


Article

Substation Operation Sequence Inference Model Based on Deep Reinforcement Learning

Tie Chen , Hongxin Li *, Ying Cao and Zhifan Zhang

College of Electrical and New Energy, China Three Gorges University, Yichang 443002, China;
chent@ctgu.edu.cn (T.C.)

* Correspondence: 202108580021129@ctgu.edu.cn

Abstract: At present, substation operation ticket system is developed based on an expert system, which has some problems such as knowledge base redundancy, intelligence deficiency and automatic learning ability. To solve this problem, this paper proposes an operation sequence reasoning model based on the knowledge base of the Neo4j knowledge graph and DuelingDQN (Dueling Deep Q Network) algorithm. Firstly, the diagram structure model of substation main wiring was established using the Neo4j knowledge graph. Based on the diagram structure model, the operable equipment set of the operation task was searched to form the task space, action space and action selection model of DuelingDQN. The reward and punishment function was designed based on the “five defense” rules and the state change of equipment. Make DuelingDQN model and Neo4j model interact in real time, and automatically learn the operation sequence. The test results show that the method proposed in this paper can automatically deduce the correct operation steps under different wiring modes and realize the transfer within the station, which is of great significance to the intellectualization of the operation ticket system.

Keywords: substation; sequence of operations; deep reinforcement learning; Neo4j diagram knowledge base; inference automation



Citation: Chen, T.; Li, H.; Cao, Y.; Zhang, Z. Substation Operation Sequence Inference Model Based on Deep Reinforcement Learning. *Appl. Sci.* **2023**, *13*, 7360. <https://doi.org/10.3390/app13137360>

Academic Editor: Luis Javier Garcia Villalba

Received: 22 May 2023
Revised: 10 June 2023
Accepted: 19 June 2023
Published: 21 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Electric power enterprises are undergoing digital transformation, and artificial intelligence technology is playing an increasingly important role in distribution network scheduling, operation and maintenance. The switching operation of the substation is an important work in the dispatching operation of the distribution network. When the power network fails or is repaired, the load transfer and operation mode adjustment should be realized through the switching operation. Any maloperation may cause serious security accidents. Drawing up the operation ticket in advance and determining the operation steps can improve the correct rate of switching operation, thus avoiding maloperation. However, with the continuous expansion of the power system, the number of power equipment pieces involved has soared, the workload of employees to fill in the operation ticket, the quality requirements of employees grow increasingly larger, and the mistakes of ticketing personnel are inevitable. Therefore, it is necessary to develop an intelligent operation ticket generation system which can adapt to various electrical main wiring and improve the efficiency and accuracy of operation ticket generation.

An intelligent operating ticket system can automatically form operation steps. The current operating ticket system is implemented based on an expert system [1–5]. This type of method needs to build a knowledge base with a large number of operation rules in advance, and then deduce the knowledge base to obtain the operation steps. However, the knowledge base is highly dependent on the experience of the operator. Because the operation rules are highly related to the main wiring mode, equipment composition, initial state of the equipment, and whether it needs to be transferred, there will be different

operation rules under different conditions even if the wiring structure is the same. Therefore, the knowledge base has insufficient completeness, high redundancy and poor flexibility, and manual intervention must be carried out in the formation of operation steps.

Scholars have carried out a significant amount of research on simplifying rules and designing templates, and have made some achievements, but cannot solve the fundamental problem. Rough set, rule tree and other techniques simplify the operation rules, reduce the redundancy of rules and reduce the time required for inference [6,7]. Shu-chen Dong, Dongying Zhang et al. establishes typical interval template and specific main wiring template, solidifies its operation rules, and forms operation steps through rule matching [8,9]. These studies focus on specific main connections and topologies, and the rules need to be refined as new connections emerge.

Deep learning is widely used in power system because of its excellent ability in data analysis [10,11]; in order to shorten the reasoning process of operation tickets, some scholars use deep learning to guide the generation of operation tickets [12,13]. By learning the relationship between the environment and the operation sequence before and after the substation operation, a model that can match similar historical tickets according to the operation task and the initial state of the substation is obtained, and the generation of operation tickets is guided. Although the method has a very good accuracy, its essence does not break away from the artificial, thus not being able to achieve the realization of real intelligence. At the same time, the training of the model relies on a large number of historical operation tickets, which is still not friendly to the new plant station and the expansion part. Therefore, it is necessary to construct a universal and intelligent inference model of operation sequence.

Reverse brake operation can be regarded as a sequential decision-making process in uncertain scenarios, and its results are uncertain. Deep reinforcement learning has excellent performance in sequential decision-making problems due to its strong decision-making ability [14]. As a typical model of deep reinforcement learning, DuelingDQN has been widely applied to sequential decision problems such as demand response [15], energy management [16] and operation control [17] in power systems, and has achieved excellent results. Therefore, the DuelingDQN algorithm is considered for inference solving.

Compared with the model composed of relational database, the graph model is more suitable for describing the ontology model intuitively due to its advantages [18]. As a graph knowledge base development platform, based on its good dynamic updating ability and efficient query ability, Neo4j graph knowledge base is currently widely used in the research of distribution network data storage of power system, and has achieved good results compared with the relational database in updating and querying distribution network data [19,20]. The substation and the distribution network have great requirements for the integrity of topological information. At the same time, as the interactive environment of deep reinforcement learning, it needs to constantly update the received actions; therefore, it is appropriate to choose Neo4j to complete the construction of the substation knowledge base model.

In summary, this paper designs a substation operation sequence reasoning model framework based on deep reinforcement learning and combined with the Neo4j knowledge graph. Compared with mainstream methods, this model has stronger universality and intelligence. The main innovations are:

- (1) Based on the natural similarity between the power network and the graph network, the main wiring model of the substation is designed using the knowledge base of Neo4j graph. The model can respond to operation and update the state of the main wiring in real time, providing an interactive environment for deep reinforcement learning.
- (2) A task-state perception module is designed to identify the working state of the device from the knowledge network and obtain action space for the deep reinforcement learning model to make decisions.
- (3) The inference model is constructed using the deep reinforcement learning model, and the reward function and penalty function oriented to reverse switching operation are

designed according to the general rules of switching operation so as to deduce the sequence of operations.

2. Design of Substation Operation Ticket Inference Model

The reasoning model consists of four modules, including the main wiring module based on the knowledge base of Neo4j map, the task-state perception module, the operation sequence reasoning module based on deep reinforcement learning, and the action module.

- (1) The main wiring module converts the main wiring into the graph knowledge model, which is driven by the action module, receives the operation from the action module, changes the topological state of the main wiring, updates the state of the device, and sends the real-time state of the main wiring and the device to the task-state awareness module.
- (2) Task-state awareness module: Firstly, the operation task is analyzed, and the task device is obtained. Then, starting from the task device, the action space and associated device required by the operation are searched according to the real-time state sent by the main wiring module, and the action space and associated device state are sent to the operation sequence module for reasoning.
- (3) The operation sequence reasoning module firstly evaluates the state of associated devices through action-state evaluation, then obtains reasonable operations from the action space through deep reinforcement learning, and finally gradually forms a complete and correct operation sequence.
- (4) The action module mainly updates the knowledge base model of substation diagram with the operation action of the reasoning model, including node attributes and relations.

The structure and working mode of the model composed of the above four modules are shown in Figure 1.

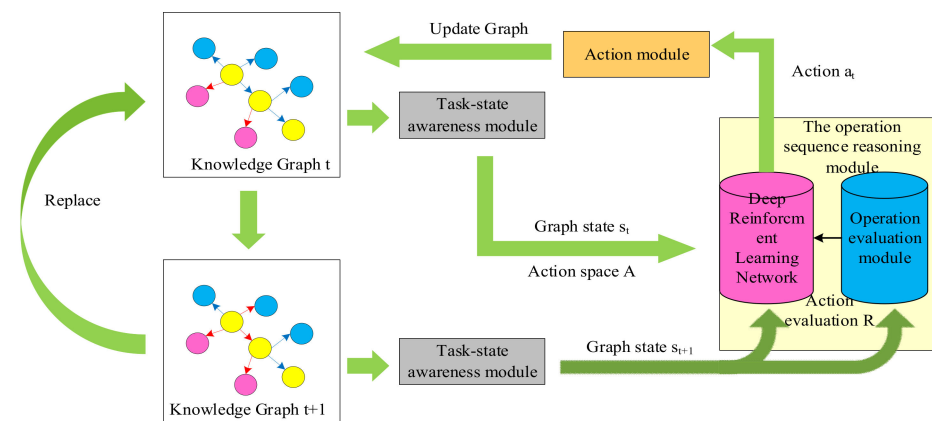


Figure 1. Structure and working mode of substation operation ticket inference model.

3. Diagram Model of the Main Wiring

The main electrical connection is topologically similar to the knowledge graph, and the knowledge graph has a good dynamic updating performance and high query efficiency. It can quickly synchronize and search for changes in the wiring mode and device status caused by operations. The essence of reinforcement learning is “trial and error”. In this paper, the knowledge base of the Neo4j diagram is used to build the substation main wiring model, and the model feeds back the state changes of the main wiring and equipment caused by the “trial and error operation”.

3.1. Modeling

The graph network model of the main connection is constructed by taking the device on the main connection as the entity, and the connection relationship between the devices

as the relationship between the entities; the status, number and device type of the entity are represented by attributes. Figure 2 shows the graph knowledge base model of a switch group in operation and its connected busbar, in which busbar B1 is connected to the incoming line.

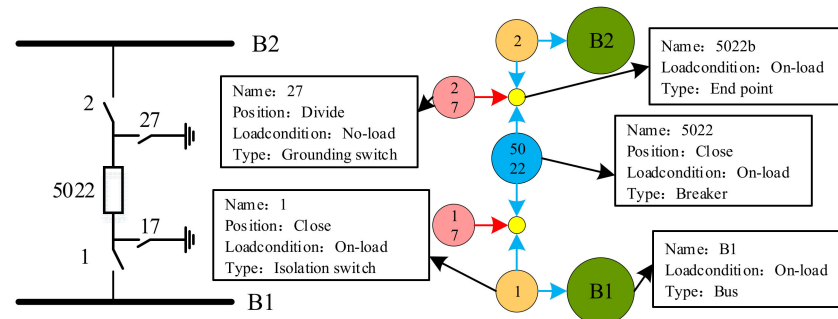


Figure 2. An example of a diagram model of the main wiring.

- (1) Entity. The reverse operation is to set the working state of the task equipment as the target state by operating the switch group. The switch group is set as an operable entity. The switch group includes the circuit breaker, isolation switch and grounding brake. Other devices such as busbars, transformers, incoming cables, and outgoing cables are set to inoperable entities. Introducing an “endpoint” entity ensures the topological integrity of the main wiring.
- (2) Entity attributes. Entity attributes describe the characteristics of entities in the graph model, including static attributes and dynamic attributes. Static attributes include device type, voltage level, operable or unavailable. Dynamic attributes include the following: on/off, off/on, maintenance/cold standby/hot standby/operation/transition/bad state.
- (3) The maintenance/cold standby/hot standby/operation/transition/bad state attribute of the entity is related to the shutdown state of the connected switch group. If the entity is a switchgear, its running state is determined by the state of the switching group in which it is located.
 - (i) Operation σ_1 : The circuit breaker and the isolation switch are closed, and the grounding switch is open.
 - (ii) Hot standby σ_2 : The circuit breaker and ground switch are open, and the isolation switch is closed.
 - (iii) Cold standby σ_3 : The circuit breaker, isolation switch and grounding brake are all open.
 - (iv) Maintenance σ_4 : The circuit breaker and isolation switch are open, and the grounding switch is closed.
 - (v) Transition σ_5 : It is in between four states. For example, the circuit breakers are open, but the isolation switch is not closed, or the circuit breaker and the isolation switch are all closed, but the grounding switch is not fully closed.
 - (vi) Bad state σ_6 : This state includes the following: (a) the isolation switch in the switch group operates but the circuit breaker does not operate; (b) repeated action of the same switch; (c) when the load is powered off during the operation when a power recovery path exists; (d) when two or more switch groups are in transition state at the same time.
- (4) Relationships. The relationship indicates the connection between the devices. The value can be connected or disconnected.

3.2. Rules for Model Updates

Actions from the action module change the close/divide on properties of the actionable entities, thereby changing the relationship between the entities. The Cypher statements

built into the Neo4j diagram repository are used to update the attributes of the entities, and the relationships between the entities are updated according to the device switch off/switch on attributes.

If and only if an entity can retrieve the power entity through the “connected” relation, determine that the current device is live; otherwise, there is no power. After the circuit breaker 5022 in Figure 2 is disconnected to enable the switch group to enter the hot standby state, the relationship between the devices and the device properties are shown in Figure 3.

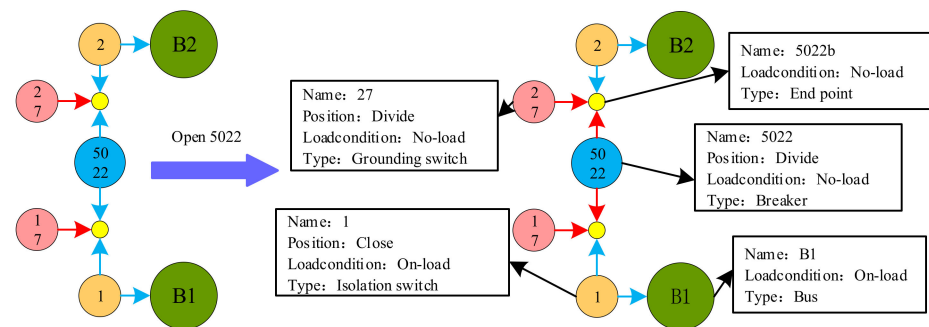


Figure 3. Example of updating rules of knowledge graph.

4. Task—State Awareness Module

The operation task involves only parts of the devices on the main wiring. If the complete main wiring is included into the action space of reinforcement learning, the solving efficiency will be greatly reduced. According to the operation task, determining the required switch group set to complete the operation can improve the solution efficiency. This model first determines the target according to the operation task, and then obtains the action space from the target.

The task is divided into target + state. The target represents the task object and the state represents the state to be achieved by the target, including operation, hot standby, cold standby and maintenance.

From the task objective, the Neo4j search algorithm was used to obtain the action space. The action space is composed of switch groups. When performing maintenance tasks, equipment power outage is inevitable. If the conditions exist, it is necessary to ensure that the user’s power supply is not interrupted. Therefore, the action space needs to include the task space and the diversion space. A task path is the set of switches required to complete a task regardless of the user’s power supply. The transfer space is a switch group that provides backup power to users to prevent power outages caused by the exit of the target devices.

4.1. Get Task Space

The search process is as follows:

- (1) Starting from the target device, search the power supply path connected to the non-switching device, and record the device set in the path.
- (2) Select circuit breakers and isolation switches in the path, and record the switchgear in the same path as a switch group.
- (3) Arrange the isolation switchgear in the switch group in order from near to far from the load.
- (4) Through the circuit breaker device, search the ground switch device with its common endpoint, and record the switch group of the circuit breaker.
- (5) If the task device is a non-switching device, search for the ground tool switch that is directly related to the task device and refer to it as the ground tool switch group of the task device.

4.2. Get the Transfer Space

The transfer space is related to the form of the main wiring, and the adequacy of the transfer needs to be judged. The basis of judgment is as follows:

- (1) Whether there is a path for restoring power supply to the non-switching devices supplied through the path of the task device, that is, whether power devices such as a live bus or an incoming line can be retrieved from the path.
- (2) Whether the path of the task device is the only power supply path of the load device.

When the above two conditions are met, the transfer space can be judged to be available. The search process is as follows:

- (1) Starting from the non-switching devices affected by the device, the attributes of the device nodes and the topology of the graph are used to search for the path through which devices such as live bus can be retrieved.
- (2) Pick the path. Select the transfer path with circuit breaker to ensure the safe operation of restored power supply. As a priority, select the path on the same side of the power supply device to ensure the stability of the power supply quality. Finally, record the device collection.
- (3) Pick out the circuit breakers and isolation switchgear in the path and record them as a switch group.
- (4) In contrast to the tripping space, the isolation switchgear in the switching group is arranged in order from far to near the load.
- (5) Through the circuit breaker device, search the ground switch device with its common endpoint, and record the switch group of the circuit breaker.

The complete action space includes task space and transfer space, and the specific operation process is shown in Figure 4.

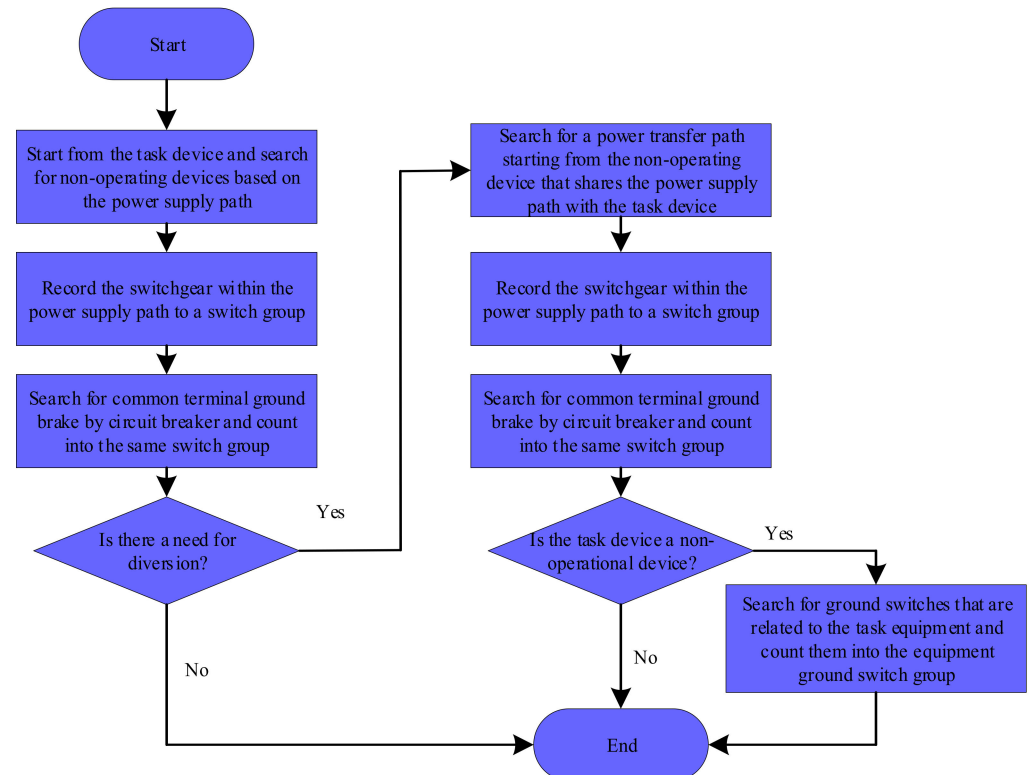


Figure 4. Flowchart of action space search.

4.3. Identifying the Status of the Target Device

Update the maintenance/cold standby/hot standby/operation/transition/bad status attributes of the target device based on the switch status of the target device in the task

space. If the target is a switching device, its operating state is determined by the state of the switching group in which it resides. As shown in Figure 2, bus B1 and circuit breaker 5022 are in operation.

5. Action Module

The action module converts the “trial and error” actions formed by reinforcement learning into operations and changes the close/divide on properties of workspace switches.

When the switchgear is operated, the state transition is based on the state before the operation. For example, if it is closed before the action, it is open after the action, and the action is “open” + the name of the switch device; otherwise, it is “close” + the name of the switch device.

- (1) After the action of the switching device, change the opening and closing attributes of the action device itself. Before the action, “close” is changed to “divide”, and before the action, “divide” is changed to “close”.
- (2) When the switch device is finished, the properties of the relationship between it change. For example, when the switch device is switched from close to open, the relationship changes from “connected” to “disconnected”.
- (3) After the updating of relation and attribute, each entity in the substation model is retrieved through “Cypher” statement to judge whether the entity in the model can retrieve the incoming line entity only through the “connected” relation. If it can, the live attribute is updated to “On-load”; otherwise, it is updated to “No-load”.

6. Operation Sequence Inference Module

This module is a sequential reasoning model built based on deep reinforcement learning. Under the constraint of basic rules, it operates the graph network model of the main connection through “trial and error”, evaluates the operation steps and the state of the target device, and forms the optimal operation sequence through continuous iteration.

6.1. Deep Reinforcement Learning Module

6.1.1. Deep Q Network

Reinforcement learning refers to the learning process in which the agent obtains the maximum cumulative reward from the environment by learning from the environment state to the action strategy set. In essence, the reversing operation is regarded as a Markov decision process, which can be expressed as a quadruple (S, A, R, π) , where:

- (1) S represents the set of all environmental states in the decision-making process, and $s_t \in S$ represents the perceived state of the agent in the environment during time t . S is the close/divide state set of the action space switch group.
- (2) A is the set of all executable actions of the agent, and $a_t \in A$ represents the actions taken by the agent in time t . In order to simplify the size of the motion space, the size of A is the size of the motion space. The action content is to select the switch device from the action space for the on-off state attribute transformation.
- (3) R represents the reward function, and $r_t \in R(s_t, a_t)$ represents the immediate reward that the agent gets by performing a_t action in state s_t .
- (4) π is the policy set of the agent, representing the mapping from state space S to action space A .

The agent iteratively updates the action value function Q through the feedback information from the environment:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right] \quad (1)$$

where Q is the value function of state action, which refers to the value of taking action a_t to the environment under the current state s_t . t is the current period; α is the learning rate. γ is the future decay rate of Q value at present. a_{t+1} is the action that can be performed in

state s_{t+1} . The algorithm approximates the optimal Q function according to the next state information and reward obtained after the agent interacts with the environment.

On the basis of Q learning, deep Q learning (DQN) approximates value function Q by constructing function Q composed of parameter ω , i.e.,

$$Q'(s, a, \omega) \approx Q(s, a) \quad (2)$$

At the same time, in order to reduce the correlation of samples and improve the stability of training, DQN introduced the experience replay mechanism and the parameter-freezing mechanism. The empirical replay mechanism means that the algorithm will create an experience replay library, and when the experience of the replay library reaches a certain number, it will extract the experience for training. The specific structure of DQN network is shown in Figure 5.

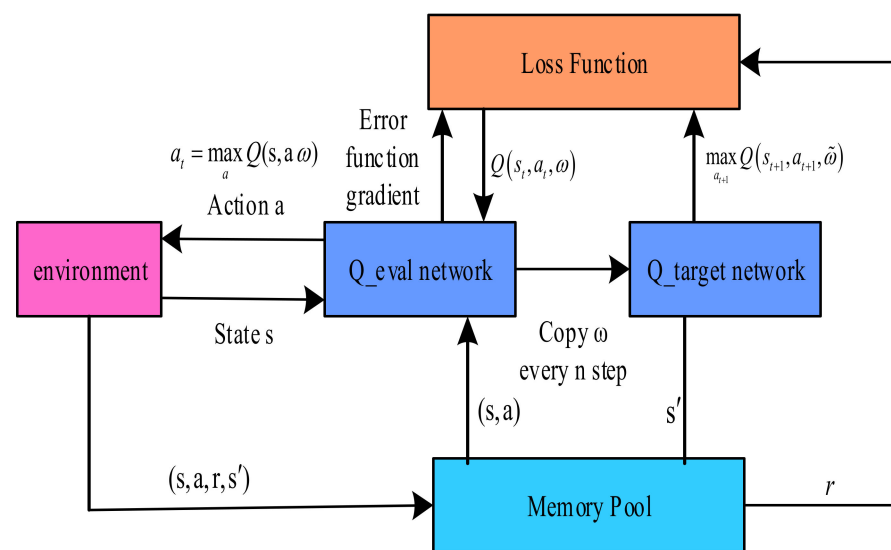


Figure 5. Network model of DQN.

6.1.2. Dueling Deep Q Network

Dueling Deep Q Learning (DuelingDQN), a variant of the DQN model, is adopted as the training model in this paper. DuelingDQN is the introduction of duel networks based on DQN. A duel network divides a Q network into two parts, value function $V(s, \omega, \beta)$ and advantage function $A(s, a, \omega, \alpha)$, so as to improve the generalization of deep reinforcement learning network. The value function V is the evaluation of the current state of the action space, and the action function A is the evaluation of the action taken in the current state. The final result of the network consists of two parts:

$$Q(s, a; \omega, \alpha, \beta) = V(s; \omega, \beta) + A(s, a; \omega, \alpha) \quad (3)$$

where ω is the common network parameter of the value function and the advantage function, α is the network parameter of the independent part of the value function, β is the network parameter of the independent part of the advantage function. The Dueling network structure is shown in Figure 6.

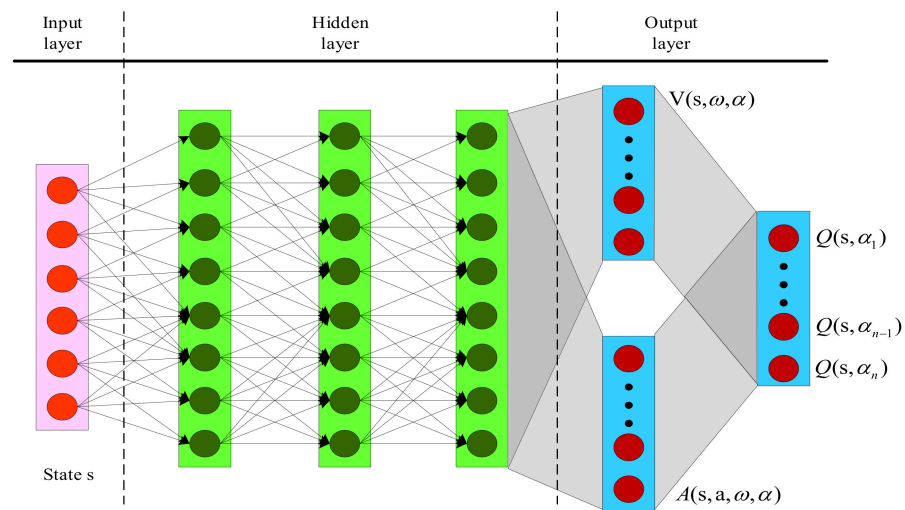


Figure 6. Model of the duel network.

6.2. Procedure Real-Time Evaluation Module

The evaluation of operational steps is mainly accomplished through the reward and punishment function. In the value-based Dueling DQN model, the reward and punishment value is the immediate feedback obtained by the agent after the action selection. In order to set the reward and punishment function reasonably, this paper sets the operation evaluation function $O(s)$. According to the requirements in the process of operation ticket formation, a reward function $R(s)$ is set in this paper to guide the model towards the destination direction, and a penalty function $P(s)$ is set in order to avoid maloperations in the operation sequence or make the environment enter a bad state. The environment state function is the accumulation of reward function and punishment function, as follows:

$$O(s) = \sum R(s) + \sum P(s) \quad (4)$$

The task of the reverse operation is to realize the operation of the conversion between the four basic operating states of the task equipment. According to the updating rules of the operating state of the task equipment, the reward function $R(s)$ can be divided into the switching device state reward $R_S(s)$ and the task device state $R_D(s)$ reward, where:

$$R_S(s) = \begin{cases} 1 & S_s = \sigma_1, \sigma_2, \sigma_3, \sigma_4 \\ 0.5 & else \end{cases} \quad (5)$$

$$R_D(s) = \begin{cases} 0.5 & S_D = \sigma_1, \sigma_2, \sigma_3, \sigma_4 \\ 1 & S_D = S_{aim} \end{cases} \quad (6)$$

$$R(s) = R_S(s) + R_D(s) \quad (7)$$

In the formula, S_s represents the running state of the switch group, S_D represents the running state of the task device, and S_{aim} represents the target state of the task device.

For the penalty function $P(s)$, according to the safety requirements in the switching operation, this paper divides the penalty function $P_I(s)$ in the bad state and $P_O(s)$ in the wrong operation sequence, where:

$$\begin{cases} P_I(s) = -2 & S_s \text{ or } S_D = \sigma_6 \\ P_O(s) = -1.5 \end{cases} \quad (8)$$

The wrong operation sequence refers to that the isolation switches in the same switch group do not operate in the sequence of load side followed by power side in the case of power failure and power side followed by load side in the case of power supply.

7. Numerical Example Verification

Figure 7 shows the main wiring diagram of a substation. Bus WBI and bus WBII are running in sections, the bypass circuit breaker QFP is in maintenance state, and the bypass busbar WP is not live. Select transformer T1, outgoing circuit breaker QF3, and bus WBI for maintenance operation.

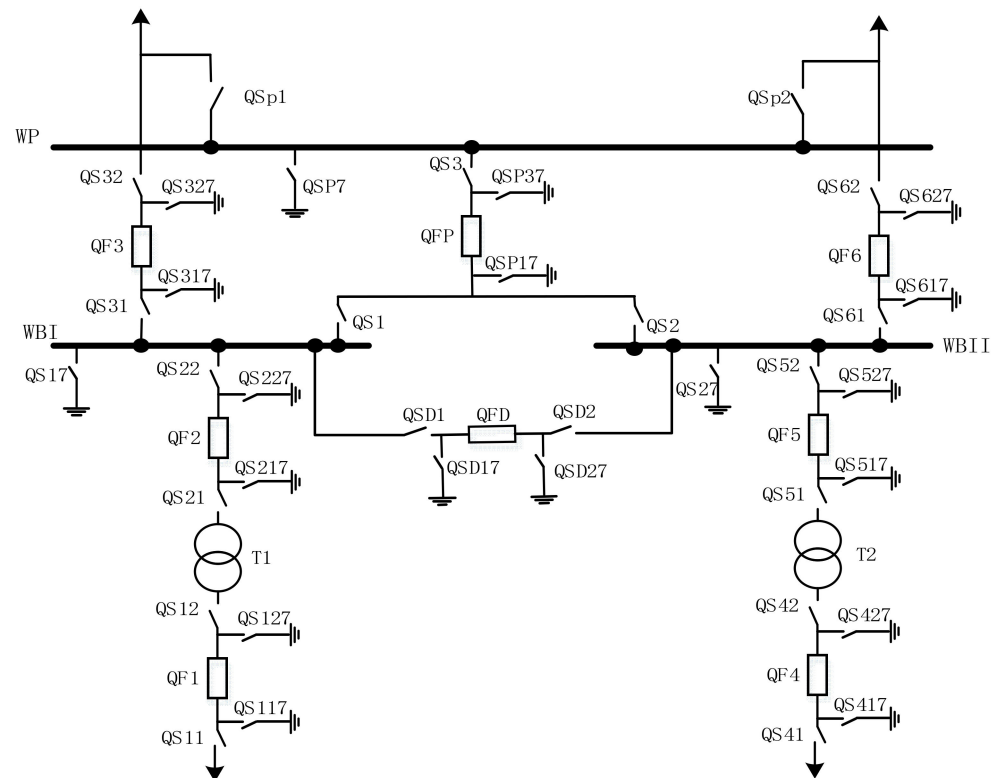


Figure 7. Main wiring of a substation.

In this paper, Python is used as the underlying language of program compilation, and Tensorflow open source software library is used to form a deep reinforcement network, and Py2neo toolkit is used to complete the linkage work between Python platform and Neo4j knowledge graph. The specific versions of the above tools are Neo4j4.4.5, Python3.6, Py2neo2021.2.3 and Tensorflow2.0.0.

According to the task awareness module, the switching operation path and power supply recovery path obtained in the above three cases are shown in Figure 8.

Figure 8, from top to bottom, respectively, represents the action space of the outgoing circuit breaker QF3, transformer T1 and busbar WBI when they enter the maintenance state, in which the switchgear in the red dotted line frame is the action space of task equipment and the switchgear in the blue dotted line is the action space of power transfer.

As can be seen from the figure, the motion space search model accurately identifies the path of mission equipment withdrawal and power supply recovery. When T1 enters the maintenance state, the model chooses the busbar switch group where the circuit breaker QFD is located as the power supply recovery path. When QF3 and WBI enter the maintenance state, the path through the bypass bus is selected as the power supply recovery path to meet the requirement that circuit breakers exist in the path. At the same time, comparing the power supply recovery paths of QF3 and WBI, the model chooses the WBI bus on the

same side as the load as the power supply when QF3 is undergoing maintenance, which meets the requirement of giving priority to the power supply on the same side.

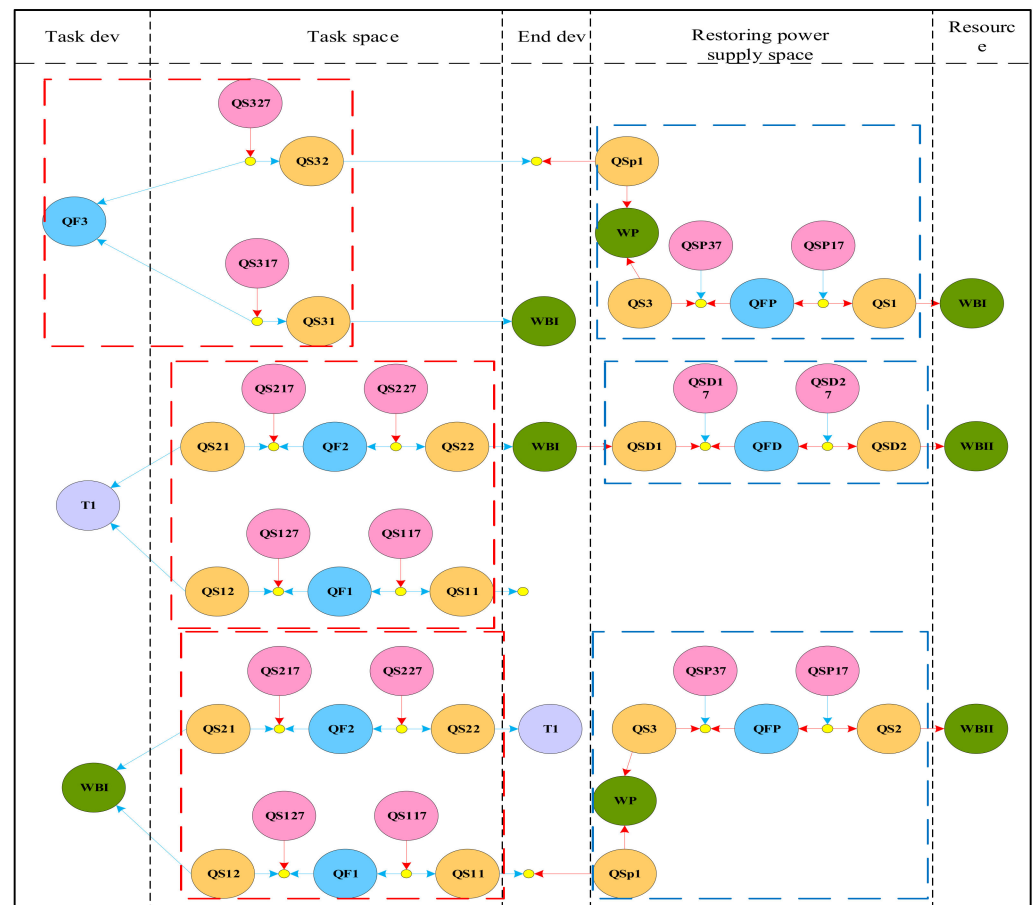


Figure 8. Action space acquisition graph.

After acquiring action space, the average reward curve obtained after DuelingDQN and DQN model training is shown in Figure 9.

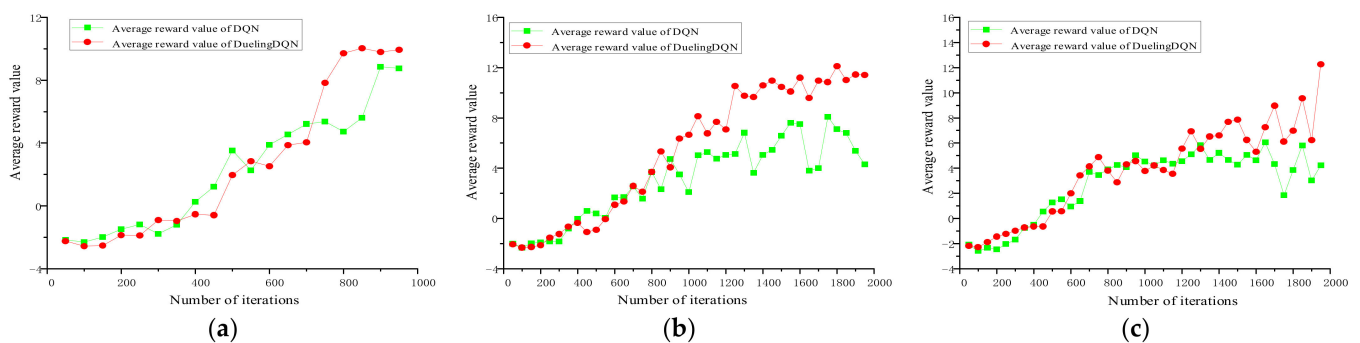


Figure 9. Reward curve of deep reinforcement learning training. (a) the training curves of outgoing circuit breaker QF3; (b) the training curves of transformer T1. (c) the training curves of busbar WBI.

In Figure 9a–c are, respectively, the training curves of outgoing circuit breaker QF3, transformer T1 and busbar WBI. As can be seen from the figure, DuelingDQN obtains rewards significantly faster than DQN in the same operating space. Meanwhile, with the increase in operation space, the acquisition efficiency of reward value decreases somewhat, but DuelingDQN's performance is still better than that of the DQN algorithm.

After the training of reinforcement learning model, all optimal operation sequences obtained are as shown in Tables 1–3.

Table 1. Optimal operating sequence of outlet circuit breaker QF3 maintenance.

The Optimal Operation Sequence when QF3 Exits the Running Overhaul	
Step 1	Open grounding switch QSP17/QSP37
Step 2	Open grounding switch QSP37/QSP17
Step 3	Close disconnecting switch QS1
Step 4	Close disconnecting switch QS3
Step 5	Close disconnecting switch QSp1
Step 6	Close breaker QFP
Step 7	Open breaker QF3
Step 8	Open disconnecting switch QS32
Step 9	Open disconnecting switch QS31
Step 10	Close grounding switch QS317/QS327
Step 11	Close grounding switch QS327/QS117

Table 2. Optimal operating sequence for transformer T1 overhaul.

The Optimal Operation Sequence when T1 Exits the Running Overhaul	
Step 1	Open grounding switch QSD17/QSD27
Step 2	Open grounding switch QSD27/QSD17
Step 3	Close disconnecting switch QSD2
Step 4	Close disconnecting switch QSD1
Step 5	Close breaker QFD
Step 6	Open breaker QF2/QF1
Step 7	Open breaker QF1/QF2
Step 8	Open disconnecting switch QS12/QS22
Step 9	Open disconnecting switch QS11/QS21
Step 10	Open disconnecting switch QS22/QS12
Step 11	Open disconnecting switch QS21/QS11
Step 12	Close grounding switch QS127/QS117/QS227/QS217
Step 13	Close grounding switch QS117/QS127/QS217/QS227
Step 14	Close grounding switch QS227/QS217/QS127/QS117
Step 15	Close grounding switch QS217/QS227/QS117/QS127

Table 3. Optimal operating sequence for bus WBI overhaul.

The Optimal Operation Sequence when WBI Exits the Running Overhaul	
Step 1	Open grounding switch QSP17/QSP27
Step 2	Open grounding switch QSP27/QSP17
Step 3	Close disconnecting switch QS2
Step 4	Close disconnecting switch QS3
Step 5	Close disconnecting switch QSp1
Step 6	Close breaker QFP
Step 7	Open breaker QF3/QF1
Step 8	Open breaker QF1/QF3
Step 9	Open disconnecting switch QS12/QS32
Step 10	Open disconnecting switch QS11/QS31
Step 11	Open disconnecting switch QS32/QS12
Step 12	Open disconnecting switch QS31/QS11
Step 13	Close grounding switch QS227/QS217/QS327/QS317
Step 14	Close grounding switch QS217/QS227/QS317/QS327
Step 15	Close grounding switch QS327/QS317/QS227/QS217
Step 16	Close grounding switch QS317/QS327/QS217/QS227
Step 17	Close grounding switch QS17

As can be seen from the sequence of three operations in the above list:

- (1) The equipment and steps of three maintenance operations are different, but the accuracy of 100% is maintained.
- (2) The three maintenance are connected to the transfer channel first, to ensure that the load will not outage, and then maintenance. For example, during the maintenance of QF3, firstly power the line with QFP through steps 1–6, and then overhaul QF3 through steps 7–11. To repair T1, first use steps 1–5 to supply power to WBI, and then use steps 6–15 to repair T1. To overhaul the WBI, first power the line with QFP through steps 1–6, and then overhaul the WBI through steps 7–17.
- (3) The operation of the circuit breaker and disconnecting switch complies with the rule that the circuit breaker is first on and then off, and the busbar disconnecting switch is first on and then off to ensure safe operation.

It can be seen that it is feasible to generate operation sequence using the model described in this paper, and the obtained results are also credible.

8. Conclusions

In this paper, a reasoning model of substation switching operation steps is proposed. First, the main wiring is converted into a Neo4j model, and the task space is then obtained from the task awareness model. Then, DuelingDQN automatic reasoning operation steps are used to update the Neo4j model through the action model, so that the operating steps can be obtained automatically under the condition of no interference. The specific conclusions are as follows:

- (1) Using Neo4j to build a graph structure model containing the main wiring structure and status information, and using its own path search tool, can improve the search efficiency of information such as device connection relationship, device status, transfer path, outage path, load distance, etc.
- (2) The Neo4j diagram structure model is used to search the task state space, and then the required environment space and action space of DuelingDQN are constructed according to the task space, which can reduce the required resources of DuelingDQN and improve the computational efficiency of DuelingDQN.
- (3) The DuelingDQN model and Neo4j model interact in real time, and the operation rules of “five defense” and isolation switch are used to automatically complete the reasoning of operation sequence, avoiding the problem of rule redundancy in the expert system, ensuring the universality of reasoning, and giving play to the decision-making ability of reinforcing learning to eliminate human intervention.

It can be observed from the results of the example that the method described in this paper can actively show excellent reasoning ability in the face of different main wiring and different operation tasks, and the operation steps obtained also meet the requirements of substation safety production, which has important significance for the intelligent operation ticket system.

The reward acquisition curve of deep reinforcement learning model shows that the rewards efficiency decreased significantly with increase in the operation space. The efficiency of deep reinforcement learning may not be guaranteed when faced with more complex operations in a distribution network within a large operation space. Transfer operation space and equipment drop back are independent of each other. The next study aims to train the transfer task and the equipment drop back task separately, and adopt the hierarchical learning method to ensure the correctness of the operation sequence and improve the ability of the model to face complex situations.

Author Contributions: Method, T.C. and H.L.; simulation and verification, H.L. and Y.C.; writing, H.L.; review, T.C. and Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (51907104), The Opening Fund of Hubei Province Key Laboratory of Operation, and Control of Cascade Hydropower Station (2019KJX08).

Data Availability Statement: All the data supporting the reported results have been included in this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, Y.; Zhang, Q.; Zheng, D.; Xu, P.; Wu, B. On operation ticket expert system for a substation using object oriented technique. In Proceedings of the 2008 27th Chinese Control Conference, Kunming, China, 16–18 July 2008; pp. 2–5.
2. Zhang, F.; Guo, B. Discussion on the combination of intelligent operation ticket system and programmed operation system in substation. *Electr. Power Autom. Equip.* **2010**, *30*, 117–120.
3. Chen, K.; Zhang, B.; Wu, W.; Sun, H. An intelligent checking system for power system operation tickets. In Proceedings of the 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), Weihai, China, 6–9 July 2011; pp. 757–762.
4. Jiang, C.; Zhou, H.; Deng, Q. A new system of operation ticket for generation and misoperation prevention in smart distribution network. In Proceedings of the 2011 International Conference on Electric Information and Control Engineering, Wuhan, China, 15–17 April 2011; pp. 60–64.
5. Xu, J.; Zhao, J.; Rao, M.; Shi, J.; Ren, Y. Design of Regional Power Grid Dispatching Operation Ticket System based on SCADA System. *Power Syst. Prot. Control.* **2010**, *38*, 104–107, 112.
6. Zhang, X.; Cheng, X.-T.; Zhao, D.-M. Power Grid Operation Ticket Rule Extraction based on Rough Set. *Power Grid Technol.* **2014**, *38*, 1600–1605.
7. Luo, P.; Xia, W.; Lin, J.-K.; Zheng, W.-H. A New Method of Automatic invoicing Based on Rule Tree. *Electr. Power* **2011**, *44*, 79–84.
8. Dong, S.-C.; Cheng, J.; Peng, B.; Chen, X. Automatic Generation method of Graph-Library-Rule-operation ticket based on Interval Model. *Autom. Electr. Power Syst.* **2015**, *39*, 84–89.
9. Zhang, D.; Yang, J.; Huang, J.; Wang, W. General method of Automatic Generation of operation ticket task. *Autom. Electr. Power Syst.* **2018**, *42*, 169–174.
10. Wang, R.; Li, C.; Fu, W.; Tang, G. Deep Learning Method Based on Gated Recurrent Unit and Variational Mode Decomposition for Short-Term Wind Power Interval Prediction. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 3814–3827. [[CrossRef](#)] [[PubMed](#)]
11. Fang, P.; Fu, W.; Wang, K.; Xiong, D.; Zhang, K. A compositive architecture coupling outlier correction, EWT, nonlinear Volterra multi-model fusion with multi-objective optimization for short-term wind speed forecasting. *Appl. Energy* **2022**, *307*, 118191. [[CrossRef](#)]
12. Ren, S.; Lv, X.; Sun, K.; Xu, Z.; Long, T. Chinese Operation Ticket Automatic Verification Method Based on CNN-BiGRU-Attention in Intelligent Substations. In Proceedings of the 2022 IEEE International Conference on High Voltage Engineering and Applications (ICHVE), Chongqing, China, 25–29 September 2022; pp. 1–4.
13. Liu, T.; Li, S.; Gu, X.; Wang, T.; Lu, P.; Cao, X.; Yang, X.; Wang, W.; Lv, H.; Feng, C. Historical Similar Ticket Matching and Extraction used for Power Grid Maintenance Work Ticket Decision Making. In Proceedings of the 2019 3rd International Conference on Data Science and Business Analytics (ICDSBA), Istanbul, Turkey, 11–12 October 2019; pp. 302–307.
14. Zhang, Z.; Zhang, D.; Qiu, R.C. Deep reinforcement learning for power system applications: An overview. *CSEE J. Power Energy Syst.* **2019**, *6*, 213–225.
15. Wei, T.; Chu, X.; Yang, D.; Ma, H. Power balance control of RES integrated power system by deep reinforcement learning with optimized utilization rate of renewable energy. *Energy Rep.* **2022**, *8*, 544–553. [[CrossRef](#)]
16. Damjanović, I.; Pavić, I.; Puljiz, M.; Brcic, M. Deep Reinforcement Learning-Based Approach for Autonomous Power Flow Control Using Only Topology Changes. *Energies* **2022**, *15*, 6920. [[CrossRef](#)]
17. Zhang, Y.; Ai, Q.; Li, Z. Intelligent demand response resource trading using deep reinforcement learning. *CSEE J. Power Energy Syst.* **2021**, *1*–10. [[CrossRef](#)]
18. Elbattah, M.; Roushdy, M.; Aref, M.; Salem, A.B.M. Large-scale ontology storage and query using graph database-oriented approach: The case of Freebase. In Proceedings of the 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 12–14 December 2015; pp. 39–43.
19. Jiang, W.; Wang, M.; Chen, J.; Liu, J.; Pu, S.; Xu, Z. Reliability calculation of Distribution network based on Neo4j graph database. *Autom. Electr. Power Syst.* **2022**, *46*, 104–111.
20. Shi, P.; Fan, G.; Li, S.; Kou, D. Big data storage technology for smart distribution grid based on neo4j graph database. In Proceedings of the 2021 IEEE 4th International Conference on Electronics Technology (ICET), Chengdu, China, 7–10 May 2021; pp. 441–445.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.