Event Prediction Technology Based on Graph Neural Network

Qiang Fu and Yongchao Wei*
Civil Aviation Flight University of China, Guanghan, Sichuan, China

*Corresponding author e-mail: weiyongchao@cafuc.edu.cn

Abstract. With the rapid development of the mobile Internet and the continuous expansion of network scale, the network security situation is becoming increasingly severe, and the endless network security threats have put forward higher requirements for network security performance. Based on the above background, the purpose of this paper is to explore the event prediction technology based on graph neural network. Due to the slow convergence of the network event prediction and evaluation model, the untimely risk assessment and inaccurate safety prediction caused by the incomplete parameter setting of the prediction model have become prominent problems in this field. This paper proposes an event prediction technology based on graph neural network. This method first uses genetic algorithms to optimize the weights in the training process of the graph neural network, which overcomes the blindness of initial weight selection and improves the training efficiency of the graph neural network; the KDDCup99 data set is used to conduct experiments on the above two methods respectively. Verification and analysis. The simulation and comparison experiments respectively verify that the neural network-based network security situation assessment and prediction method proposed in this paper can realize the assessment and prediction of the network situation more efficiently and accurately.

Keywords: Graph Neural Network, Network Security, Situation Assessment, Event Prediction

1. Introduction
The frequent occurrence of various security incidents such as network vulnerabilities and attacks often causes serious consequences, continuously affecting the development of the national economy of all countries in the world, and even endangering national security and social stability. Therefore, countries all over the world have begun to pay attention to and pay close attention to the issue of network security [1-2]. Faced with increasingly serious network security issues, on the one hand, countries all over the world are formulating various laws and regulations related to network security [3]. On the other hand, many newer methods have also been proposed at the technical level, such as continuously improving the defense against network attacks through various software and hardware systems, and deploying machine learning-based packet identification firewalls to automatically learn and filter abnormal data packets to strengthen the network. Access control, dynamic upgrade of intrusion
detection system, timely detection of abnormal behavior in the network operation process, etc. [4-5]. Traditional network security measures (such as intrusion detection, firewall, data encryption, vulnerability scanning, etc.) also emphasize the use of static protection schemes for passive defense, and at the same time, there is a lack of necessary internal correlation between various defense measures, although they can be solved to a certain extent. Certain security issues, but technically and functionally, it has been unable to cope with the more severe network security situation [6].

In this context, network security situational awareness technology has emerged in the field of information security, and has received extensive attention and research once it appears [7-8]. This technology takes the network security status and development trend as the main entry point. It first collects key network data information, and then performs fusion analysis on it, and uses complex algorithms to calculate a comprehensive score to obtain information on the dangers suffered by the entire large-scale network. Carry out comprehensive monitoring and effective evaluation [9]. Through the evaluation results, it provides a reliable decision-making basis for estimating the current network security status, and helps network administrators realize timely perception, comprehensive evaluation, and overall control of network operation conditions [10].

Based on the basic algorithm, this article comprehensively adopts a variety of improvement measures, including adding inertia coefficients and adjusting the learning rate method to ensure that the network training has a better convergence effect. In the early stage of network training and during the training process, combined genetic algorithm and particle swarm algorithms and other intelligent swarm algorithms optimize the initial training parameters of the network and adjust the network topology to meet the requirements of accelerating convergence and improving prediction accuracy. Combined with practical applications, it is designed for network event prediction system.

2. Graph Neural Network and Cyber Security Incident

2.1. Graph Neural Network Structure

For an undirected graph \( G=(V,E,W) \), where \( V=[v_1,v_2,...,v_n]^T \) is the node in the graph; there may be edges between two nodes, denoted as \( E \); And \( W \) is a parameter describing the similarity of graph nodes. Graph neural network takes graph structure as input. For each node in \( G \), \( v=[p,t]^T \), where \( p \) is used to represent the location information corresponding to the node, and \( t \) is used to represent the texture information of the node. For the two nodes \( v_i \) and \( v_j \) in the undirected graph, the weight \( w_{ij} \) between them is expressed as:

\[
    w_{ij} = \|v_i - v_j\|_2^2
\]

(1)

In the graph neural network, each node has only the \( k \) nearest neighbors whose weight is closest to the edge \( E \). For each node, their local feature \( F \) is expressed as:

\[
    F = [\delta_{i1}f_{i1}, \delta_{i2}f_{i2},...\delta_{ik}f_{ik}]^T
\]

(2)

Where \( f_{i1}, f_{i2},...,f_{ik} \) correspond to the features of the \( k \) nearest neighbors \( v_{i1}, v_{i2},...,v_{ik} \) of node \( v_i \), and their weights satisfy \( w_{i1} < w_{i2} < ... < w_{ik} \), and \( \delta \) is the control node. The parameters of the point feature contribution, in order to highlight the contribution of the central node and weaken the contribution of the distant node, this paper defines the \( \delta_{ij} \) calculation method as:

\[
    \delta_{ij} = 1 - \frac{j}{k}, j = 1,2,...,k
\]

(3)

2.2. Network Security Situation

The network security situational awareness process is divided into several modules as shown in Figure 1. The main work of each module is shown in Figure 1:
Figure 1. Cybersecurity situational awareness framework model

1. Data collection: Collect various security element information that affects network operation through various network security monitoring equipment, as a prerequisite for security situational awareness;

2. Situational awareness: The collected security influence factors are combined with the rules of network security situation research, and the processed data is fused and analyzed. The result of the analysis is the basis for security situation assessment and prediction;

3. Situation assessment: First, through qualitative and quantitative analysis of various network element information, and then further quantitative assessment of the current network security status, and finally obtain the security status of the network for a certain period of time, which is the core of security situation awareness. The improved graph neural network algorithm will be used for network security situation assessment, and the current network status will be evaluated more quickly and accurately;

4. Situation prediction: Based on the situation assessment, by analyzing the results of the situation assessment, predicting the development trend of network security is the ultimate goal of security situation awareness. This article will use the improved RBF neural network algorithm for network security Situation forecasting work to accurately predict the development trend of the network.

5. Response and early warning: Based on the results of the situation assessment and prediction, the network administrator can obtain the security status of the network in time, and do a good job of network protection in advance.

3. Event Prediction Experiment of Graph Neural Network

3.1. Experimental Environment Construction
In order to verify the effectiveness and efficiency of the network security situation assessment method based on the graph neural network of attack index factors, this paper builds an experimental environment for verification. In this experiment, an attacker will be simulated to launch an attack on the network, and network sensors will be used to collect the router's Snort attack information.

3.2. Experimental Data Set
The KDDCup99 data set is currently the most commonly used intrusion detection standard data set. This data set has annotated specific network behaviors and can accurately detect the accuracy of the method. The five network behaviors are:

1) "Normal", use "Normal" mark;
2) "Denial of Service Attack", use the "DoS" label;
3) "Unauthorized local super user privileged access", marked with "U2R";
4) "Unauthorized access by remote host", use "R2L" mark;
(5) "Port Scan", use the "Probing" label. Each record in the KDDCup99 data set consists of 42 fields, of which the first 41 bits are characteristic fields, and the last one is the network behavior label field. The network attack types in this data set are classified as shown in Table 1. Different attacks are based on the attack. The acquired system permissions and the degree of influence on the network operation are different, and the threat level is also different. With the increase of the threat level, the influence of different attack types on the network security operation threat increases.

| Threat level | Attack type | Xi   | Label field                                      |
|--------------|-------------|------|-------------------------------------------------|
| low          | DoS         | 0.3  | back, land, neptune, pod, smurf, teardrop       |
|              | U2R         | 0.5  | rootkit, buffer_overflow, loadmodule, perl      |
|              | R2L         | 0.6  | guess_passwd, warezmaster, ftp_write, imap,     |
|              |             |      | phf, multihop, warezclient, spy                 |
| high         | Probing     | 0.8  | ipsweep, nmap, portsweep, satan                 |

At the same time, the complete data set contains a total of more than 4.9 million records. If all used as the verification data set of this article, it is too large, so 10% of the KDDCup99 data set is selected as the sample data of this article for experimental verification.

Network situation assessment has a strong periodicity. Set the assessment cycle to T, extract the KDDCup99 sample set data for intrusion detection operations within T time, and use Snort to obtain it. After Matlab7.0 was used for simulation and the situation assessment factors were obtained, the graph neural network proposed in this paper was used for training to make it have situation assessment capabilities, and then the situation assessment was carried out based on the data in the test set. According to the above method, the situation assessment experiment is carried out, and various parameters of the experimental model are given.

4. Analysis of Event Prediction Technology Based on Graph Neural Network

4.1. Network Security Situation Prediction Analysis
In this paper, genetic algorithm is used to optimize the weights of graph neural network, and then the optimization results are compared with the graph algorithm without improvement to verify the advantages of the optimization algorithm. The result is shown in Figure 2, which shows the convergence of the two algorithms in the iterative training process. In Figure 2, the abscissa is the number of iterations, and the ordinate represents the mean square error.

![Figure 2. Comparison of iterative convergence](image-url)
It can be seen from the results in Figure 2 that the convergence speed of the improved graph neural network algorithm proposed in this chapter is significantly better than that of the traditional graph neural network, and the final convergence results of the two are basically the same, and will not affect the training results of the neural network. Therefore, under the premise of ensuring the correct training results, the method in this paper is more efficient.

4.2. Analysis of Evaluation Experiment Results
Using the trained graph neural network to evaluate the test data set, combined with the classification of the security evaluation level in Table 1, randomly select the output results of 5 sets of test samples and the expected output results, as shown in Table 2.

| Sample number | Actual output value | Expected output value | Actual threat level | Expected threat level |
|---------------|---------------------|-----------------------|---------------------|-----------------------|
| 1             | 0.23                | 0.25                  | Mild                | Mild                  |
| 2             | 0.36                | 0.34                  | Mild                | Mild                  |
| 3             | 0.29                | 0.38                  | Mild                | Mild                  |
| 4             | 0.26                | 0.20                  | Mild                | Mild                  |
| 5             | 0.08                | 0.07                  | Safety              | Safety                |

Using the improved graph neural network to evaluate the network security situation is almost the same as the actual situation value of the network, and the accuracy of the evaluation results is very high, which proves the effectiveness and accuracy of the improved graph neural network algorithm for evaluating the network situation.

With the help of genetic algorithm to obtain the global optimal solution more efficiently, it is used to train the connection weights of graph neural network, optimize the neural network, accelerate the network convergence speed, and give the algorithm realization. At the same time, with the help of graph neural network to solve nonlinear mapping problems, the genetic algorithm-graph neural network proposed in this chapter is used to evaluate the network security situation. The experimental results show that when the graph neural network improved by the genetic algorithm is trained, the convergence speed is significantly better than that of the traditional graph neural network; at the same time, the trained network model is used for the network security situation assessment, and the accuracy of the evaluation result is very high.

5. Conclusion
The work of network security situation assessment and event prediction is the focus of this article. Through the monitoring of the network, various factors affecting the operation of network security are obtained, and these factors are analyzed, and the network security situation event assessment model is used to obtain the current network operation status; at the same time, the network situation prediction work is carried out by using the results of the assessment. It enables network administrators to understand the future development trend of the network in a more timely manner, promptly respond to possible security problems, and maintain the security and stability of the network. This paper introduces genetic algorithm into the training process of graph neural network, optimizes the weights, improves the training efficiency of graph neural network, and at the same time, realizes the classification of network security level by constructing situation indicator evaluation factor, and uses the improved graph neural network. The network conducts network security situation assessment to improve the accuracy and efficiency of the situation assessment results.

Acknowledgments
This work was financially supported by civil aviation security capacity building fund (14002600100018J002).
References

[1] Xie T, Grossman J C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Physical Review Letters, 2018, 120(14):145301.1-145301.6.

[2] Wang M, Hu G. A Novel Method for Twitter Sentiment Analysis Based on Attentional-Graph Neural Network. Information (Switzerland), 2020, 11(2):92-111.

[3] Yanlin, Hamed, Karimian. A hybrid model for spatiotemporal forecasting of PM2.5 based on graph convolutional neural network and long short-term memory. The ence of the total environment, 2019, 664(10):1-10.

[4] Xiao, Qinkun. Recurrent neural network system using probability graph model optimization. Applied Intelligence, 2017, 46(4):889-897.

[5] Ding X P, Wu S J, Liu J, et al. Functional neural networks of honesty and dishonesty in children: Evidence from graph theory analysis. Scientific Reports, 2017, 7(1):12085-12100.

[6] Lu Y, Chen Y, Zhao D, et al. CNN-G: Convolutional Neural Network Combined with Graph for Image Segmentation with Theoretical Analysis. IEEE Transactions on Cognitive and Developmental Systems, 2020, PP (99):1-1.

[7] A Y S, B Q L, C X X Z B. Building segmentation through a gated graph convolutional neural network with deep structured feature embedding. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159(6):184-197.

[8] Hu K, Wang Z, Wang W, et al. Graph Sequence Recurrent Neural Network for Vision-Based Freezing of Gait Detection. IEEE Transactions on Image Processing, 2019, 29(11):1890-1901.

[9] Vu H T, Wilkinson R H, Lech M, et al. A hybrid neural network for graph-based human pose estimation from 2D images. IEEE Access, 2020, PP (99):1-1.

[10] Bogaerts T, Masegosa A D, Angarita-Zapata J S, et al. A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data. Transportation Research Part C Emerging Technologies, 2020, 112(March 2020):62-77.