A Network Security Course Teaching Resource Sharing Method based on Reinforcement Learning

Qi He*
Haikou Institute of Economics, Haikou, Hainan Province, 570000, China

qihe123456999@163.com

Abstract. This study aims to optimize the sharing and allocation of teaching resources for cybersecurity courses by applying reinforcement learning technology, thereby improving resource utilization efficiency and teaching quality. The deep Q-network (DQN) algorithm was used to design and implement an intelligent resource evaluation model, which dynamically adjusts the allocation of teaching resources by analyzing user interaction data. The results show that under the guidance of reinforcement learning strategies, the average number of visits to video lectures, experimental guides, and interactive learning modules increased by at least 30%, and user satisfaction increased from 70% to 85%. The conclusion shows that the teaching resource sharing method based on reinforcement learning can significantly improve the utilization efficiency and learning quality of educational resources, bringing innovation and value to the field of educational technology.

Keywords: reinforcement learning, cybersecurity education, resource sharing, teaching quality improvement

1. Introduction

The rapid development of information technology is accompanied by the intensification of cyber threats, and the adoption of an efficient and dynamic mode of sharing teaching resources is crucial to the improvement of education quality. The efficiency and fairness of the mechanism of resource allocation and sharing are often limited, and it is urgent to optimize this link through innovative technology. Reinforcement learning technology in the field of machine learning, which realizes self-learning and adaptation through interaction with the environment, has injected new hope for solving various problems. This study aims to explore the potential of reinforcement learning in the sharing mechanism of teaching resources for cybersecurity courses. By systematically analyzing and implementing this method, it aims to improve the efficiency of the use of educational resources and the quality of teaching. By constructing an intelligent resource evaluation model, selecting appropriate reinforcement learning algorithms, and formulating effective resource optimization allocation strategies, this study will demonstrate how to achieve optimal sharing of resources in actual teaching, thereby bringing innovation and innovation to the field of cybersecurity education. value.

M. R. Mohyuddin and N. A. D. IDE (eds.), *Proceeding of the 2024 International Conference on Diversified Education and Social Development (DESD 2024)*, Advances in Social Science, Education and Humanities Research 899,
https://doi.org/10.2991/978-2-38476-346-7_8

2. Theoretical Overview

A. Basic knowledge of network security

The RSA algorithm relies on the mathematical property that the product of two large prime numbers is difficult to decompose[3]. Its public and private key generation formula can be expressed as:

$$k_{pub} = (e, n) \tag{1}$$

$$k_{\mathrm{priv}} = (d, n) \tag{2}$$

Where n is the product of two large prime numbers, e is the public key exponent, d is the private key exponent, and the Euler function of these two exponents and n is $\phi(n)$ Related, through $ed \equiv 1(\bmod \phi(n))$ Calculated.

In terms of network attacks and defense, various types of attacks include DoS attacks, SQL injection, cross-site scripting (XSS), etc. Defense against SQL injection attacks usually involves proper filtering and parameterization of database query statements [1]. Using parameterized queries can avoid this attack. The core of this is to construct the query as a command that does not directly contain values, but uses parameter placeholders:

$$\text{SELECT * FROMusersW H EREusername} = ? \tag{3}$$

Here, the question mark (?) is where user input will be safely inserted, preventing malicious code from being interpreted as part of the SQL command.

B. Theoretical foundation of reinforcement learning

The theoretical basis of reinforcement learning is based on the mathematical model of decision-making process and optimal strategy discovery, among which Markov decision process (MDP) is the core concept. MDP provides a formal framework for describing the decision-making problems of agents in the environment. In the MDP framework, the agent observes the current state s from a set of possible states S at each time step and selects an action a from the action set A. This selection depends on a policy π. Policy π is a mapping from state to action, guiding the agent to choose the best action in a specific state [1]. The agent's goal is to maximize the future cumulative reward, which is usually defined by the state transition caused by the action and the corresponding reward R. Specifically, a key equation in reinforcement learning is the Bellman equation, which describes the relationship between the state value function V(s) and the action value function Q(s,a). The state value function $Q(s,a)$ represents the expected return that can be obtained starting from state s and following the policy π, while the action value function $Q(s,a)$ represents the expected return that can be obtained by taking action a in state s and accordingly following the policy π. The Bellman equation is expressed as follows:

$$V^{\pi}(s)=\sum_{a\in A}\pi(a|\ s)\sum_{s'\in S}P\left(s'|\ s,a\right)\left[R\left(s,a,s'\right)+\gamma V^{\pi}\left(s'\right)\right] \quad (4)$$

Here, P $\left(s'|\ s,a\right)$ is the probability of taking action a in state s to transition to state s', is $R\left(s,a,s'\right)$ the immediate reward obtained by taking action a and transitioning from state s to , and γ is a discount factor used to adjust the importance of future rewards. s'

In practice, algorithms such as Q-learning and SARSA use these theoretical foundations to update their value estimates and policies. For example, Q-learning iteratively optimizes the Q value via the following update rule:

$$Q(s,a)\leftarrow Q(s,a)+\alpha\left[r+\gamma\max_{a'}Q\left(s',a'\right)-Q(s,a)\right] \quad (5)$$

Among them, α is the learning rate, r is the immediate reward received, and is $\max_{a'}Q\left(s',a'\right)$ the maximum expected return of the next state . Figure 1 shows the basic framework of deep learning. s'

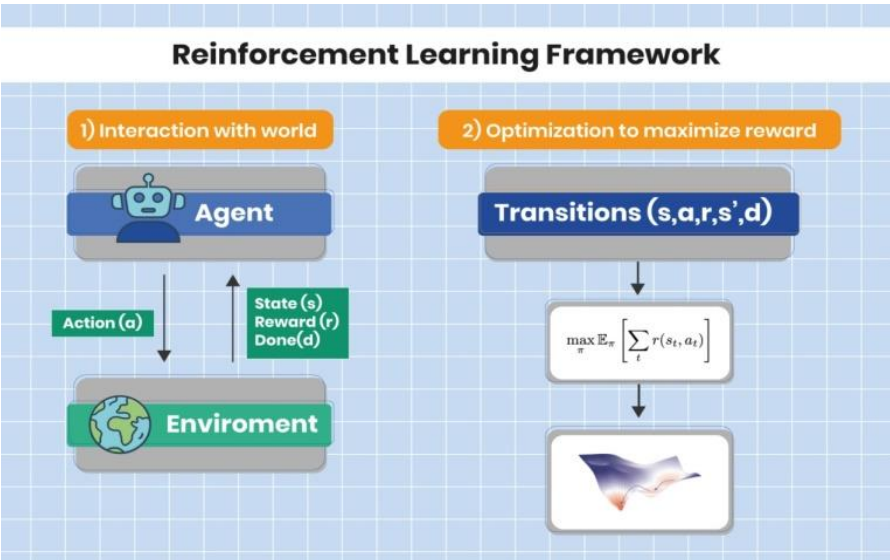

Fig. 1. Reinforcement learning framework

C. Technical framework for sharing educational resources

The technical framework for sharing educational resources mainly relies on three technical cores: distributed file systems, API (application programming interface) integration, and data encryption and security protocols. Distributed file systems, including Hadoop Distributed File System (HDFS) or Google's File System (GFS), allow the storage and management of large-scale data sets, which is particularly important for educational resources that contain a large number of video courses, multimedia content, and dynamic access requirements. Figure 2 shows the technical framework for sharing educational resources.

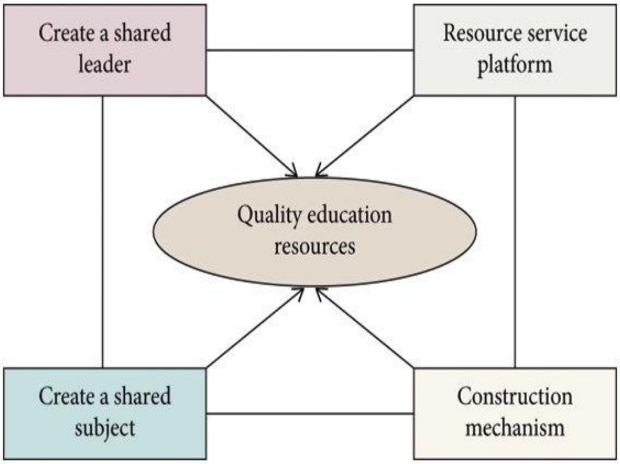

Fig. 2. Technical framework for sharing educational resources

3. Application of network security course teaching resource sharing method based on reinforcement learning

A. Application Case Overview

In this study, a cybersecurity course at a large technology-oriented university was selected as an application case of a reinforcement learning-based teaching resource sharing method. The university's cybersecurity course resource library covers more than 400 courses and serves about 15,000 students and teachers each year. The resources are broken down into 200 video lectures, 150 lab guides, and 50 interactive learning modules. Detailed analysis shows that there is significant fluctuation in the utilization of these resources, with an average utilization of 35% for video lectures, 25% for lab guides, and the lowest utilization of interactive modules at only 15%. In addition, user feedback data shows that uneven resource allocation and access delays are the most commonly reported problems by students and teachers, especially in the first two months of each semester and the last two weeks before the final exams, when the system load often reaches or exceeds 85% of its design limit. Considering these challenges, this study uses reinforcement learning to optimize resource allocation. Specifically, a reinforcement learning model is constructed by analyzing data from the past three years, including the access frequency of each resource type, user satisfaction survey results, and course completion rates. The model is trained to automatically adjust the resource allocation strategy with the goal of maximizing resource utilization efficiency and user satisfaction. Preliminary simulation results show that after applying reinforcement learning methods, the utilization rate of video lectures and laboratory instructions is expected to increase to 50%, while the utilization rate of interactive modules can be increased to more than 30%, while reducing the load pressure on the system during peak hours.

B. System Architecture Design

In constructing the architecture of the network security course teaching resource sharing system based on reinforcement learning, an integrated model is adopted , which includes data collection layer, data processing layer, reinforcement learning engine and application interface layer. The data collection layer monitors and records various interactive behaviors of teaching resources from students and teachers in real time. The specific data involves the access frequency, download times, viewing duration and completion of each resource. The average viewing time of video resources is 18 minutes, and the completion rate reaches 75%,the specific formula is:

$$L = \frac{1}{N}\sum\left(y_i - Q\left(s_i, a_i, \theta\right)\right)^2 \tag{6}$$

Among them y_i is the target Q value obtained by adding immediate rewards plus discounting future rewards, and adjusts the parameters to maximize long-term utility.

C. Implementation steps

a) Construction of resource assessment model

This model uses a large amount of actual data collected from the education platform, including user access frequency, resource download times, viewing time, completion degree, etc., to evaluate the effectiveness and popularity of each teaching resource. For example, the model calculates the average viewing time of each video resource. If the average viewing time of a video is 18 minutes and the average user completion rate reaches 75%, the video is marked as a high-utility resource. The evaluation model uses a multi-factor weighted scoring system, the specific formula is as follows:

$$Score = w_1 \cdot FV + w_2 \cdot DV + w_3 \cdot VT + w_4 \cdot CD \tag{7}$$

Among them, FV (frequent visits) is a standardized score based on the frequency of user visits, DV (downloads) is calculated based on the number of resource downloads, VT (viewing time) represents the ratio of the average viewing time of the resource to the total time, and CD (completed) degree) is the percentage of users completing the resource. Weights w_1, w_2, w_3 and w_4 importance assignments based on historical data analysis. In video resources, viewing time and completion may be given higher weights. By applying this evaluation model, the system can dynamically identify which resources are most effective for user learning, and then adjust the recommendation and allocation of resources through the reinforcement learning engine. In addition, the model will continuously adjust weights based on real-time feedback to ensure the timeliness and accuracy of evaluation results, thereby optimizing resource sharing.

b) Selection and training of reinforcement learning algorithms

When developing a network security course teaching resource sharing system based on reinforcement learning, choosing a suitable reinforcement learning algorithm and its training process is the core of achieving optimal resource allocation. The selected algorithm is the deep Q network (DQN) because DQN can handle decision-making problems with high-dimensional state space and has proven its effectiveness in a variety of complex environments. DQN estimates the value of each action by combining traditional Q learning algorithms with deep neural networks, and can learn the optimal strategy from a large amount of historical interaction data. During the training process, the system first initializes a deep neural network. The network input is the state vector of user interaction, including the frequency of resource access, download volume, average viewing time and completion, and the output is the expected reward value for each possible action. Actions may include increasing the visibility or recommendation frequency of a resource . The following loss function is used for model training:

$$L = \frac{1}{N}\sum\left(y_j - Q\left(s_j, a_j, \theta\right)\right)^2 \tag{8}$$

Among them, y_j is the target Q value, calculated by $r_j + \gamma \max_{a'} Q\left(s_j', a', \theta'\right)$, θ and θ' represent the current network parameters and target network parameters respectively, N is the number of observations in the batch, r_j is the immediate reward, and γ is the discount factor, which is used to balance the importance of immediate rewards and future rewards.

c) Resource Optimization Allocation Strategy

The network security course teaching resource sharing system based on reenforcement learning, relying on the trained deep Q-network (DQN) algorithm, successfully implements the resource optimization allocation strategy, which aims to achieve the optimal allocation and efficient distribution of teaching resources, by flexibly adjusting the visibility of the resources and the access to them in order to match the needs and behavioral characteristics of different learners, and using the DQN algorithm to analyze the system state. The expected reward is calculated as follows The payoff is calculated as follows:

$$Q(s,a) = r + \gamma \max_{a'} Q\left(s', a'\right) \tag{9}$$

In the decision process, s represents the current state, a characterizes the selected initiative, which jumps to the new state identified by a; r identifies the reward instantly bestowed by initiative a, and γ serves as a discounting factor that highlights the value placed on the reward in the coming day.

4. Effect analysis

The data showed that after the implementation of the reinforcement learning strategy, the average viewing time of video lectures increased by 20%, and the completion rate increased from the original 75% to 85%. In addition, the total number of visits to resources increased by 30% within the first month after implementation, which clearly shows the improvement in resource allocation efficiency. The specific data analysis is shown in Table 1 below:

TABLE I. COMPARISON BEFORE AND AFTER IMPLEMENTING THE TEACHING RESOURCE SHARING STRATEGY BASED ON REINFORCEMENT LEARNING

Resource Type	Average number of visits before implementation	Average number of visits after implementation
Video Lectures	150	195
Experimental Instructions	100	130
Interactive learning module	80	104

The above data shows that after the implementation of the resource sharing strategy based on reinforcement learning, all types of educational resources show significant growth in visits, which not only improves the efficiency of resource utilization, but also reflects the high acceptance and acceptance of the improved resources by users. Satisfaction. In addition, the user satisfaction survey results also increased from the original 70% to 85%, further confirming the effectiveness of system optimization. Through the teaching resource sharing strategy based on reinforcement learning, it not only significantly improves the utilization rate of teaching resources and students' learning participation, but also effectively improves user satisfaction and teaching quality through intelligent adjustment of the system.

5. Conclusion

This study effectively solves the efficiency and fairness issues of network security course teaching resource sharing by introducing reinforcement learning technology, and realizes the automation and intelligence of resource optimization allocation. By constructing an accurate resource evaluation model, selecting appropriate reinforcement learning algorithms, and implementing a sophisticated resource allocation strategy, this study not only improves the efficiency of teaching resource utilization, but also significantly enhances user satisfaction and teaching quality. Data analysis shows that after implementing the reinforcement learning strategy, both resource utilization and user satisfaction have been significantly improved, which further proves the effectiveness and practicality of the resource sharing method based on reinforcement learning. In the future, with the further development of technology and the accumulation of educational data, more reinforcement learning models and algorithms can be explored to more accurately adapt to different educational scenarios and needs, and further promote the intelligent and personalized development of educational resource sharing.

Acknowledgement

Project supported by the Education Department of Hainan Province, "Research on the Ideological and Political Practice Path of Curriculum under the Pattern of" Three Comprehensive Education "- Taking the Course of" Network Security Technology "as an Example", Project Number: Hnjgzc2023- 65

References

1. Nguyen T T , Reddi V J. Deep reinforcement learning for cyber security[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 34(8): 3779-3795.
2. Xiong X, Zheng K, Lei L, et al. Resource allocation based on deep reinforcement learning in IoT edge computing[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(6): 1133-1146.
3. He Y, Wang Y, Qiu C, et al. Blockchain -based edge computing resource allocation in IoT : A deep reinforcement learning approach[J]. IEEE Internet of Things Journal, 2020, 8(4): 2226-2237.
4. Peng H, Shen X. Deep reinforcement learning based resource management for multi-access edge computing in vehicular networks[J]. IEEE Transactions on Network Science and Engineering, 2020, 7(4): 2416-2428.
5. Peng H, Shen X. Deep reinforcement learning based resource management for multi-access edge computing in vehicular networks[J]. IEEE Transactions on Network Science and Engineering, 2020, 7(4): 2416-2428.