International Journal of Multiphysics
Volume 18, No. 2, 2024
ISSN: 1750-9548

Research on Load Intelligent Transfer Strategy of Urban Power Grid Based on Improved DQN

Yunfei Chu[1], Jing Wang[1], Qimeng Cui[2,*], Yong Wang[1], Boqiang Li[1]

[1]Economic and Technical Research Institute, State Grid Jilin Electric Power Co., Ltd., Changchun 130022, Jilin, China

[2]School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China

*Corresponding Author.

Abstract

With the expansion of the scale of the urban power grid, its structure is becoming more and more complex. After the occurrence of N-1 failure, the operation risk is likely to increase, resulting in a large area of power failure at N-1-1. Therefore, this paper proposes a load transfer strategy of urban power grid based on Double Deep Q-Network to generate load transfer scheme of power grid intelligently. Firstly, the load transfer problem of urban power grid is modeled as a sequential decision problem which is easy to be learned and trained by agents, and the intelligent load transfer decision model is constructed. Secondly, a learning framework of agent load transfer knowledge is proposed by using a large amount of interaction information between agents and line switches in the simulation environment to continuously accumulate power grid operation knowledge. Finally, in order to improve the effectiveness and generalization of the algorithm, a strategy of pre-action-change exploration value selection is added. The effectiveness of the proposed method is verified through the analysis of actual power grid examples. The proposed method can reduce the serious consequences such as the loss of voltage in all substation stations after N-1-1 occurrence, and provide support and guarantee for the operation safety of urban power grid.

Keywords: Diversion strategy, intelligent generation, generalization; N-1-1.

1. Introduction

In recent years, with the rapid economic development, the social demand for electricity is increasing, so to ensure the safety and reliability of power supply is the most important. In the whole power system, dispatching operation is an important part to ensure the normal operation of the power system [1, 2]. Up to now, the dispatching department has developed a perfect plan and measures for the occurrence of N-1 in the urban power grid, so as to ensure that there will be no power outage when N-1 occurs, but it is difficult to ensure that some substations are not in the state of single supply risk after the occurrence of N-1. For example, when the substation bus breaks down, the action of the backup automatic transmission device in the system will ensure that the system will not have a power failure under the current state, but if these substations once the line breaks down again, there is no alternative path available, it is very easy to have a large area of power failure in a certain area. In order to prevent the above situation, the dispatching department adopts the means of load transfer to transfer the load carried by the substation with risks to other alternate paths. However, due to the relationship between a large number of switches and voltage levels in the urban power grid, some traditional methods take a long time to solve and cannot timely formulate decision plans for the dispatching department. Therefore, it is very important for this paper to study how to quickly develop load transfer scheme for dispatching department and reduce the risk of N-1-1 in urban power grid.

At present, the methods for solving load transfer at home and abroad are roughly divided into the following kinds, including heuristic algorithm, expert system method, mathematical programming method and so on. At present, most researches on heuristic algorithms are based on simulated annealing, genetic algorithm, particle swarm optimization algorithm and ant colony algorithm [3-5]. However, heuristic algorithms rely heavily on in-depth understanding of specific problems and domain knowledge. Although feasible solutions can be narrowed to a certain space according to such knowledge, and suitable routing paths can be found in a short time, once the lack of such knowledge, the robustness and universality of the system will be greatly reduced or even the algorithm will be unable to find effective solutions. The main feature of expert system method is that it can call the pre-saved policy library, so it has good real-time performance and wide applicability. Enables it to quickly provide solutions when dealing with complex problems [6]. However, it takes a lot of time and effort to build and integrate these strategy libraries, and the expert system needs to be constantly adjusted to the new operating conditions as the city grid undergoes load transfer. This adjustment process is tedious and time-consuming, which increases the complexity and workload of system maintenance. Mathematical programming method is a popular method in recent years because of its good ability to search for optimization in the whole world, which can transform the complicated load transfer problem of urban power grid into mathematical problem. And when solving the problem of load transfer, it is supported by very strict mathematical theory and formula deduction, which can ensure the real and effective and optimal solution. However, when dealing with a large and complex structure, high dimensional power grid, such methods are extremely prone to "combination explosion" situation [7-10].

In recent years, artificial intelligence has developed rapidly in various fields, and its technology has become increasingly mature, which has also laid the foundation for solving the problem of load transfer in the power system. Among them, the application of deep reinforcement learning technology in power system has a very significant advantage. It is not limited by complex physical models, and a large number of sample data can be automatically generated after the information of system topology and power flow is known. Moreover, deep reinforcement learning can deal with complex nonlinear environment and changeable state space through neural network model [11, 12]. This capability enables it to flexibly adapt to various situations when facing practical problems. The model does not require prior knowledge, and the corresponding strategies can be gradually optimized through the interactive self-help learning between the agent and the power grid environment. The load transfer process of power grid is regarded as a Markov decision process, and the framework of Markov decision process is helpful to optimize the load transfer strategy systematically and quickly in the complex power grid environment.

Therefore, this paper proposes an improved Double DQN load switching research method based on deep reinforcement learning, which significantly improves the convergence speed of the algorithm by introducing an instant reward mechanism and adding a change exploration value selection strategy to the pre-action [13-15]. When N-1 failure occurs in the power grid, this method can respond quickly and provide high-quality load switching decision schemes in real time. This method not only effectively deals with different N-1 fault scenarios, but also ensures the reliability of the system [16-19].

2. Reinforcement Learning Model of Load Transfer

Reinforcement learning is a process of constantly interacting with the environment, obtaining feedback, updating strategies, and iterating until the optimal strategy is learned. The power grid is the environment of reinforcement learning, the agent provides the current power grid state space S, and the result of the agent's analysis and decision is the switching action A. The action is applied to the environment, and the reward value of the environment is R [20-22]. The objective of the reinforcement learning agent is to maximize the cumulative reward value through a limited number of steps, so as to find the optimal strategy [23]. The reinforcement learning model of load transfer is shown in Figure 1.

International Journal of Multiphysics
Volume 18, No. 2, 2024
ISSN: 1750-9548

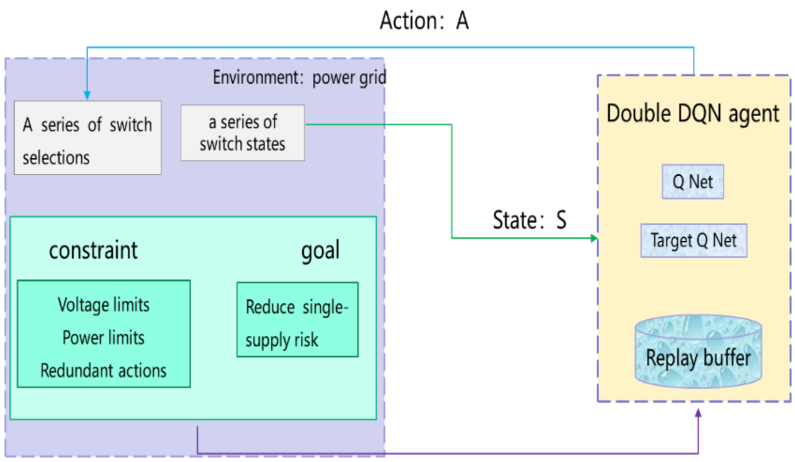

Figure 1 Intelligent load transfer decision architecture based on reinforcement learning

2.1 State space

The state space should consider as far as possible the factors that will affect the decision, so in reinforcement learning, these data are selected to build the state space S, which refers to the limited set of the state of the environment, and S_t represents the state of the environment at the current moment:

$$S = [K_g, G_i] \tag{1}$$

Where, K_g is the switching state vector of the branch in the power grid, and G_i is whether each substation has a single supply.

2.2 Action space

Action space A refers to the set of all actions that the agent can perform on the environment, and the action taken by the agent a_t time t is represented by AT. A complete load switching operation consists of a series of switching switches. In order to prevent the state space from being too large, this paper selects the mode of switching only one switch in one operation. In addition, the load switching should end within a limited number of operations, and the action of actively ending this switching should be set.

$$A = [a_1, a_{2,} ..., a_k, ...a_n] \tag{2}$$

Where, $a_1, a_{2,} ..., a_k, ...a_n$ indicates whether the switch operates, when the vector is 1, it indicates that the switch operates, and n is the number of switches.

2.3 Return function

(1) Bonus part

The main purpose of load transfer in the network is to reduce the risk of N-1-1 occurrence in the network after N-1 failure.

$$R_1 = 10 \times e^{-1n2s} \tag{3}$$

In the formula, s is the number of single supply, and the maximum reward value is when the single supply is 0.

$$R_2 = P_{\text{transform}} \tag{4}$$

Where, $P_{\text{transform}}$ is the amount of load transformed from risk load to non-risk load.

The transfer of supply should be completed in as few operation times as possible to reduce the cost of operation and maintenance and the possibility of error, but also save the operation time, prevent the distribution network structure from changing too much, and increase the difficulty of restoring the original operation mode after the fault is eliminated.

$$R_3 = -(\frac{x-a}{2})^2 + b \tag{5}$$

Where, x is the number of actions, a is the optimal number of actions, ranging from 1 to 8, and b is the maximum value of reward.

(2) Penalty part

In order to ensure the normal operation of the distribution network, it is essential to maintain the node voltage within the specified range. The voltage should be kept within the allowable deviation range of ±7% to ensure the stability of the system. In order to prevent the agent from acting without meeting this voltage constraint, severe penalties are given for voltage values outside this range, but no penalties are set for voltage values within this range. This strategy is designed to guide the agent to select operations that meet the requirements of voltage stability, thereby avoiding behaviors that may cause system instability and ensuring the reliability of grid operation.

The voltage exceeds the limit and enters the failed exit state directly. Formula () is the calculation formula of voltage penalty P_{dy} ,

$$P_{dy} = \begin{cases} P_d, U_i \le U_{i\min} \\ 0, U_{i\min} \le U_i \le U_{i\max} \\ P_d, U_i \ge U_{i\max} \end{cases} \tag{6}$$

Where U_i is the per unit value of the voltage of node i; P_d is the penalty value after the voltage exceeds the limit; $U_{i\min}$ and $U_{i\max}$ are the lower limit and upper limit of node i voltage respectively.

When the transmission power exceeds the limit value, it is easy to cause secondary failure of the equipment in the power grid, and there should be corresponding limit value for each branch, and corresponding punishment should be done after exceeding the limit value. The calculation formula of power penalty P_{gl} is as follows:

$$P_{gl} = \begin{cases} -\frac{P_i}{P_{\max}} r, P_i > P_{\max} \\ 0, P_i < P_{\max} \end{cases} \tag{7}$$

Where, $P_{\max}$ is the maximum transmission power of a branch, and P_i is the power transmitted on a branch.

When the agent performs invalid redundant actions, such as repeatedly turning off the switch that has been turned off or the switch that has been turned off because of the N-1 fault is turned off again, these actions will have an adverse impact on the recovery of the power grid and economic cost in the later stage, so the penalty for invalid actions is given in formula (8).

$$P_{ope} = \begin{cases} P_o, a_i \in C_A \\ 0, a_i \notin C_A \end{cases} \tag{8}$$

Where, P_o is the penalty value for the agent to perform an invalid redundant action, a_i is the agent to perform the i th action, and C_A is the set of switching actions.

Finally, the reward function of the model is composed of the sum of the reward part and the punishment part:

$$R = R_1 + R_2 + R_3 + P_{dy} + P_{gl} + P_{ope} \tag{9}$$

3. Load transfer method based on deep reinforcement learning

3.1 Double DQN algorithm

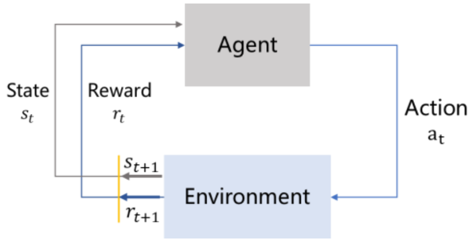

Figure 2 Reinforcement learning process

The whole process of reinforcement learning can be simplified into Markov decision process (MDP), and all states are Markov. At each time step, the agent receives a state. It selects an action according to a preset policy mechanism, receives a corresponding reward according to the dynamic model of the environment, and then moves on to the next state. The policy mechanism defines how the agent determines actions based on the current state. In reinforcement learning, MDP can be represented by a quintuple $\langle S, A, P, R, \gamma \rangle$, and in a fragmented environment, this process continues until it reaches a termination state. The model here refers to the dynamic characteristics of the environment, including the law of state transition and the way rewards are distributed [24], as shown in Figure 2.

The network structure of DQN is composed of target network and estimate network. The two networks are architecturally identical, but have different parameters. The estimation network uses the latest parameters to calculate the value of the current state-action pair and periodically updates these parameters to the target network to calculate the target Q value. With this dual network structure, correlations between data can be broken, enabling DQN to learn from different data distributions. The experience playback unit can store the historical behavior information of the agent, including the current moment state s , action a , reward r and the next moment state. When updating the DQN algorithm, the system will randomly extract some behavior sequences from the experience pool for playback training. This replay training method can alleviate the problem of high correlation between data in the experience pool, and solve the problem of insufficient generalization ability caused by non-static data distribution. The DQN algorithm directly computes the target Q value using a greedy strategy, which quickly converges to the best possible target by maximizing the Q value. However, this method can easily lead to overestimation of Q value and introduce large deviation, which is unfavorable to the research of robot operation behavior. In order to solve this problem, the Double DQN algorithm decouples the action selection from the calculation of the target Q value to avoid the overestimation of the Q value [25, 26].

Double DQN algorithm is an improved version of DQN algorithm, which solves the problem of overestimating behavior value of DQN algorithm. In the DQN algorithm, when the state at a certain time is a non-terminating state, the calculation formula of the target Q value is as follows:

$$y_j = r_j + \gamma \max_{a'} Q(s_{j+1}, a'; \theta') \tag{10}$$

The Double DQN algorithm does not directly select all possible Q values calculated by the target network in a maximized way, but first selects actions corresponding to the maximum Q value by estimating the network. The formula is expressed as follows:

$$a_{\max} = \arg\max_a Q(s_{t+1}, a; \theta) \tag{11}$$

Then the target network calculates the target Q value according to $a_{\max}$, and the formula is expressed as follows:

$$y_j = r_j + \gamma Q(s_{j+1}, a_{\max}; \theta') \tag{12}$$

Since the final goal of Double DQN is to minimize the gap between the estimated Q value and the real target's Q value, the formula () can be obtained and the loss function can be defined:

$$\delta = |Q(s_t, a_t) - y_i| = |Q(s_t, a_t; \theta) - (r_t + \gamma Q(S_{t+1}, \arg\max_a Q(s_{t+1}, a; \theta); \theta'))| \tag{13}$$

$$loss = \begin{cases} \frac{1}{2}\delta^2, |\delta| \le 1 \\ |\delta| - \frac{1}{2}, else \end{cases} \tag{14}$$

There are two neural network models in the Double DQN framework, namely the training network and the target network. The structure of the two neural network models is exactly the same, but the weight parameters are different. After each training interval, the weight parameters of the training network are copied to the target network. During the training of the whole model, the training network is responsible for estimating the current $Q(s_t a_t)$, and the target network is responsible for estimating the $\max_a Q(s_{t+1}, a_t)$. The target network ensures that the estimate of the true value $Q_{target}(s_t, a)$ will not change too quickly with the continuous update of the training network. Double DQN also supports offline learning. The past experience is learned offline by constructing an experience pool. The loss function of the training network is mean-square error $MSE(Q_{train}, Q_{target})$, and the training model is updated by backpropagation of gradient descent method. Finally, after several rounds of experience pool sampling, the weight of the training network model is assigned to the target model, and the model self-learning of Double DQN is carried out, as shown in Figure 3.

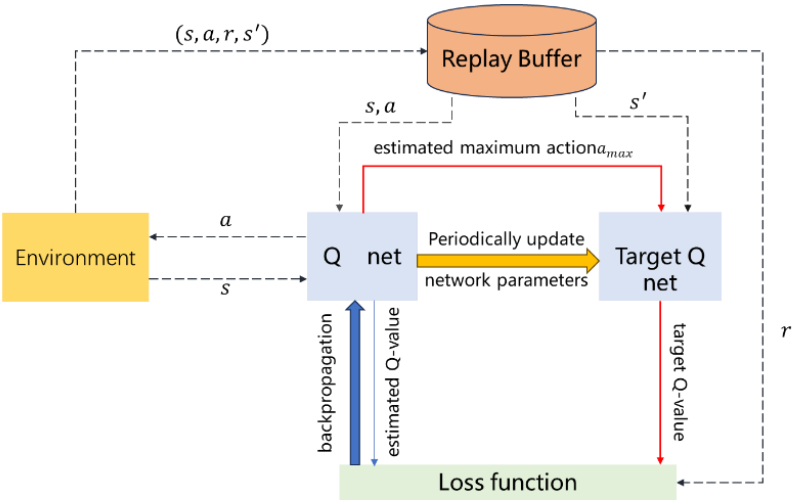

Figure 3 Double DQN algorithm structure diagram

3.2 Explore the greed mechanism in pre-action

Consider that rewards play a key role in action evaluation and Q-value optimization in reinforcement learning. Before the neural network is fully trained, when the Q value may not be stable, using instant rewards as a guide to choose the next action can help the agent explore the environment more effectively and discover the optimal strategy. The pre-action exploration strategy guides the action selection by evaluating the immediate reward of the predicted action, enabling the agent to quickly accumulate high-quality experience samples in the early learning period. This approach not only improves the efficiency of exploration, but also helps agents adapt and optimize decision-making strategies quickly in complex environments.

4. Example Analysis

4.1 Training process

As shown in Figure 4, in order to prove the effectiveness of the proposed model method, a local power grid in a certain region is used as an example. The power grid topology contains 4 power stations, 6 220kV substations, 12 110kV substations, 12 220kV grade lines and 36 110kV grade lines. 158 nodes, 196 operable switches. The simulation environment is Python3.7, pytorch1.10, CPU is i7-11800, and memory is 16GB.

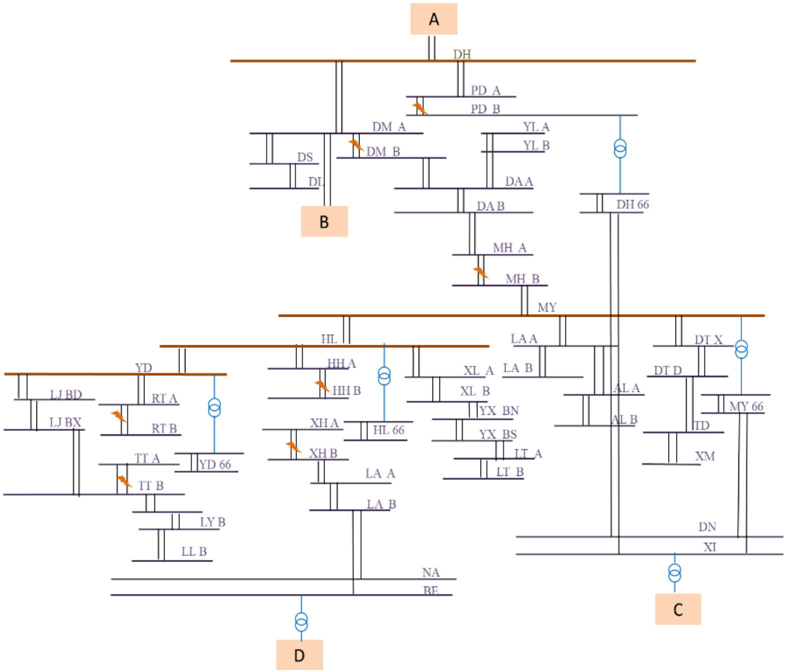

Figure 4 Topology diagram of a regional power grid and break position

Double DQN neural network is composed of input layer, hidden layer and output layer. The hidden layer of the two agents has 64, 48 neurons, 24 and 16 neurons respectively. The discount coefficient is 0.98, the capacity of experience playback pool is 1000, and the number of learning samples is 100. The learning rate, which starts at 0.001, decreases over time.

In order to verify the effectiveness of the load transfer control method based on Double DQN proposed in this paper, the scheduling method based on Double DQN algorithm is compared with the DQN algorithm, the same reward function is set for them, and the average reward value obtained during the training process is compared.

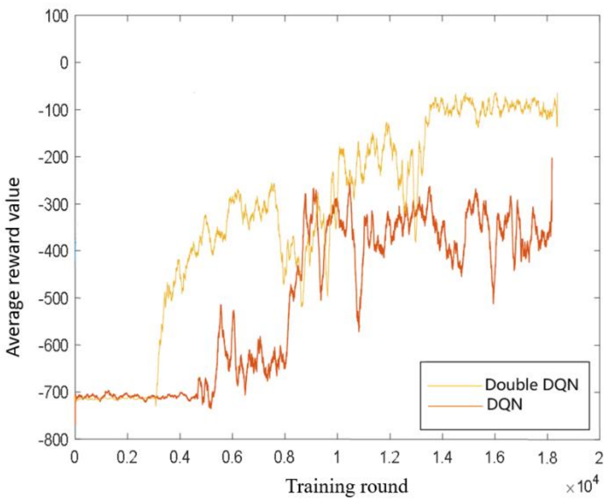

Figure 5 Average reward value comparison

As can be seen from Figure 5, in the early stage of training, due to the tendency of the flow to not converge, the reward level of the comparison method is low, while the action selection mechanism of the proposed method introduces the immediate reward mechanism to obtain a higher average reward, while the pre-action-change exploration value selection strategy improves the global optimal convergence ability. At the later stage of training, certain oscillations are due to the probability that the agent attempts random actions to avoid falling into local optimality. Double DQN algorithm and classical DQN algorithm are prone to pursue risk prevention and control effect, which makes it difficult to restore the radiation network and obtain a relatively perfect transfer strategy. As a result, it is difficult to provide the optimal convergence speed and higher reward value of the method.

4.2 Analysis of multi-scenario intelligent decision strategy results

In order to further verify the applicability of intelligent online decision making in different N-1 scenarios concerned by scheduling departments, multiple N-1 topology scenarios are set for testing. After training, the proposed method and model still have good decision-making performance in different scenarios.

Table 1 Transfer strategy 1 (DH Station: PD Line A and B, DM Line A and B Outages)

Breakdown	DH Station: PD Line A and B, DM Line A and B Outages		
Risk	Apart from the DS and DL lines, all other loads at the DH station are supplied only by the YL line. If this line goes down, it will lead to a significant loss of load.		
Procedure	Switch Operation		Turn (string) supply path
1	DA Line A to AL Line A Switch	Open → Close	DH station east bus load transfers to MY station
	YL Line A Switch	Close → Open	
	DH Station 66kV Bus Coupler		
Transfer result	The loads on the two buses at the DH station are supplied by CY Station and MY Station respectively, reducing the likelihood of simultaneous power outages.		

Table 2 Transfer strategy 2 (MY Station: MH Line A and B Outages)

Breakdown	MY Station: MH Line A and B Outages		
Risk	The loads at MY Station are supplied solely by the LA line. If this line goes down, it will result in a complete loss of load.		
Procedure	Switch Operation		Turn (string) supply path
1	DT Line West Bus Side Switch	Open → Close	TD load changes to XM
	DT Line East Bus Side Switch	Close → Open	
2	DA Line A to AL Line A Switch	Open → Close	MY station east bus load transfers to DH station
	LA Line A Switch	Close → Open	
	MY Station 66kV Bus Coupler		
Transfer result	The loads on the two buses at MY Station are supplied by DH Station and HL Station, and the DT line load is transferred to prevent equipment overload, reducing the likelihood of simultaneous power outages.		

Table 3 Transfer strategy 3 (HL Station: HH Line A and B, XH Line A and B Outages)

Breakdown	HL Station: HH Line A and B, XH Line A and B Outages		
Risk	At HL Station, aside from the LA Line A, the remaining loads are only supplied by the XL line. If this line goes down, it will result in a significant loss of load.		
Procedure	Switch Operation		Turn (string) supply path
1	TX Line B North Bus Side Switch	Open → Close	TX line B load transfer to BM
	TX Line B South Bus Side Switch	Close → Open	
2	LY Line B Switch	Open → Close	HL station west bus load transfers to TD station
	XL Line B Switch	Close → Open	
	HL Station 66kV Bus Coupler		
Transfer result	The loads on the two buses at HL Station are supplied by TX Station and TD Station, and the TX Line B load is transferred to prevent equipment overload, reducing the likelihood of simultaneous power outages.		

As can be seen from Table 1 to Table 3, under the load switching strategy proposed in this paper, there is no substation single supply risk, and the load loss risk is at a relatively low level. N-1-1 check conforms to the requirements of the risk strict control event level (there is no whole station voltage loss of substations above 110kV, and the load loss is less than 50MW). However, after the empirical decision optimization, there are high risks of the whole substation voltage loss and load loss risk, which is due to the lack of a systematic evaluation system for N-1 new risks when the plan is formulated.

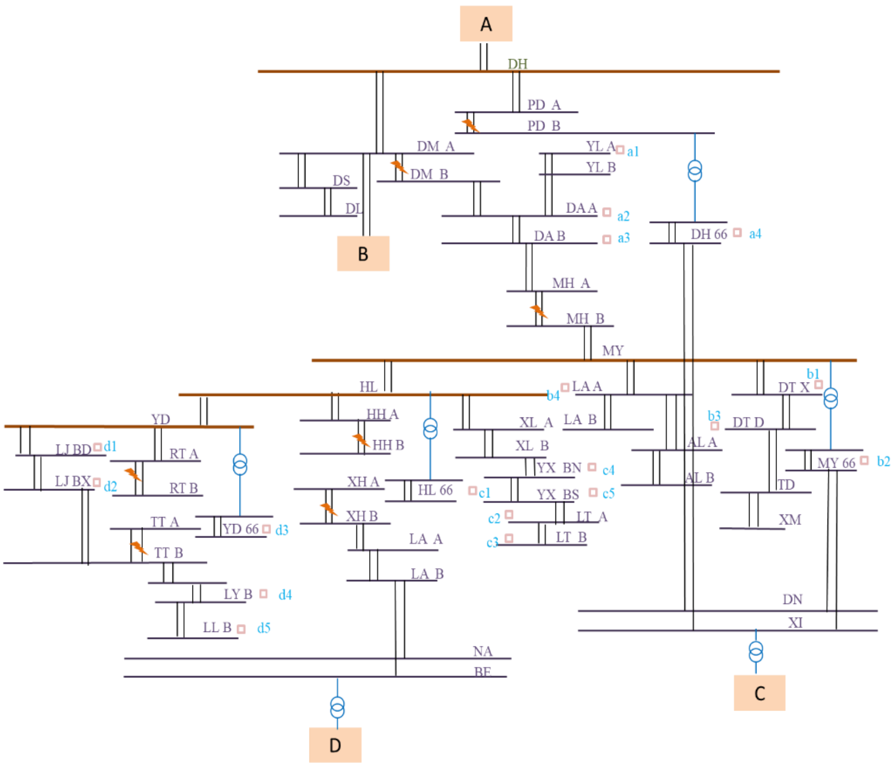

Figure 6 Switch switching position

Table 4 Comparison of transfer results

Transfer method	Transfer scheme	Number of substations at risk (units)	Maximum load rate (%)
non-optimization	-	4	42.78
Textual method	①open a1 ②close a2 ③open b1 ④close b3 ⑤close d1 ⑥open d4 ⑦close c1 ⑧open c2	0	67.17

The load transfer results of the urban power grid after N-1 are analyzed and compared with the non-optimization method. The load transfer results are shown in Table 4, and the switch positions are marked in Figure 6. In this paper, the maximum load ratio of the line with the switching strategy is 67.17%, which indicates that the load ratio of the power network is relatively uniform.

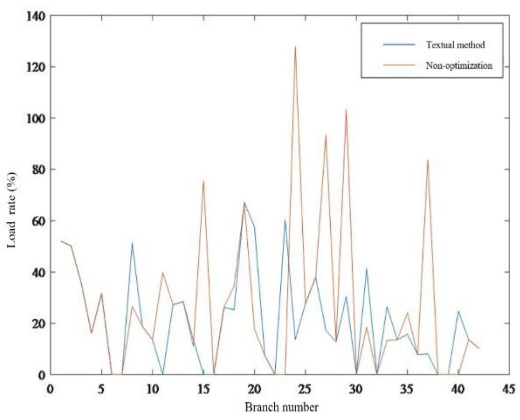

Figure 7 The comparison of load rates after transfer under different methods

As shown in Figure 7, the Double DQN switching strategy is effective for reducing the load ratio. Although the load ratio of some light-load branches is greater than that of the non-optimization method, this is the inevitable result of transferring the load of heavy-load branches.

5. Conclusion

In this paper, the deep reinforcement learning method is applied to the process of power grid load transfer. Through direct and effective analysis of the operating environment information of the power grid, effective data is extracted to build a deep reinforcement learning switching model, and a series of switches of the power grid are reasonably and effectively controlled to realize load switching. This method can adapt to the N-1 situation across multiple scenarios in the power grid, and does not need to modify the model for different fault types, thus providing a more generalized switching strategy.

In the algorithm design, this paper improves the effectiveness of decision making and speeds up the training speed significantly by using Double DQN method. Compared with the traditional algorithm, this method has the dual advantages of offline learning and online application. By transferring a large number of calculations to the offline stage, the algorithm can accumulate rich experience through offline learning of massive data. When an N-1 failure occurs in the power grid, the algorithm can quickly perform online calculation and provide accurate and effective control strategy for operators in a very short time. The power failure loss is reduced and the operating cost is reduced.

In addition, the improved algorithm proposed in this paper can also handle more complex load switching scenarios, such as different line break fault locations. The algorithm can dynamically adjust its own parameters to adapt to different operating environments and conditions. This flexibility and adaptability make the algorithm more applicable and valuable in real-world scenarios.

To sum up, the innovation of this paper lies in the combination of artificial intelligence deep reinforcement learning technology and power system load transfer, and puts forward a method with high efficiency and adaptability. Through the in-depth experimental verification, the method shows remarkable superiority in improving the reliability of power grid operation.

Acknowledgements

This research is funded by the 2024 High-Quality Development Strategy Project of State Grid Jilin Electric Power Co., Ltd., Contract No. SGJLJY00ZLJS2400058.

References

[1] Li Peng, YU Tao, LI Li-port, “Evolution and prospect of electric power artificial Intelligence: From professional intelligence to general intelligence”, Automation of Electric Power Systems, pp. 1-21, June 2024. (Article)

[2] Gao Guanzhong, Yang Shengchun, Guo Xiaorui, “Application of deep reinforcement learning in power grid optimal scheduling with distributed flexible resources”, Proceedings of the CSEE, pp.1-21, June 2024. (Article)

[3] Zhang Haibo, Zhang Xiaoyun, Tao Wenwei. “Distribution network fault recovery Algorithm based on breadth-first search”, Network Technology, Vol. 34, No. 7, pp. 103-108, 2010. (Article)

[4] Wang Zheng. “Power Supply restoration of Distribution Network based on Heuristic Search Method”, North China Electric Power University (Hebei), 2007.

[5] Li Hongwei, LIU Yulu, Jin Yong, “Decision Strategy of switch action timing in Load Transfer of High voltage distribution Network”, Power System Automation, Vol. 42, No. 1, pp. 82-90, 2018. (Article)

[6] Liu Weijia, Lin Zhenzhi, Wen Fusuan, “Considering the multi-objective optimal strategy for power system Recovery supported by electric vehicles” Power System Automation, Vol. 39, No. 20, pp. 32-40, 2015. (Article)

[7] Bo C, Chen C, Jianhui W, “Sequential Service Restoration for Unbalanced Distribution Systems and Microgrids”, IEEE Transactions on Power Systems, Vol. 33, No. 2, pp. 1507-1520, 2018. (Article)

[8] Dai Zhi-Hui, CHONG Zhi-qiang, JIAO Yanjun, “Multi-objective power supply recovery in distribution network with Distributed Power Supply”, Power Network Technology, Vol. 38, No. 7, pp. 1959-1965, 2014. (Article)

[9] Zeyu W, Jianhui W, Chen C, “A Three-Phase Microgrid Restoration Model Considering Unbalanced Operation of Distributed Generation”, IEEE Transactions on Smart Grid, Vol. 9, No. 4, pp. 3594-3604, 2018. (Article)

[10] Farzin, Hossein, Fotuhi-Firuzabad, “Enhancing Power System Resilience Through Hierarchical Outage Management in Multi-Microgrids”, IEEE transactions on smart grid, Vol. 7, No. 6, pp. 2869-2879, 2016. (Article)
[11] Oh H S, Yoon T Y, Kim W S, “Online reconfiguration scheme of self-sufficient distribution network based on a reinforcement learning approach”, Applied Energy, 2020. (Article)
[12] L. N D, B. A J, Abdollah S, “Intelligent hurricane resilience enhancement of power distribution systems via deep reinforcement learning”, Applied Energy, pp. 285, 2021. (Article)
[13] Jun A, Liang Z, Yibo Z, et al. “Transient Stability Margin Prediction Under the Concept of Security Region of Power Systems Based on the Long Short-Term Memory Network and Attention Mechanism”, Frontiers in Energy Research, 2022, 10. (Article)
[14] J. C. Bedoya, Y. Wang and C. -C. Liu, “Distribution System Resilience Under Asynchronous Information Using Deep Reinforcement Learning”, IEEE Transactions on Power Systems, DOI: 10.1109/TPWRS.2021.3056543, Vol. 36, No. 5, pp. 4235-4245, Sept. 2021. (Article)
[15] M. Kamruzzaman, J. Duan, D. Shi and M. Benidris, “A Deep Reinforcement Learning-Based Multi-Agent Framework to Enhance Power System Resilience Using Shunt Resources”, IEEE Transactions on Power Systems, DOI: 10.1109/TPWRS.2021.3078446, Vol. 36, No. 6, pp. 5525-5536, Nov. 2021. (Article)
[16] S. Safiri, A. Nikoofard, M. Khosravy and T. Senjyu, “Multi-Agent Distributed Reinforcement Learning Algorithm for Free-Model Economic-Environmental Power and CHP Dispatch Problems”, IEEE Transactions on Power Systems, DOI: 10.1109/TPWRS.2022.3217905, Vol. 38, No. 5, pp. 4489-4500, Sept. 2023. (Article)
[17] M. S. Massaoudi, H. Abu-Rub and A. Ghrayeb, “Navigating the Landscape of Deep Reinforcement Learning for Power System Stability Control: A Review”, IEEE Access, DOI: 10.1109/ACCESS.2023.3337118, Vol. 11, pp. 134298-134317, 2023. (Article)
[18] J. Khalid, M. A. M. Ramli, M. S. Khan and T. Hidayat, “Efficient Load Frequency Control of Renewable Integrated Power System: A Twin Delayed DDPG-Based Deep Reinforcement Learning Approach”, IEEE Access, DOI: 10.1109/ACCESS.2022.3174625, Vol. 10, pp. 51561-51574, 2022. (Article)
[19] Fengzhang Luo, Chengshan Wang, Jun Xiao, “Rapid evaluation method for power supply capability of urban distribution system based on N−1 contingency analysis of main-transformers”, International Journal of Electrical Power and Energy Systems, Vol. 32, No. 10, pp. 1063-1068, 2010. (Article)
[20] Cheng Lin, Wan Yuxiang, Qi Ning, “Coordinated operation strategy of distribution network with the multi-station integrated system conside-ring the risk of controllable resources”, International Journal of Electrical Power and Energy Systems, pp. 137, 2022. (Article)
[21] Beenish Sultana, M. W. Mustafa, U. Sultana, Abdul Rauf Bhatti, “Review on reliability improvement and power loss reduction in distribution system via network reconfiguration”, Renewable and Sustainable Energy Reviews, vol. 66, pp. 217-310, 2016. (Article)
[22] M. A. Tavakoli Ghazi Jahani, P. Nazarian, A. Safari, M. R. Haghifam, “Multi-objective optimization model for optimal reconfiguration of distribution networks with demand response services”, Sustainable Cities and Society, pp. 47, 2019. (Article)
[23] Yibo Z, Jun A, Gang M, et al. “An improved constraint-inference approach for causality exploration of power system transient stability”, Protection and Control of Modern Power Systems, 2023, 8(1). (Article)
[24] An Jun, Li Zicong, Zhou Yibo, et al. “Power Grid N-1 New Risk Prevention and Control Strategy Based on Improved Dual-Agent D3QN”, Proceedings of the Chinese Society for Electrical Engineering, 1-13 [2024-07-06]. (Article)
[25] Zhang Liang, An Jun, Zhou Yibo. “Transient Stability Assessment of Power Systems Based on Temporal Convolution and Graph Attention Networks”, Automation of Electric Power Systems, 2023, 47(07): 114-122. (Article)
[26] Zhou Yibo. “Research on Data-Driven Methods for Identifying the Relationships and Causality Between Variables in Transient Stability Assessment of Power Systems”, Northeast Electric Power University, 2023. (Dissertation)