2024 10th International Conference on Machinery, Materials and Computing Technology (ICMMCT 2024)

A Method for Constructing Region Sensitive Models Based on Deep Reinforcement Learning Networks

Kaili Shao, Jifeng Qin*

Faculty of Engineering, Huanghe Science and Technology University, Zhengzhou, 450003, China

Keywords: Deep reinforcement learning; Network construction; Regional sensitivity; Model

Abstract: With the continuous development of current internet technology, the requirements for wireless sensor network signal performance are gradually increasing. Affected by external and internal noise factors, sensor network signals are prone to significant fluctuations, resulting in reduced stability and weakened network performance. Therefore, this article conducts research on the method of constructing region sensitive models based on deep RLnetworks, improving the convergence speed of the network, reducing routing load, and providing routing control methods for each connection. RL(Reinforcement Learning) refers to learning a suitable set of actions for selection, with the aim of maximizing the reward value. Given the continuous interaction with the external environment, the algorithm proposed in this article is based on a set of RLagents to learn the most suitable spectrum allocation scheme, in order to maximize the given reward in each access network. The research results indicate that the algorithm proposed in this paper reduces the scale of long convergence events in the network. The convergence times of AGA (Adaptive Genetic Algorithm) and AR (Augmented Reality) are both very short, generally around 0.58 seconds. However, the convergence time of the algorithm proposed in this paper is the shortest, only around 0.43 seconds. The experimental results verify that the algorithm proposed in this paper effectively improves the efficiency of network convergence time.

1. Introduction

With the rapid development of network technology, the scale and complexity of the network are constantly expanding, and the heterogeneous problems of the network are becoming more and more prominent, which makes the network management face great challenges. Regional sensitivity is different from hot topics, because hot topics are divided into positive, neutral and negative hot topics; Regional sensitivity is different from harmful information, because harmful information refers to aggressive hackers, viruses, etc. Regional sensitivity is different from interest information, because interest information is the user's online behavior [1]. Network construction has become a very important part of the whole security system. Therefore, sensitive information filtering is the basis of developing network public opinion monitoring system, which has extremely important academic significance and application value. People require the new network management model to manage any network at any time, at any place and under any platform, to solve the problems of integration and expansibility of network applications, to make the network scalable, and to flexibly distribute and automatically execute network management tasks [2]. With the continuous development of Internet technology, the requirements for signal performance of wireless sensor networks are gradually increasing. Affected by external and internal noise factors, the signal of sensor network is prone to obvious fluctuation, which leads to the decrease of sensor network stability and network performance [3]. Therefore, this paper studies the method of constructing a region-sensitive model based on deep RLnetwork, which improves the convergence speed of the network, reduces the load of routing, and provides the routing control mode for each connection. However, all new routing systems are faced with an important problem: because organizations and entities have different requirements for improving routing controllability, it is unrealistic to require all entities to update to a new protocol at the same time. Only those countries or organizations that have a strong demand for routing controllability are willing to adopt the new routing system, and other nodes in the network are not necessarily willing to deploy it, which limits its feasibility [4-5].

  DOI: 10.25236/icmmct.2024.033

Traditionally, non-linked static failure is the dominant failure type. With the continuous development of technology, some old failure modes seem to become more and more problems. Therefore, the sensitive failure of region model is very important not only for dynamic random access memory, but also for static random access memory.

2. Correlation model

2.1. Node energy consumption model

The energy consumption comes from the sensitive components and signal conditioning circuits of sensor nodes. Under the same hardware circuit and detection program, the energy consumption is related to the sampling frequency and sampling accuracy. When a node switches between the same access point, it does not need to be a node with mobility support. Only when a mobile node requires global mobility support, it must be a sensitive support node, that is, in global sensitivity support, while sensitive nodes do not need global mobility support in regional mobility management. The processing energy comes from the program execution and data storage of the microcontroller module [6]. The energy consumed by network nodes to execute instructions is much lower than the energy consumed by communication module data transmission and reception, so processing energy consumption is ignored in this article. The current levels for receiving, monitoring, and sleeping states are related to the circuit characteristics of the communication module and the communication method used, and can be determined based on the data manual and relevant protocol standards of the chip used. The received current and monitoring current can be considered equal [7].

2.2. Path loss model

Path loss represents the power attenuation in the process of electromagnetic wave propagation. When the propagation channel is free space, according to the Friis free space model, the received signal power is low. Same transmission signal power. The relationship between can be expressed as:

$$P_{Rx} = P_{Tx} \frac{G_t G_r}{\beta d^2} \tag{1}$$

Where: G_t and G_r are the antenna gains of the transmitting node and the receiving node respectively, β is the system loss factor, and d is the propagation distance. Set the gain of both the transmitting antenna and the receiving antenna to 1, regardless of the system loss. When the channel is in non-free space, the received signal strength can be expressed as:

$$P_{Rx} = \frac{\lambda^2}{(4\pi)d^\alpha} P_{Tx} \tag{2}$$

Where: α is the path loss coefficient when the channel is non-free, and its value is related to the physical environment in which the channel propagates; when α is 2, it corresponds to the path loss in free space.

The protocol mainly introduces the concept of access service domain to manage the access network hierarchically. In this way, the position of the correspondent node in the network will have a great influence on the handover management between protocol domains, and different handover management steps must be designed according to the network structure determined for different handover scenarios [8].

3. Regionally sensitive models

The so-called RLrefers to learning that maps from the environment to behavior, in order to maximize the cumulative reward value of behavior. In the application of complex nonlinear practical systems such as air traffic control and automatic manufacturing systems. Designers find it difficult to design computable systems for them. Therefore, it is required that intelligent systems can learn online from dynamic environments through trial and error to find solutions to problems, and RLrefers to learning appropriate action sets for selection, with the aim of maximizing the reward value [9]. How to provide real-time statistical model analysis for complex and massive network traffic, how to effectively identify the statistical characteristics of traffic containing intrusion attacks, and how to effectively improve detection accuracy are all issues that need to be addressed. How to provide real-time statistical model analysis for complex and massive network traffic, how to effectively identify the statistical characteristics of traffic containing intrusion attacks, and how to effectively improve detection accuracy are all issues that need to be addressed. Given the continuous interaction with the external environment, the algorithm proposed by the author is based on a set of RLagents to learn the most suitable spectrum allocation scheme, in order to maximize the given reward in each access network.

Assuming that the goal of network operators is to maximize the spectral efficiency of the entire heterogeneous network while ensuring the quality of service for each community user. To address this issue, RLsystems include RLagents based on RLmethods. As shown in Figure 3. The action unit generates a function that maps the state of the cell to an action; The random action selection unit randomly selects an action based on the output of the action unit to achieve the function of action exploration; The evaluation unit evaluates the adopted actions based on external reinforcement signals and the current state of the community, and its output is generally a predicted value, thus forming an internal reinforcement signal [10]. The self-organizing state space representation unit generalizes continuous sensor information into a representation suitable for processing action units and evaluation units.

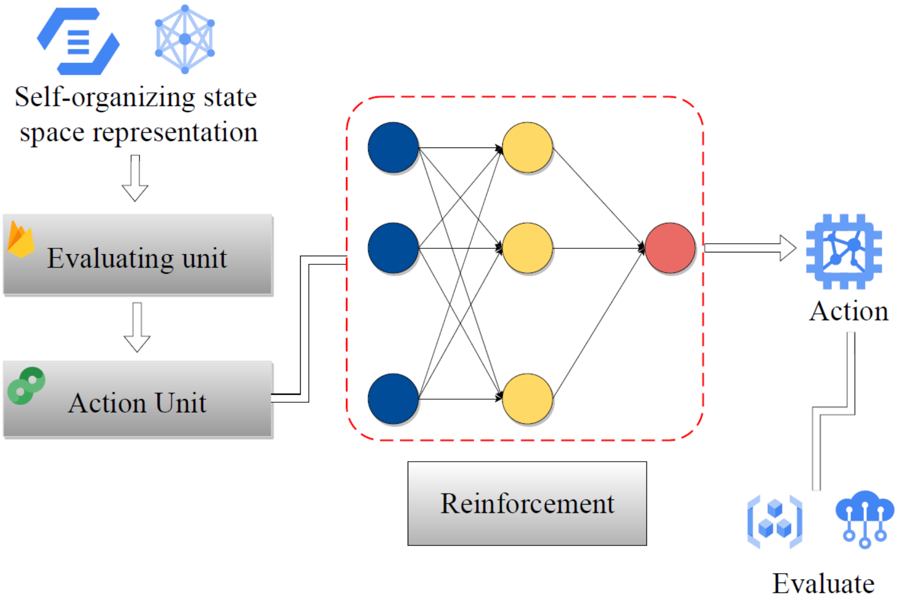

Figure 1 Region-sensitive model based on deep reinforcement learning

According to their respective functions, corresponding learning algorithms are designed to adjust the connection weights between the action unit and evaluation unit of the network output layer and the hidden layer of the network. Gaussian function is selected as the excitation output of the hidden node of the network;

$$a_i(x(t)) = \exp\left(-\frac{1}{2}M_t\right) \tag{3}$$

Where M_t is the matrix that determines the shape of the excitation output. The corresponding normalized output is:

$$b_i(x) = \alpha_i(x(t)) / \sum_{j=1}^{K} a_j \tag{4}$$

The i th action selection function of the action network is:

$$t_i^0 = \sum_{j=1}^{N} w_j b_j \tag{5}$$

In reinforcement learning, there is a problem: which experimental strategy can produce the most effective learning, that is, the learner faces a process of empowerment. In order to solve this problem, the output of the action unit does not directly act on the system, but is first sent to the random action selection unit, which is equivalent to the search mechanism acting on the actual control quantity of the system. The model evaluates the output of the network, and after a delay, the evaluation signals of several times are provided to the neural network in batches to adjust the weight, and the corresponding traffic data with the largest weight is the predicted value of the scale coefficient. As the rate of the controller designed by RLmethod fluctuates greatly, the queue length in the buffer basically fluctuates around the set value, so that the network resources are fully utilized without the situation that the queue is full and the network is congested, and the queue is empty and the network resources are idle.

The training set obtained by deep RLis trained by the exercise value, and then used for multi-step prediction of actual network traffic. It is slow to respond to the change of input, which means that the transfer function is not as accurate as the middle part in mapping the data with values at both ends. In order to solve the above problems, it is often necessary to preprocess the network input sample data. In order to achieve accurate, comprehensive and high-quality results, the underlying model of sensitive information must be able to effectively judge the synonymous relationship, that is, the system can make necessary reasoning, based on the semantic relationship of keywords and reasoning, combined with certain database information, to analyze and reason the text, so as to find words similar to sensitive words.

4. Result analysis

Once a new network update arrives, select a new optimal path from the candidate paths and notify the neighboring area. If a certain connection within the region fails, the neighboring quickly notifies the entire region of the update status of the "downstream connection", and quickly reconstructs the path within the region after receiving the update status of the downstream connection. Real network traffic has fractal characteristics, and traditional short correlation models cannot accurately describe and predict real network traffic. The tree shaped multifractal structure is a mathematical description of multifractals, with multiplicative cascades as its core. Through multipliers, a given set is refined into smaller subsets. The model evaluates the output of the network, and after delay, batches of evaluation signals from several environments are provided to the neural network to adjust the weights. The traffic data corresponding to the maximum weight is the predicted value of the scale coefficient. Then, the training set obtained by deep RLalgorithm is used for weight training, and then used for multi-step prediction of actual network traffic. This study compared the convergence time of AGA, AR, and our algorithm on the network. We removed a connection in the network and waited for the network to converge; Move the connection back until the network stabilizes and calculate the consumption time of the entire process. We ignore the processing time of the CPU and only calculate the convergence time caused by connection latency. The cumulative distribution function of network convergence time under various algorithms is shown in Figure 2.

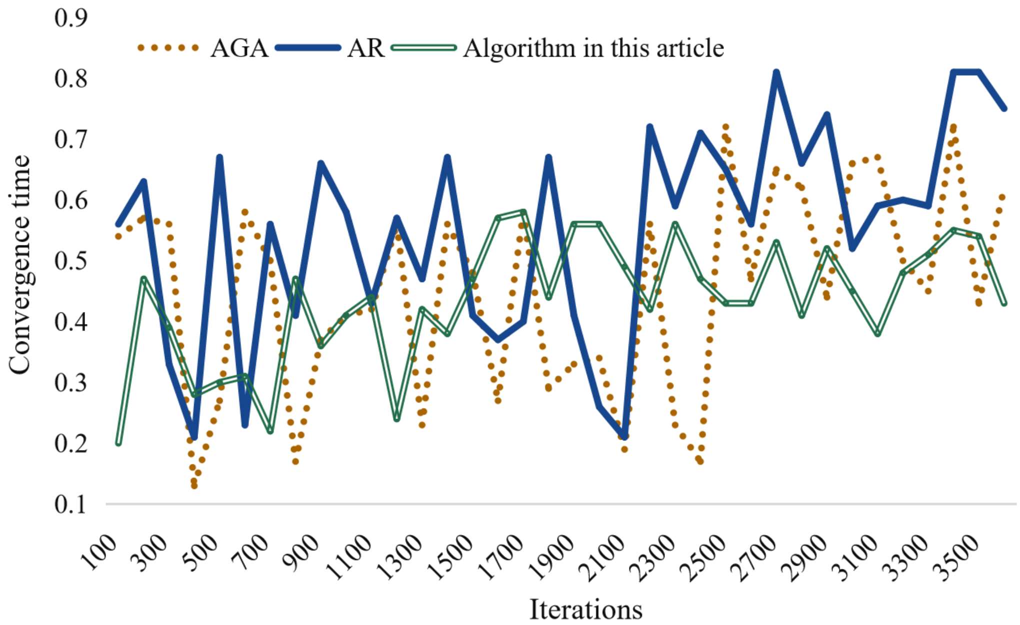

Figure 2 Comparison of network convergence time

As shown in Figure 2, the algorithm proposed in this paper reduces the scale of long convergence events in the network. The convergence times of AGA and AR are both very short, generally around 0.58 seconds. However, the convergence time of the algorithm proposed in this paper is the shortest, only around 0.43 seconds. The experimental results verify that the algorithm proposed in this paper effectively improves the efficiency of network convergence time.

5. Conclusions

In this paper, a regional sensitive model based on deep RLnetwork is given. The model first identifies sensitive words according to their characteristics, and then filters sensitive information according to the tendency analysis of sensitive words. Therefore, sensitive information filtering is the basis of developing network public opinion monitoring system, which has extremely important academic significance and application value. People require the new network management model to manage any network at any time, at any place and under any platform, to solve the problems of integration and expansibility of network applications, to make the network scalable, and to flexibly distribute and automatically execute network management tasks. The research results show that this algorithm reduces the scale of long convergence events, and the convergence time of AGA and AR is very short, generally about 0.58 seconds, but the convergence time of this algorithm is the shortest, only about 0.43. The experimental results show that the proposed algorithm can effectively improve the efficiency of network convergence time. The algorithm in this paper provides control means for each connection, and can control the influence of some network events in the area, which reduces the global sensitivity of routing and improves the convergence speed of the network. Because RLtheory is universal, this method can be applied to adaptive control, network routing, intelligent transportation, resource allocation, robot navigation and other fields.

Acknowledgements

2021 Basic Research and Applied Basic Research Project of Zhengzhou City: "Research on Network Models and Intelligent Decision-making Methods Based on Deep Reinforcement Learning in Complex Situations (Project No.: ZZSZX202107)"; Key Research Project of Henan Higher Education Institutions, "Complex Situational Intelligence Decision-making Method and Application Research Based on Deep Reinforcement Learning (Project No.: 22A520033)"

References

[1] Panou D N, Reczko M. DeepFoldit -- A Deep RLNeural Network Folding Proteins[J]. IEEE Transactions on Network Science and Engineering, 2020, 36(15):36-44.

[2] Indhumathi V, Kumar K. A Survey on Deep RLNetwork for Traffic Light Cycle Control[J]. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 2020, 45(4):16-27.

[3] Liang X, Du X, Wang G, et al. A Deep RLNetwork for Traffic Light Cycle Control[J]. IEEE Transactions on Vehicular Technology, 2019, 68(2):1243-1253.

[4] Santis E D, Giuseppi A, Pietrabissa A, et al. Satellite Integration into 5G: Deep RLfor Network Selection[J]. Machine Intelligence Research, 2022, 19(2):127-137.

[5] Chkirbene Z, Awad A, Mohamed A, et al. Deep RLfor Network Selection over Heterogeneous Health Systems[J]. IEEE Transactions on Network Science and Engineering, 2021, PP(99):11-19.

[6] Wu J, He H, Peng J, et al. Continuous RLof energy management with deep Q network for a power split hybrid electric bus[J]. Applied Energy, 2023, 222(35):799-811.

[7] Serrano W. Deep RLwith the Random Neural Network[J]. Engineering Applications of Artificial Intelligence: The International Journal of Intelligent Real-Time Automation, 2022, 63(110):110-123.

[8] Chen X L, Cao L, Li C X, et al. Ensemble Network Architecture for Deep Reinforcement Learning[J]. Mathematical Problems in Engineering,2018,(2018-4-5), 2018, 18(4):1-6.

[9] Yuan S, Zhang Y, Qie W, et al. Deep RLfor resource allocation with network slicing in cognitive radio network[J]. National Library of Serbia, 2021, 12(3):20-31.

[10] Bertsekas D P. Feature-based aggregation and deep reinforcement learning: a survey and some new implementations[J]. IEEE/CAA Journal of Automatica Sinica, 2019, 6(001):1-31.