Fast Reconfiguration of Distribution Network Based on Deep Reinforcement Learning Algorithm

Bincheng Zhao1*, Xueshan Han¹, Yiran Ma¹ and Zhiqi Li¹

¹School of Electrical Engineering, Shandong University, Jinan, Shandong, 250061, China
*Corresponding author’s e-mail: 201834405@mail.sdu.edu.cn

Abstract. In the context of large-scale grid connection of distributed energy, during the reconfiguration of the distribution network, the availability of distributed energy and the load of the distribution system may be inconsistent with the prediction due to the influence of environmental factors and human factors. If the distribution network reconfiguration is still carried out according to the expected offline optimization scheme, there may be reliability problems of voltage over-limits and economic problems of increased network loss in the actual reconfiguration process. Therefore, the reconfiguration plan formulated in advance can give some guidance to the dispatch operator, but it may not be directly used in the actual reconfiguration process. This paper proposes a deep reinforcement learning approach to solving the electric distribution network reconfiguration. Based on the uncertainty of distributed energy output and network load in the distribution network, the online algorithm of distribution network reconfiguration realizes the second-level solution of distribution network reconfiguration, through day-ahead training of the neural network.

1. Introduction

Electric distribution network reconfiguration is an optimization decision-making process for the purpose of reducing losses, balancing loads, improving power supply reliability, safety, and economy by adjusting the position of tie switch of the distribution network. The results of distribution network reconfiguration generally require that the distribution network radial constraints be met. In recent years, distributed energy has developed rapidly, and its economic and environmental characteristics have also been paid more and more attention, and will be more widely used.

At present, domestic and foreign scholars have conducted a lot of research on the reconfiguration method of distributed power access to the distribution network. Reference [1] introduces interval numbers to describe various uncertainties, and uses interval numbers to describe the minimum network loss as the objective function, and establishes a description method of distribution network reconfiguration with DG and electric vehicles. Reference [2] proposes to make full use of active distribution network rapid reconfiguration and distributed energy control measures to reduce operating costs and guide more distributed energy consumption. Reference [3] considers the difference in credibility periods, and proposes to formulate a reasonable scheduling plan through coordinated optimization, leaving greater margin for periods with poor credibility and reducing operational risks. These studies provide a theoretical basis for the distribution network reconstruction after distributed energy access, but they all reduce the uncertainty of distributed energy and load on the distribution network by increasing the conservativeness of the distribution system operation. Operational risk. They
failed to solve the problem of how to quickly realize a reasonable distribution network reconfiguration when the source load level of the distribution network is determined and does not meet expectations.

The day-ahead reconfiguration of the active distribution network is based on the day-ahead forecast data of distributed energy and load, and a day-ahead reconfiguration plan that meets the expected goals is formulated. However, there is a certain error between the current forecast and the actual operating conditions, which may lead to operational risks such as voltage over-limit, power overload, and increased network loss in the actual operation of the recent reconfiguration plan, which requires fast online adjustment. The voltage management of the distribution network must meet greater challenges.

In order to improve the applicability and safety of the distribution network reconfiguration scheme in actual operation, this paper uses deep reinforcement learning (DRL) algorithm to pre-learn the optimal switch state selection for the power supply and load level in the distribution network operation history. After the learning is completed, the optimal topology state of the distribution network with different source and load states can be quickly obtained. Finally, the IEEE14 node system is simulated to verify that the proposed algorithm has a faster network reconfiguration capability than other heuristic algorithms when dealing with the deviation of the actual situation of distributed energy and load from the forecast.

2. Distribution network reconfiguration model with distributed energy

2.1. Objective function of distribution network reconfiguration

In distribution network reconstruction, the objective function is usually the least active power loss in the network.

\[
\begin{align*}
\min P_{\text{loss}} &= \min \sum_{i=1}^{br} k_{i,jk} R_{i,jk} P_{i,jk}^2 + Q_{i,jk}^2 \overline{U}_{i,jk} \\
\end{align*}
\]

Among them, \(br\) represents the total number of branches in the distribution network; \(j\) and \(k\) represent the first and last node numbers of branch \(i\); \(k_{i,jk}\) represent the running status of the line, \(k_{i,jk}\) takes 1 to indicate branch \(jk\) normal operation, \(k_{i,jk}\) takes 0 to indicate that the line tie switch is off and exit the operation. \(R_{i,jk}\) represents the resistance of the branch \(jk\); \(P_{i,jk}\), \(Q_{i,jk}\) represent the active and reactive power flowing through the line \(jk\); \(U_{i,k}\) is the voltage amplitude of the node \(k\) at the end of the line \(i\).

In the case of distributed energy, the direction of power flowing through the branch \(jk\) may change. At this time, the first and last node numbers of the branch \(jk\) should be considered to be swapped to obtain the branch active power loss.

2.2. Constraint Conditions for Distribution Network Reconfiguration

In the optimization process of distribution network reconfiguration determined in this paper, the constraints to be considered include: radial reconstruction constraints, line capacity constraints, node voltage constraints, power flow constraints, and DG output constraints. The specific mathematical expressions are as follows:

1) Power flow constraints

Power flow calculation model of distribution network considering distributed energy.

\[
\begin{align*}
P_i &= U_i \sum_{j=1}^{N} U_j \left( G_{i,j} \cos \theta_{i,j} + B_{i,j} \sin \theta_{i,j} \right) - P_{\text{DG}_i} \\
Q_i &= U_i \sum_{j=1}^{N} U_j \left( G_{i,j} \sin \theta_{i,j} - B_{i,j} \cos \theta_{i,j} \right) - Q_{\text{DG}_i} \\
\end{align*}
\]

(2) Node voltage constraints

\[
U_{i,\text{min}} \leq U_i \leq U_{i,\text{max}} \quad i = 1, \ldots, n
\]

That is, the actual value of the node voltage should meet the upper and lower limit values of the node voltage. Generally the upper limit value is 1.05, and the lower limit value is 0.9.

(3) Branch capacity constraints
\[ P_{i,j}^2 + Q_{i,j}^2 \leq S_{i,j}^2 \quad i = 1, \cdots, br \]  

(4)

\( P_{i,j} \) and \( Q_{i,j} \) respectively represent the active and reactive power flowing through the branch \( jk \), and \( S_{i,j}^{\text{max}} \) represents the maximum complex power transmission capacity of the line.

(4) DG output constraints

\[ P_{i,\text{DG}}^2 + Q_{i,\text{DG}}^2 \leq S_{i,\text{DG}}^2 \quad i = 1, \cdots, x \]  

(5)

\( P_{i,\text{DG}} \) and \( Q_{i,\text{DG}} \) represent the active and reactive power output of the \( i \)-th grid-connected distributed energy, and \( S_{i,\text{DG}} \) represents the capacity of the \( i \)-th distributed energy grid-connected inverter.

(5) Network radial constraints

After the contact switch is closed to form a ring network, a segmented switch in the ring network should be opened to ensure the radial topology of the electric distribution network.

2.3. Power model of load and distributed energy

Assuming that both load power and distributed energy output follow a normal distribution, as shown below

![Figure 1. Probability distribution of load and output power of distributed energy](image)

In the figure, \( f_{p,l}(P_{p,l}) \) represents the probability density function of the distributed photovoltaic power prediction value and load prediction value of each node at time \( t \). As shown in the figure, the distribution interval of expected power in period \( t \) can be expressed as \([P_{l,i} - \beta \cdot \sigma_{l,i}, P_{l,i} + \alpha \cdot \sigma_{l,i}]\) and \([P_{p,l} - \beta \cdot \sigma_{p,l}, P_{p,l} + \alpha \cdot \sigma_{p,l}]\). Among them: \( \sigma_{p,l} \), \( \sigma_{l,i} \) is the standard deviation of the power probability distribution; \( \alpha \), \( \beta \) is the confidence coefficient of the actual power above and below the predicted value.

Generally in the distribution network, distributed photovoltaic often adopts constant power control, taking power factor \( \cos \phi = 0.9 \).

3. DRL algorithm for distribution network reconfiguration

3.1. Description of distribution network reconfiguration problems

In fact, the different power and load states that may appear in the distribution network are continuous, and the actions of the distribution network reconstruction are discrete. For each source and load state of the distribution network, there must be optimal and sub-optimal switch selection, so that in each reconfiguration process, experience can be accumulated in the switch selection. In view of the large number of typical source and load sample data saved in the operation history of the distribution network, the reinforcement learning theory to solve the distribution network reconfiguration problem includes the following parts:

(1) Step

There is a limited number of tie switches in each distribution network. Closing a tie switch to form a ring network, and then opening a segmented switch in the ring network to form a radial network is called a step. Therefore, the number of tie switches in the distribution network is the total number of stages in
the distribution network reconfiguration. The variable describing the step is called the step variable, denoted by $i$, and there are $n$ tie switches in total.

(2) Operation status of distribution network

During the operation of the distribution network, a certain topology structure of the distribution network, the distributed energy of each node and the power of the load are expressed as the operation state of the distribution network. In the $i$-th step of the distribution network reconfiguration, the distribution network operation state $s_i$ is transferred to the operation state $s_{i+1}$ through the distribution network reconfiguration action.

(3) Distribution network reconfiguration action

It is defined that in a certain operation state of the distribution network, the breaking of a section switch in step $i$ is an action, denoted as $a_i$. The action space of each stage is defined as the environment of the next stage, that is, the determination of the switch operation in the previous stage will greatly reduce the number of the entire switch operation in the next stage and reduce the operation space.

(4) Reward function $R$

For distribution network reconfiguration, when the final step of the switch state is selected, the network losses of the distribution network are different. There should be a greater selection trend for distribution network reconfiguration methods with good objective function improvement effects, so positive values are given to reconfiguration measures with good reconfiguration effects, and higher values are given to measures with better reconfiguration effects, negative values are given to poorly effective reconfiguration measures to form a reward function based on the advantages and disadvantages of the reconfiguration scheme.

$$R = \frac{P_{loss} - P_{loss,r}}{P_{loss}}$$

In the formula, $P_{loss,r}$ represents the network loss value of the actual reconstruction result, and $P_{loss}$ represents the network loss value of the initial operation state of the distribution network.

3.2. Deep reinforcement learning algorithm (DRL)

Reinforcement learning has special requirements for approximate functions: The approximation function must be differentiable; in order to work online, the approximation function must be able to adapt to the incremental learning method; the learning framework based on generalized strategy iteration causes the algorithm objective to change continuously during the iteration process. Even if it does not change, bootstrapping will also lead to non-stationarity, that is, the reinforcement learning training process belongs to the optimization problem of variable objective function, so the approximate function Handle non-stationary objective functions. The use of neural networks to replace the $Q$ table makes the DRL algorithm have the potential to deal with complex scenes, that is, high-dimensional state and action space. At the same time, the use of approximation functions makes the DRL algorithm have the ability to derive unknown states from the known state to a certain extent, so that the algorithm has the ability to deal with partially observable problems.

3.3. Distribution network reconfiguration based on DRL

Based on the generalized strategy iteration framework, the convergence of RL depends entirely on the strategy promotion method, and its commonly used methods are greedy method and strategy gradient method. In this paper, the strategy of distribution network reconfiguration is improved by Monte Carlo tree search (MCTS), a classic algorithm. The DRL strategy is used to guide the truncation and selection of MCTS depth and breadth. MCTS helps DRL achieve prediction and strategy improvement, and the two complement each other.

MCTS algorithm is very suitable for online decision-making problem of distribution network reconfiguration. In the process of distribution network reconfiguration, due to the uncertainty of distributed energy output and load level, the actual distribution network reconfiguration time and plan
may be different from expected. The selection of MCTS can realize the termination of the algorithm and provide reconstruction Program at any time. On the other hand, MCTS not only gives the optimal reconstruction plan, but also sorts each reconstruction plan according to the value of the income, and provides some alternative reconstruction plans. When the main reconstruction scheme based on DRL does not meet the reconstruction constraints, the rapid reconstruction can be achieved through the backup scheme.

Based on the content in 1.3, this paper discretizes the confidence intervals of the distributed energy and load power prediction values at each time into interval forms, and the length $\Delta P$ of each interval is expressed as follows

$$\Delta P = \frac{P_{p,d}}{K/2}$$

In the formula, $K$ is an even number, $P_{p,d}$ is the predicted value of the distributed energy or load at time $t$. After discretization, it includes $[0, \Delta P), \ldots, [P_{p,d}, P_{p,d} + \Delta P), \ldots, [P_{p,max} - \Delta P, P_{p,max}]$, a total of $K$ intervals. If the actual output of distributed energy or load falls within interval $[a, b)$, the state value is taken as $\frac{a+b}{2}$, and the number of segment switches of the ring network formed in each step is $n_1, n_2, \ldots, n_K$. Let the distribution network have a total of node nodes and $x$ distributed nodes in the distribution network. For energy access, in section 2.1, the state input space of the distribution network contains $n_1, n_2, \ldots, n_K \cdot K^{mode+1}$ states.

Based on the input state space set as described above and the action set in Section 2.1, the pre-learning of the neural network is first performed, and the learning content is first of all the various operating states that appear in the historical operation of the distribution network. Secondly, according to the probability distribution within the interval of distributed energy and load power, corresponding probability samples are taken for different power values to obtain new state space samples. Through the above two steps of pre-learning, the neural network can be optimized. When the $K$ value is larger, the state space of the distribution network reconstruction is larger, the learning time is longer, but the corresponding online calculation in the real environment is obtained The higher the feasibility of the solution.

The calculation steps are as follows:

1. In offline learning, the state-action value functions of the switching probabilities of the Actor neural network and the topological states of the Critic neural network are initialized randomly. In online learning, the neural network value obtained by offline learning is the initial value.

2. According to the current switch state and the current distributed energy output and load power, determine the current distribution network operating state. Under the guidance of Actor and Critic neural network, MCTS searches for the optimal switching action at each stage, and the Actor network can be updated at each step until the final step. Select the switch with the highest $Q$ value.

3. Perform power flow calculation based on the optimal switching action obtained in step (2), find the network loss in this switch state, get the actual action reward, and correct the Critic neural network.

4. To judge whether the distribution network reconfiguration process has converged, the judgment criterion is when the distribution network reconfiguration finally reaches the preset maximum number of iterations or the neural network action probability and state action value update are less than the set threshold

4. Example analysis

In order to test the actual performance of the proposed algorithm, this paper calculates the distribution network reconstruction of IEEE14 nodes based on python3.7, and sets nodes 9 and 14 to access distributed energy, and the inverter capacity is 800MVA. The topology of the IEEE14 node distribution network is as follows, and tables 1 and 2 are the difference between the actual value and the predicted value and the result of the distribution network reconstruction.
Table 1. Forecast and actual value of distributed energy and load (kW)

| Node | Forecast / Actual distributed energy | Forecast / Actual load |
|------|-------------------------------------|------------------------|
| 3    | 0/0                                 | 4000/1500              |
| 6    | 0/0                                 | 3000/5000              |
| 8    | 0/0                                 | 5000/1500              |
| 9    | 0/600                               | 4500/2500              |
| 14   | 0/300                               | 2100/2100              |

Table 2. Predicted optimal reconstruction and actual optimal reconstruction.

|                      | Open branch | Active power loss(kW) |
|----------------------|-------------|-----------------------|
| Before reconfiguration | 14, 15, 16  | 511                   |
| Prediction reconfiguration | 9, 10, 16  | 466/253               |
| Actual reconfiguration  | 14, 15, 16  | 250                   |

5. Conclusion
Towards distribution network reconfiguration, different power and load levels have their potential characteristics, which will lead to different optimal switching states. This potential feature may be difficult for people to directly observe. Reinforcement learning finds this potential optimal switch combination state discrimination basis through continuous exploration and trial in the complex distributed energy and load state. Solving complex multivariate decision problems has great advantages. Reinforcement learning has great advantages for solving such complex multivariate decision-making problems. Relevant scholars are also constantly exploring its application in power system optimization, such as online decision-making of power system unit recovery [4], cooperative decision of wind storage [5], etc., but its application research in topology switching decision-making is still scarce. In this paper, the effectiveness of the algorithm is demonstrated through theoretical analysis and calculation examples.

Acknowledgement
Teacher Han is a knowledgeable person. From him, I can feel the real scholarship. He sublimated his professional knowledge to a more systematic and deeper understanding and passed it on to me. In the future, I will also remember the teachings of Teacher Han. Here, I would like to express my most sincere admiration and sincere gratitude to my teacher!

References
[1] Li, Y., Wei, G., et al. (2018) Dynamic Reconfiguration of Active Distribution Network Considering Electrical Vehicles and Distributed Generations. J. Automation of Electric Power Systems. Commun., 42(05):102-110.
[2] Zhou, H.Z., Tang, F., et al. (2016) Active Distribution Network Dynamic Reconfiguration and DG Dynamic Control Strategy Considering Time-Variant Load. J. Power System Technology. Commun., 40(08):2423-2430.
[3] Ma, J.W., Sun, Q., et al. (2018) Dispatching Strategy for Distribution Network with Distributed Renewable Energies Considering Forecasting Credibility Interval Difference. J. Proceedings of the CSU – EPSA. Commun., 30(05):56-62.
[4] Sun, R.J., Liu Y.T., et al. (2018) Online Decision-making for Generator Start-up Based on Deep learning and Monte Carlo Tree Search. J. Automation of Electric Power Systems. Commun., 42(14):40-47.

[5] Liu, G.J., Han, X.S., et al. (2016) Optimal Decision-Making in the Cooperation of Wind Power and Energy Storage Based on Reinforcement Learning Algorithm. J. Automation of Electric Power Systems. Commun., 40(09):2729-2736.