Generation Method of Power Network Security Defense Strategy Based on Markov Decision Process

Wang Yang[1], Liu Dong[2,*], Wang Dong[1] and Xu Chun[3]

[1]*Xinjiang Normal College, Xinjiang, China*
[2]*Xinjiang Ploytechnical College, Xinjiang, China*
[3]*Xinjiang University of Finance and Economics, Xinjiang, China*
E-mail: 623044261@qq.com
[*]*Corresponding Author*

Received 21 April 2021; Accepted 11 May 2021;
Publication 13 July 2021

Abstract

Aiming at the problem that the current generation method of power network security defense strategy ignores the dependency relationship between nodes, resulting in closed-loop attack graph, which makes the defense strategy not generate attack path, resulting in poor defense effect and long generation response time of power network security defense strategy, a generation method of power network security defense strategy based on Markov decision process is proposed. Based on the generation of network attack and defense diagram, the paper describes the state change of attack network by using Markov decision-making process correlation principle, introduces discount factor, calculates the income value of attack and defense game process, constructs the evolutionary game model of attack and defense, solves the objective function according to the dynamic programming theory, obtains the optimal strategy set and outputs the final results, and generates the power network security defense strategy. The experimental results show that the

Distributed Generation & Alternative Energy Journal, Vol. 36_3, 287–300.
doi: 10.13052/dgaej2156-3306.3635

proposed method has good defense effect and can effectively shorten the generation response time of power network security defense strategy.

Keywords: Markov decision, power network security, defense strategy, offense and defense game.

1 Introduction

The network has penetrated into every aspect of people's life. While bringing great convenience to people, the network also puts forward the network security problem worth paying attention to. In the increasingly complex network security environment, there are a certain degree of hidden dangers in the network infrastructure such as power grid control and monitoring. With the continuous occurrence of power network intrusion and malicious attacks in the world, the network security problems faced by power system have attracted people's attention. Power network security is mainly reflected in the main form of network attack and defense. The main idea of network attack and defense is to strengthen the power network defense to ensure the network from infringement or minimize the loss caused by the attack. Therefore, it is very important to generate the optimal power network security defense strategy according to the network attack and defense state to ensure the safe operation of power network. Attack graph is a basic way to solve the problem of power network security. When the traditional method uses attack graph to generate power network security defense strategy, it is limited by the dependency relationship between nodes in the network, resulting in the closed loop of attack graph [1]. When eliminating the closed loop of attack graph, the final generated defense strategy will ignore some important attack paths, which will affect the defense effect and response time of power network security defense strategy.

Markov decision process is the DE facto standard method for sequential decision making, which has the advantages of allowing online solutions, allowing approximate solutions based on computational resources, and allowing numerical measurements of the strategic quality and learning effect of decision theory [2]. Adding Markov decision process into the generation process of power network security defense strategy can improve the effectiveness and applicability of the strategy generation method. Therefore, in order to improve the defense effect of power network security defense strategy and shorten the response time of defense strategy generation, according to the

above analysis, this paper will study the generation method of power network security defense strategy based on Markov decision process.

2 Generation Method of Power Network Security Defense Strategy Based on Markov Decision Process

2.1 Generate Network Attack and Defense Diagrams

The network attack and defense graph is a directed graph, which contains the access nodes and defense cost nodes respectively. One kind of directed edge represents the transition of network state or the attacker's attack action under the attacker's attack, and the other type represents the defense strategy adopted for the corresponding attack action [3]. The node information set includes the initial state of the power network node, the state of the attacker searching for the target node, the state information of the intermediate node in the attack and defense graph, and the state set of the power network defense behavior. The attack and defense map is defined as binary group (N, E), wherein N is the set of node information, the nodes in the attack and defense graph correspond to the security state of the power network nodes. E is the information set of the edge of the attack and defense diagram, represents the transition relationship between states in power network.

According to the above definition of network attack and defense chart, the correlation algorithm is used to generate the attack and defense chart. The attacker takes the host as the initial location to attack the target network. By looking for the security loopholes in the power network and the connection relationship of nodes in the network, the jump attack is carried out. Before the network attack and defense chart is generated, the host information of the network, network connection relationship, network vulnerability attack rules exploited by the attacker and the target host of the attacker are all known. By attacking the security holes in the network, attackers gain their privileges on the attacked host [4]. Determine if the host privilege obtained by the current attack has been obtained, and if it is determined to be a duplicate activity, stop the attack. After the end of one attack, the attacker can judge whether the atomic attack can achieve the enhancement of attack privilege and get the minimum path represented by permission sequence. The final attack and defense map can be obtained by matching the corresponding defense behaviors according to the generated attack path. Using the generated network attack and defense map, the game model of attack and defense based on Markov decision process is constructed.

2.2 Build Attack Defense Game Model

The process of network attack and defense is the evolutionary game between the attacker and the attacked, so the network attack and defense behavior will reach a stable state within a certain period of time. The attacker's attack on the power network is a multi-stage and jumping attack. By using the relevant principles of Markov decision process, the change of network state when the attacker attacks the power network is described, and the attack defense game model based on Markov decision process is constructed.

According to the relevant definition of evolutionary game, if the probability of the defender chooses the network security defense strategy DS_n is q_n, and the probability of the attacker chooses the attack strategy AS_i is p_m. In the process of power network security attack and defense game, the defender chooses the corresponding strategy to calculate the benefits of the game as follows [5]:

$$\begin{cases} U_{DS_1} = p_1 b_{11} + p_2 b_{21} + \cdots + p_m b_{m1} \\ U_{DS_n} = p_1 b_{1n} + p_2 b_{2n} + \cdots + p_m b_{mn} \\ \bar{U}_D = q_1 U_{DS_1} + q_2 U_{DS_2} + \cdots + q_n U_{DS_n} \end{cases} \tag{1}$$

In formula (1), U_{DS_1} is the expected revenue when the defender adopts the corresponding defense strategy; b_{mn} is the game strategy benefit of the defender; $\bar{U}_D$ is the average defense gain. According to the above formula, the game gains of the attacker can be obtained. After determining the evolution game process of single-stage Markov decision, a broken line factor is introduced to construct a multi-stage Markov decision evolution game model. If the total number of stages in the multi-stage game is T, and the value interval of discount factor ξ is [0,1], its value is determined by the relationship between the current game stage returns and the initial game stage returns [6]. Therefore, the objective function of the multi-stage Markov decision evolutionary game model is as follows:

$$\begin{cases} R_D^k(S_0^k, S_k) = U_D^k + \sum\limits_{e,h \in [k,T]} \xi^h \eta(S_h|S_e) R_D^h(S_0^h, S_h) \\ R_A^k(S_0^k, S_k) = U_A^k + \sum\limits_{e,h \in [k,T]} \xi^h \eta(S_h|S_e) R_A^h(S_0^h, S_h) \end{cases} \tag{2}$$

In formula (2), U_A^k is the return function of the attacker in the kth offensive and defensive game stage. U_D^k is the return function of the defender in the kth offense and defense game stages; η is the transition probability of

network security state; S_0^h is the transition probability of network security state; S_h is the security state set of the network [7]. The discount factor is introduced to calculate the profit of both sides in the future offensive and defensive game. The two sides of the game make their objective function reach the maximum value and get the best attack and defense strategy. After constructing the attack and defense game model based on Markov decision process, the optimal network security defense strategy is obtained by solving the model.

2.3 Generation of Power Network Security Defense Strategy

The attack defense game model based on Markov decision process established above transforms the problem of network security defense strategy selection into the dynamic programming problem with the highest profit as shown in the following formula [8].

$$\begin{cases} \max R_D^k(S_0^k, S_k) \\ \max R_A^k(S_0^k, S_k) \\ N_k(t) + I_k(t) + R_k(t) + M_k(t) = Q \end{cases} \tag{3}$$

In formula (3), $N_k(t)$ is the number of normal nodes in the kth game stages; $I_k(t)$ is the number of normal nodes infected after the failure of defense strategy in kth game stage; $R_k(t)$ represents the number of nodes identified and cleared in the kth stage; $M_k(t)$ is the number of damaged nodes unable to work normally in the kth stage; Q is the total number of nodes in the network [9]. Follow the following steps to solve the equation shown in formula (3) [10]:

(1) Initialize the relevant parameters of the offense and defense game model based on Markov decision process, and the defense behavior space and attack behavior space of power network are constructed;
(2) The security state sets S_0^h and S_h of each phase of power network are constructed.
(3) Initialize the network security state transition probability η of each game stage and the relevant constant coefficient;
(4) Start from the first game stage to calculate the game benefits of the current stage;
(5) Use discount factor to calculate discount income. According to the principle of dynamic programming, the optimal strategy set is obtained by solving the objective function, and the final result is output.

The defender strategy in the power network attack and defense strategy set obtained according to the above process is the best network security defense strategy. So far, this paper completes the research on the generation method of power network security defense strategy based on Markov decision process.

3 Simulation Experiment

In this section, the simulation experiments are carried out to verify the security defense strategy generation method based on Markov decision-making process, and the experimental data results are analyzed, and the corresponding experimental conclusions are obtained.

3.1 Experiment Content

In this simulation experiment, the generation method of power network security defense strategy based on Markov decision process is compared with the traditional generation method of power network security defense strategy, and relevant verification steps are completed. Through the experimental comparison, the actual application effect of the two defense strategy generation methods can be intuitively, truly and effectively compared to verify the feasibility and reliability of the research method in this paper.

The comparison indexes of the experiment are the power network security defense strategy generated by the power network security defense strategy generation method, the defense cost of the power network, and the response time of the two security defense strategy generation methods and the computing space occupied by the strategy reconfiguration under the condition of the power network attack parameters changing.

3.2 Experimental Environment and Process

In this paper, the simulation experiment is carried out in the experimental network with topology structure as shown in Figure 1. In the power network, the CPU of the server is a four core processor with 64g memory, and its operating system is Linux system.

During the experiment, in order to ensure the authenticity and effectiveness of the experimental results, the interference of other factors on the experimental data collection was reduced, and the data with large deviation was removed in the experimental data processing stage. In the experiment,

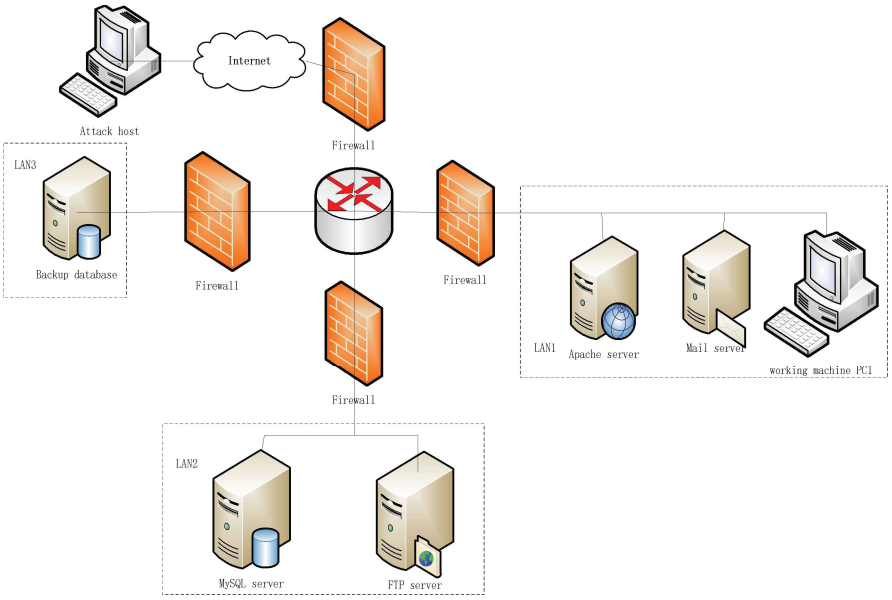

Figure 1 Schematic diagram of experimental network topology.

SSFNet simulation tool was used to set different network attack parameters and initial state by using the relevant functions in the tool, so as to simulate different network attack and defense scenarios in the experimental network. The specific parameters of attack strategy set in ssfnet tool during the experiment are shown in Table 1.

The defense strategy generation method studied in this paper is denoted as the experimental group and the traditional security strategy generation method as the control group. In the above experimental network, two generation methods of power network security defense strategy generate defense strategy according to the attack situation. In the first set of experiments, the defense cost of the actual network defense against the attack is taken as the experimental reference data. This paper quantifies the defense cost of the two security defense strategies when they are applied to the network, compares the quantified value of the two strategies with the reference value, and intuitively compares the defense effect of the two defense strategies. In the second set of experiments, the response time of the defense strategy generation method and the spatial data occupied by the policy change were collected when the attack strategy in the change simulation tool was

Table 1 Parameters of power network attack parameter strategy set

The Serial Number	Power Network Attack Code	Attack Power	Attack Types	Attack Utility
1	Http LQ-sniffer	0.75	A_H	0.73
2	Install Trojan	0.80		0.72
3	CF-exploit attack	0.65		0.73
4	Remote buffer overflow	0.95		0.82
5	Shutdown Database server	0.45	A_M	0.40
6	Oracle TNS Listener	0.35		0.3
7	Attack SSH on Web Server	0.40		0.45
8	Ftp rhost attack	0.3		0.29
9	Apache chunk overflow	0.21	A_L	0.27

used. After comprehensive processing and analysis of experimental data, the corresponding conclusions of this experiment are drawn.

3.3 Experimental Results

The experimental data of defense cost comparison between the two defense strategy generation methods is shown in Table 2. The defense cost of the experimental group method is compared with the actual defense generation value, and the defense effect is evaluated.

According to the data in Table 1, it can be seen that the difference between the defense cost of the defense strategy generated by the experimental group and the actual defense generation value is smaller, and the defense cost of the defense strategy of the experimental group is much smaller than that of the control group. The defense grade of the experimental group was the same as that of the actual defense grade, while the defense grade of the control group was different from that of the actual defense grade. The maximum difference between the defense cost of the generated strategy in the experimental group and the actual defense cost was 0.5; the maximum difference between the defense cost of the generated strategy in the control group and the actual value was 2.9 and the minimum difference was 1.0. According to the above data, compared with the control method, the defense cost of the generated strategy in the experimental group was reduced by about 50% at the minimum. In other words, the defense strategy generated by the experimental group has lower defense cost and better defense effect.

Table 2 Comparison of defense costs of two groups of defense policy generation methods

			Experimental Group Method		Control Group Method	
The Serial Number	Quantified Value of Actual Defense Cost	Actual Level	Defense Strategy Level	Defense Costs	Defense Strategy Level	Defense Costs
1	5	1	1	5.1	2	6.1
2	8	3	3	8.2	4	10.2
3	4.2	1	1	4.5	2	6.2
4	6.5	2	2	6.5	3	8.4
5	5.7	2	2	6.2	3	8.6
6	6.2	2	2	6.4	3	8.2
7	8.1	3	3	8.4	4	9.8
8	9.6	4	4	9.6	5	10.6
9	12.7	5	5	12.8	5	13.7
10	10.9	4	4	11.0	5	12.5

When the power network attack strategy changes, measure the response time and space occupation when generating the strategy. The specific data is shown in Figure 2.

The analysis of Figure 1 shows that When the attack state in power network changes, the response time of the experimental group in generating the defense strategy is significantly shorter than that of the control group. At the same time, the space occupied by the defense strategy generated by the experimental group was much less than that of the control group. In other words, the experimental method can respond to the changes of network attack status more quickly and generate the corresponding security defense strategy with less running space occupation. Compared with the control group method, the experimental group can effectively shorten the response time of power network security defense strategy generation.

To sum up, when the generation method of power network security defense strategy based on Markov decision process is applied, the generation response time of power network security defense strategy is short, the generation efficiency of security defense strategy is high, and the defense cost is low, which can effectively improve the defense effect of power network security defense strategy.

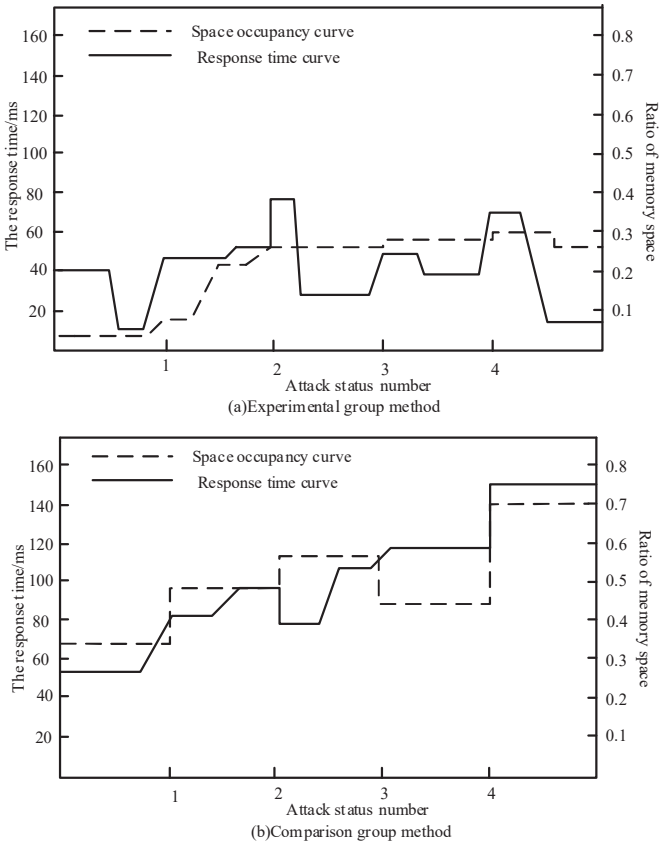

Figure 2 Defense policy response time generation and footprint comparison.

4 Conclusion

The widespread use of power network has brought a lot of convenience to people's life, work, entertainment and study. At the same time, the universal connectivity of power network makes the security problems of power network prominent with the development of network. Power network security defense has become the research focus in related fields. This paper studies the generation method of power network security defense strategy based on Markov decision process. By comparing with the traditional security strategy generation method, the response time of this research method is shorter, the defense effect is better, and the method is superior. However, this method does not consider the actual power network operation state. Therefore, in

the future research, we need to consider the actual power network operation state, improve the accuracy of the evaluation of power network vulnerability in the process of defense strategy generation, so as to optimize the strategy generation method.

References

[1] Xiong Xinli, Yang Lin, Li Kechao. A Strategy Optimization Model of Moving Target Defense Based on Markov Decision Process [J]. Journal of Wuhan University (Natural Science Edition), 2020, 66(02):141–148.

[2] Song Jiahan, Li Jingjiao, Pi Jie, et al. Research on Attack and Defense Strategy of Substation Network Security Applying Markov Decision Process[J]. Electric Power Construction, 2019, 40(10):104–110.

[3] Meng Xiaodong. Anti-theft method of sensitive data in link network in large data background[J]. Journal of Xi'an Polytechnic University, 2019, 32(02):212–217.

[4] Zhang Xingming, Gu Zeyu, Wei Shuai, et al. Markov game modeling of mimic defense and defense strategy determination [J]. Journal on Communications, 2018, 39(10):143–154.

[5] Yang Junnan, Zhang Hongqi, Zhang Chuanfu. Defense decision-making method based on incomplete information stochastic game.[J]. Chinese Journal of Network and Information Security, 2018, 4(08):12–20.

[6] Liu Jingwei, Liu Jin-Ju, Lu Yuliang, et al. Optimal Defense Strategy Selection Method Based on Network Attack-Defense Game Model [J]. Computer Science, 2018, 45(06):117–123.

[7] Based on Markov Evolutionary Game [J]. Acta Electronica Sinica, 2018, 46(06):1503–1509.

[8] Tian Jiwei, Wang Buhong, Li Xia, et al. Optimal Defense Strategy Against Load Redistribution Attack Based on Game Theory [J]. Computer Simulation, 2018, 35(05):123–127+190.

[9] Pan Qiuyu. Network intrusion detection based on hidden markov model [J]. Information Security and Technology, 2018, 9(04):65–68.

[10] Hu Hao, Liu Yuling, Zhang Hongqi, et al. Route Prediction Method for Network Intrusion Using Absorbing Markov Chain [J]. Journal of Computer Research and Development, 2018, 55(04):831–845.

Biographies

Wang Yang, graduated from school of computer science of Xinjiang Normal University in 2020 with a master's degree. At present, she works in the vocational and technical college of Xinjiang Normal College. Her main research directions are: computer science and technology, information technology teaching theory.

Liu Dong, graduated from school of telecom engineering of Beijing Jiaotong University in 2008 with a bachelor's degree. Currently, he works in the vocational and technical college of Xinjiang Ploytechnical University. His main research directions are: computer science and technology, information technology teaching theory.

Wang Dong, graduated from school of computer science of Xinjiang Normal University in 1989 with a master's degree. At present, he works in the vocational and technical college of Xinjiang Normal College. His main research directions are: computer science and technology.

Xu Chun received a Ph.D. from the University of Chinese Academy of Sciences. Her research interest is natural language processing. Currently, mainly engaged in the research of big data analysis and prediction at the Xinjiang University of Finance and Economics. Contact her at xuchun@mails.ucas.edu.cn.