural networks

3.1 Deep reinforcement learning algorithm

Reinforcement learning is to maximise the expected return as the goal, through the continuous interaction of the agent (Agent) and the environment (Environment), to obtain the mapping relationship between the state variable of the environment and the action variable of the agent, and provide the agent with an optimised action strategy (Policy). This is shown in Figure 2. Essentially, the agents interaction with the environment is a Markov Decision Process (MDP). MDP is generally defined by a quadruple (S, A, R, P), where:

Figure 2 Markov decision process (see online version for colours)



1) According to the time series information, the agent can observe that the dynamic information of the microgrid is that the load consumption in different seasons is different from the daily load consumption. The time sequence is

defined according to the microgrid as follows (Mocanu et al., 2019; Wang et al., 2020):

$$\forall S_m \in S, S_m = S_m^d S_m^h \tag{13}$$

Formula: $S_m^d \in \{180, 360\}, S_m^h \in \{0, 1, ..., 23\}$; S_m^d Spring/ Summer 180d time series; S_m^h is a time series of 24h in a day.

2) The Action space represents the set of discrete actions a that the agent can perform, including the real-time action a_{bat} and the hydrogen storage action a_{hydro} of the energy storage system, and each specific action includes three states of charge, discharge, electrolysis and release, and no operation, as follows (Wang et al., 2023):

$$A : \{a_{bat}, a_{hydro}\} \tag{14}$$

3) Reward is a timely reward function that immediately evaluates the merits of actions and environments. The function is subtracted for curtailment and loss of power. At the same time, negative rewards will be received whenever short-term or long-term hydrogen storage is empty and cannot meet the energy demand.

4 Example verification and analysis

4.1 Basic data

While verifying the effectiveness of the reinforcement learning algorithm DDQN for the optimal dispatch of composite energy storage microgrid, it is necessary to list the parameters of each component of the microgrid. It is stipulated that 100 hours is used as the analysis period of coordinated dispatch of microgrid energy storage. The environment sets the main microgrid parameters and the hyper-parameters of the algorithm.

4.2 Analysis of results

In this paper, a microgrid system with 300kW photovoltaic system as the main wind power as the auxiliary system is adopted in a certain area, and the influence of wind power is ignored in this paper. At the same time, the power generation is greatly affected by solar radiation, season and time. On the other hand, due to the lack of energy supply from external power grid, the load fluctuation is more obvious, and the simulation of daily load curve is more significant. In Figure 3, the picture show the one-year load and photovoltaic power generation of the micro-grid system in this region. The data resolution is set to 1 h. The algorithm is implemented by Python 3.8 programming.

4.2.1 Energy dispatch analysis

In Figure 4, it is not difficult to see that due to the influence of time and season in the region, the photovoltaic power generation curve fluctuates significantly. The output power of photovoltaic power generation system is affected by solar radiation intensity, sky shielding, cloud cover and other weather conditions. Cloudy days, cloudy skies or changes in the position of the sun can lead to fluctuations in output power, and this uncertainty makes the output of photovoltaic power generation systems unstable.

Figure 3 Photovoltaic power generation (see online version for colours)

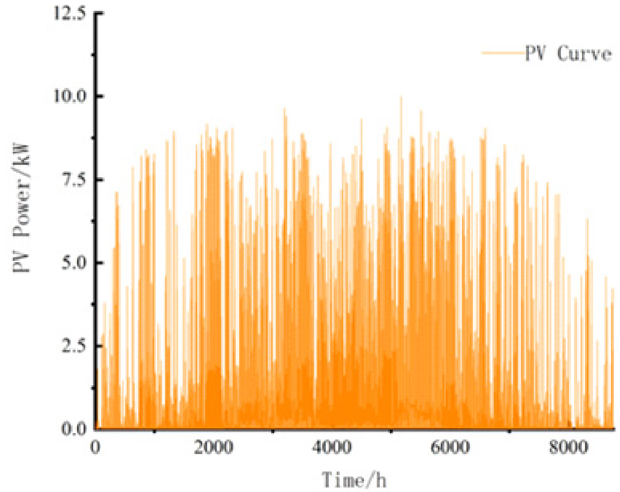


Figure 4 Reward value curve (see online version for colours)

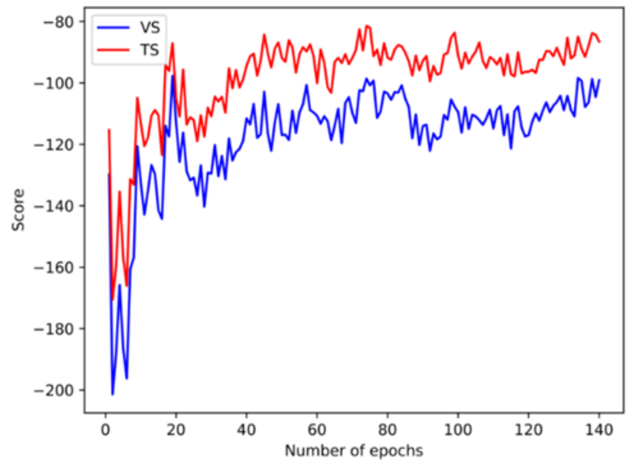
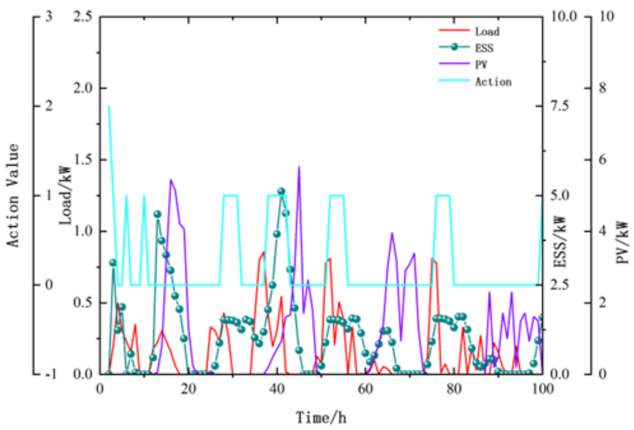


Figure 4 shows the convergence results of microgrid model training with different data sets. VS stands for effective data convergence and TS stands for test data convergence. After the 20th training set, the models showed convergence trend and the overall convergence state was stable between 40 and 140 episodes. Among them, during the training process of the test set and the effective set, the final reward value of the agent feedback does not coincide, which indicates that the training of the model has achieved the purpose of avoiding over-fitting, and the fluctuation trend of the reward value curve of the two is generally consistent, which basically achieves the accurate training results of the model.

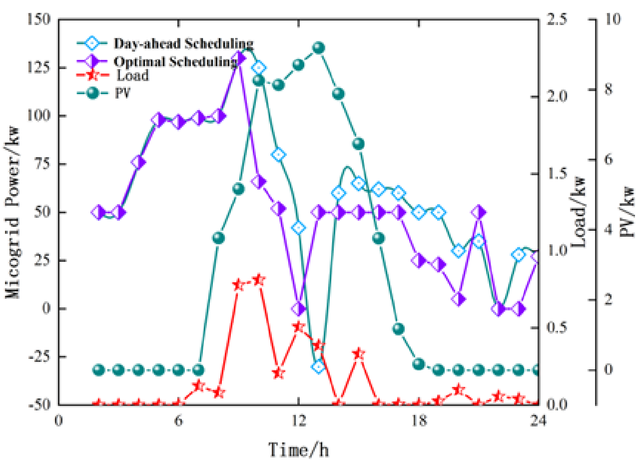
In Figure 5, when the power generation of the photovoltaic system cannot fully meet the load consumption, the hydrogen storage system begins to release the maximum rate of energy and the entire energy storage system begins to discharge, thereby reducing the loss of load. Further analysis shows that when the intelligent body measurement increases the effective information, the agent selects more effective charging and discharging actions, so as to achieve the goal of obtaining the maximum operating income of the microgrid and improve the performance of the dispatching strategy. Therefore, the decision of deep reinforcement learning relies on the current state and score, real-time decision-making can be realised, and when the algorithm is fully trained, it can get the action that should be carried out according to any current state, and this multi-objective, multi-dimensional strategy learning is unattainable by most traditional algorithms.

Figure 5 Typical winter strategies (see online version for colours)



In comparison Figure 6, the photovoltaic power around 00:00–07:00 is close to zero for the entire stage, so the energy storage charge in Figure 10 is at a low level and decreases, and the SOC drops to 0.02. Because energy storage is used as a power supply unit to provide energy for the microgrid at this time, the entire power grid is in a state of underpower and needs to be purchased from the external grid, and the grid energy drops to less than 25kW. Moreover, the electricity market price during this period is in the normal and trough hours. Meet low-price power purchase requirements.

Figure 6 Microgrid power optimisation dispatch and day-ahead dispatch (see online version for colours)



5 Conclusions

This paper introduces a deep reinforcement learning architecture to solve the problem of coordinated control of energy storage microgrids in a random environment. The conclusions are as follows:

1) In the evaluation of the coordinated control operation state of the energy storage microgrid, on the one hand, the weight value of the index is evaluated by the deep convolutional neuron network to adapt to the coordinated control operation state evaluation of the energy storage microgrid at different times and seasons. On the other hand, during the training process of the proposed model, it is not necessary to be familiar with the relationship between the control target and the state information observed by the agent, but to borrow the deep convolutional neural network to approximate the *Q* function value to determine the connection between the target variable and the state information, so it is more effective for the training control plane between the agent and the model.
2) In terms of raw data mining, this paper divides the data set into two types: training data and validation data. The validation data is validated for model validation, and the implementation of the policy on the invisible time series is regularly evaluated to ensure that the agent does not over-adapt to the limited training data.

Acknowledgements

This work was supported in part by the Science and Technology Project of Baicheng Power Supply Company of State Grid Jilin Electric Power Co (2022-13), Ltd., Shandong Provincial Natural Science Foundation (ZR2023QF036), Basic Research Projects of Science, Education and Industry Integration Pilot Project of Qilu University of Technology (Shandong Academy of Sciences) (2023PX081), Talent Research Project of Qilu University of Technology (Shandong Academy of Sciences) (2023RCKY054).

References

Duan, J.J., Yi, Z.H. and Shi, D. (2019) 'Reinforcement-learning-based optimal control of hybrid energy storage systems in hybrid ac-dc microgrids', *IEEE Transactions on Industrial Informatics*, Vol. 15, No. 9, pp.5355–5364.

Fathima, A.H. and Palanisamy, K. (2015) 'Optimization in microgrids with hybrid energy systems-a review', *Renewable and Sustainable Energy Reviews*, Vol. 45, pp.431–446.

Foruzan, E., Soh, L.K. and Asgarpoor, S. (2018) 'Reinforcement learning approach for optimal distributed energy management in a microgrid', *IEEE Transactions on Power Systems*, Vol. 33, No. 5, pp.5749–5758.

Foruzan, E., Soh, L.K. and Asgarpoor, S. (2018) 'Reinforcement learning approach for optimal distributed energy management in a microgrid', *IEEE Transactions on Power Systems*, Vol. 33, No. 5, pp.5749–5758.

He, X., Ai, Q. and Qiu, R.C. et al. (2017) 'A big data architecture design for smart grids based on random matrix theory', *IEEE Transactions on Smart Grid*, Vol. 8, No. 2, pp.674–686.

Kou, L., Li, Y., Zhang, F., Gong, X., Hu, Y., Yuan, Q. and Ke, W. (2022) 'Review on monitoring, operation and maintenance of smart offshore wind farms', *Sensors*, Vol. 22. Doi: 10.3390/s22082822.

Kou, L., Wu, J. and Zhang, F. et al. (2023) 'Image encryption for offshore wind power based on 2D-LCLM and Zhou Yi Eight Trigrams', *International Journal of Bio-Inspired Computation*. Doi: 10.1504/IJBIC.2023.10057325.

Kuznetsova, E., Li, Y.F. and Ruiz, C. et al. (2013) 'Reinforcement learning for microgrid energy management', *Energy*, Vol. 59, pp.133–146.

Leo, R., Milton, R.S. and Sibi, S. (2014) 'Reinforcement learning for optimal energy management of a solar microgrid', *Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS)*, Trivandrum, India, IEEE, pp.183–188.

Liang, H., Li, H.G. and Zhang, H.Y. et al. (2020) 'Control strategy of microgrid energy storage system based on deep reinforcement learning', *Power System Technology*, Vol. 45, No. 10, pp.3869–3877. Doi: 10.13335/J.1000-3673.pst.2020.1754.

Mason, K. and Grijalva, S. (2019) 'A review of reinforcement learning for autonomous building energy management', Comput. Electr. Eng., Vol. 78, pp.300–312.

Mbuwir, B., Ruelens, F., Spiessens, F. and Deconinck, G. (2017) 'Battery energy management in a microgrid using batch reinforcement learning', *Energies*, Vol 10, No. 11.

Mocanu, E., Mocanu, D.C. and Nguyen, P.H. et al. (2019) 'On-line building energy optimization using deep reinforcement learning', *IEEE Transactions on Smart Grid*, Vol. 10, No. 4, pp.3698–3708.

Sang, J.S., Sun, H.B. and Kou, L. (2022) 'Deep reinforcement learning microgrid optimization strategy considering priority flexible demand side', *Sensors*, Vol. 22, No. 6. Doi: 10.3390/s22062256.

Sutton, R.S., Barto, A.G. and Bach, F. (2018) *Reinforcement Learning: An Introduction*, 2nd ed., The MIT Press, Cambridge, MA, USA.

Wang, H., Yang, D.S. and Zhou, B.W. et al. (2020) 'Fault diagnosis of multi-terminal HVDC transmission line based on parallel convolutional neural network', *Automation of Electric Power Systems*, Vol. 44, No. 12, pp.84–92.

Wang, J., Zhang, X.H., Zhang, F.F., Wan, J.H., Kou, L. and Ke, W.D. (2022) 'Review on evolution of intelligent algorithms for transformer condition assessment', *Frontiers in Energy Research*, Vol. 10. Doi: 10.3389/fenrg.2022.904109.

Wang, Y., Mao, M. and Chang, L. et al. (2023) 'Intelligent voltage control method in active distribution networks based on averaged weighted double deep q-network algorithm', *Journal of Modern Power Systems and Clean Energy*, Vol. 11, No. 1, pp.132–143.

Yang, X., Su, J. and Lü, Z. et al. (2014) 'Overview on micro-grid technology', *Proceedings of the CSEE*, Vol. 34, No. 1, pp.57–70.

Zhang, R.F. and Hredzak, B. (2019) 'Nonlinear sliding mode and distributed control of battery energy storage and photovoltaic systems in ac microgrids with communication delays', *IEEE Transactions on Industrial Informatics*, Vol. 15, No. 9, pp.5149–5160.

Zhang, T., Zhang, F.X. and Zhang, Y. (2016) 'Study on energy management system of energy internet', *Power System Technology*, Vol. 40, No. 1, pp.146–155.