Data Transmission Security in Computer Network Communication

Fei Chen
School of Big Data, Chongqing Vocational College of Transportation, Chongqing, China

*Corresponding author e-mail: wyji@bjtu.edu.cn

Abstract. In recent years, computer network communication technology has made considerable progress and development, and its scale has now covered all aspects of society and people's lives. However, while computer network technology promotes social development and improves the convenience of people's lives, there are still many security problems, especially the difficulty of ensuring the security of user data, which has become a prominent problem in the further development and improvement of communication technology. This article takes the network communication security technology as an entry point, and analyzes its technical types with the aid of data encryption technology, and then analyzes and studies its security in computer network communication data transmission according to the specific points proposed. Through the substantial technical analysis of the computer network security system, the possibility of the system being attacked can be reduced to a certain extent, and the security of network data can be improved. Through the analysis of the integrity of the network system, the overall framework of a distributed network security early warning system based on the network is established, and various structures of the system are analyzed. The network load analysis module is responsible for using key performance index data, such as user traffic, number of users and wireless utilization, and analyzes and classifies the network cell load by combining the partial least square method and BP neural network, so that the key cell performance analysis module is mainly Analyze key performance index data of key communities, obtain early warning thresholds through clustering algorithms, and implement a dynamic early warning system to provide guidance for network planning and mining in key small areas. The database module is responsible for storing all performance data and performance analysis results; the performance query module is responsible for querying and analyzing performance data and static warnings, and presents them to users in the form of text and graphics. Experimental research results show that in this information age, the development of computer networks plays a very important role in the development of our country. Computer network operators make full use of cluster analysis, data mining, encryption algorithms, and early warning systems to conduct research to fully improve the security performance of data in network communications.
Keywords: Network Communication Security; Data Mining; BP Neural Network; Cluster Analysis

1. Introduction

China has entered the current Internet era. In the context of the rapid development of the Internet era, computer information technology has been widely used, especially in network communication engineering, which can effectively increase the transmission rate of information [1-2]. Although information technology has developed significantly in the operation of communication engineering, it also has a certain degree of network security. In order to eliminate the hidden dangers of data information in computer network communication security [3]. Computer network operators need to strengthen the application of data encryption technology, widely use computer network communication security instructions in communication engineering, and use data encryption technology and early warning analysis systems, mainly through physical links to lay the foundation for the development of the network communication field[4].

With the popularization of network communication technology and the rapid development of computer technology in recent years, various network security issues such as user property information leakage, corporate information theft, property information security are seriously threatened, and personal privacy information leakage continue to appear. Although people have protective measures such as firewalls and intrusion detection systems at this stage, most of them are passive defenses [5]. The BP neural network is used as the basic algorithm to identify data types, establish a network security perception system and an early warning mechanism system, which can actively discover and deal with various network attacks on its own. Before the attack occurs, it can actively discover, intercept, and automatically take corresponding solutions to ensure network security [6]. The network security perception system based on BP neural network is mainly composed of modules that find abnormal data traffic in the network through multi-level operation, modules that collect situational elements from the network security assessment results, and modules that conduct situational assessment of network security[7].

In addition, when the system is running, it needs to find network security threats and play a defensive role against these threats. To achieve this function, the system must continuously improve and update the network environment out of control control algorithm on the basis of long-term accumulation of data and information[8]. Because the information of the computer communication network is exchanged and stored in the virtual network space, it is easy to be deliberately attacked in the highly open environment of the computer network, intercepted or destroyed by hackers during transmission, or stolen and used by users. Users who have been illegally accessed in the storage state [9]. Therefore, first of all, because current Internet users access computer communication networks through mobile Internet, with the upgrade of mobile communication networks, it is difficult to accurately locate and track the behavior of such users, which brings technical difficulties to monitoring malicious networks [10].

2. Establishment of BP Neural Network Algorithm Based on Cluster Analysis

2.1. Cluster Analysis

The essence of cluster analysis is to cluster a group of physical or abstract objects into different classes. The similarity between objects in each category is higher than the similarity between objects in different categories. Cluster analysis is different from classification in many ways. This is an unsupervised learning process. Classification is based on different characteristics of the class. Clustering does not give classification features to divide categories.

(1) The steps to implement hierarchical clustering are as follows:

(2) Input data set $X = \{x_1, x_2, \ldots, x_n\}$ and the number of clusters $k$;

(3) Each data item of the input data set is treated as a separate class, and the distance between every two clusters is calculated;
2.2. BP Neural Network Algorithm

The essence of BP neural network is a gradient descent method. Its error function is constructed by the difference between the preset expected output value and the actual output value of the model. The error function is used to adjust the link weights between all nodes to make the output value get closer and closer to the expected value, and the error ends after reaching a certain accuracy. For neural network training, we must first divide training data and test data, and determine the input and output of each group of training data.

As mentioned earlier, because the number of samples in the training set is limited, this article only fine-tunes on the basis of the training set, so the correction rate cannot be too large. In this article, the initial correction rate of the model is 20%, and the correction rate increases by 85% every 3 cycles of training, namely:

\[
l_t = l_{t0} \gamma^\left\lfloor \frac{t}{3} \right\rfloor
\]

(2)

In the formula: \(l_{t0}=30\%\); \(\gamma=0.2\), \(t\) is the current number of training cycles; rounded down. The loss function is defined as cross entropy, and the formula is as follows:

\[
L = \sum_{i=1}^{K} y_i \log y_i'
\]

(3)

This article uses the stochastic gradient descent algorithm with momentum as the optimizer, and each parameter update is:

\[
\omega_l+1 = \omega_l - l_r \nabla L(\omega_l) + \gamma(\omega_l - \omega_l -1)
\]

(4)

In the formula: \(\gamma=0.8\) is the momentum; \(\nabla L(\omega_l)\) is the loss function to derive the parameter \(\omega_l\). This is the result we finally got.

Training the neural network model, the steps are as follows:

a. Calculate the output of each neuron in the hidden layer and output layer of the network.

\[
o_p = f_j(\text{net}_p) = f_j(\sum_{j=1}^{l-1} W_{pj}^j o_j^{l-1} - \theta_p^j)
\]

(5)
b. Calculate the error between the output layer and the hidden layer.

\[
\delta^{(2)}_{pj} = o^{(2)}_{pj} \left(1 - o^{(2)}_{pj}\right) \sum_{l} \delta^{(2)}_{pj} W_{jl}^{(l)} \\
\delta^{(2)}_{pj} = o^{(2)}_{pj} \left(1 - o^{(2)}_{pj}\right) \left(t - o^{(2)}_{pj}\right)
\]

2.3. Classification Results

The classification model in this article uses the data of June, July, and August to take three sets of samples for analysis and testing. The root mean square error (RMSE) is selected as the evaluation index to evaluate the error of the classification result. The formula is as follows:

\[
RMSE = \sqrt{n} \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n} = \sqrt[2]{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}
\]

Among them, RMSE represents the root mean square error of each classification method. Acc represents the correct rate of each classification method. A comparative analysis of the three-month classification results in June, July and August is now.

3. Modeling Method

3.1. Index System Model of the Network Early Warning System

There are m early-warning indicator systems, and n original data are supplied for each indicator. The early-warning indicator system processes these data to obtain a standard matrix:

\[
T = (r_{ij}) m \times n
\]

Among the m indicators, the i-th indicator is defined as:

\[
E_i = - \frac{1}{m} \ln(l_{ij}) \quad , i = 1, 2, ..., m.
\]

Among them:

\[
l_{ij} = \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}
\]

\[
W_i = \frac{1 - E_i}{m - \sum_{i=1}^{m} E_i}, 0 \leq W_i \leq 1
\]

For the security accident early warning indicators of the corresponding severe alert, moderate alert, and light alert states, the dimensionless processing is carried out by equation (1), and the dimensionless value \( x_i \) is obtained, and the index and the corresponding quantitative value are multiplied by weighted summation Get the weighted average:

\[
\Delta = \sum_{i=1}^{m} W_i x_i
\]

\[
R(t) = P[x(t) < [x]] \quad u(t) \geq [u] \\
P_{FR}(t) = \frac{u(t)}{[u]} \quad P_{RSD}(t) = 1 - \frac{u(t)}{[u]} \\
t_r - t_1 = \frac{1}{\frac{1}{(t_2-t_1)^2} - \frac{1}{(t_3-t_1)^2}}
\]

(13)
3.2. Security Risk Rating Modeling Analysis

In the computer network communication data storage security risk rating prediction model, it is a typical two-class model in the process of platform risk rating. The Logist model is the most widely used model. When \( Y = 0 \), the risk level is 0, when the value of \( Y \) represents 1, the risk level is 1. The probability can determine the cloud data storage risk level according to whether the threshold is set.

\[
\rho(Y = 1 | X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n)}}
\]

(14)

The fitting effect of the rating model is expressed by the pseudo R-square statistic, and its meaning is a proportional model explained by the self-edited amount of the total variation times model of the dependent variable, which shows that:

\[
\text{Cox}&\text{Snell} - R^2 = 1 - \frac{[\ln(L_0)]^{2n}}{\ln(L)} = 1 - \left( \frac{\sum_{i=1}^{n} \ln(L_0) - \ln(L)}{\sum_{i=1}^{n} \ln(L_0)} \right)
\]

\[
R^2 = \frac{\text{Cox}&\text{Snell} - R^2}{1 - \left[ \ln(L_0) - \ln(L) \right]^n}
\]

(15)

The coverage ratio of the result is the actual rating sample, the correct rating proportion model is as follows:

\[
\text{RESF} = \frac{TA}{TA + FA}
\]

(16)

The accuracy of the result is in the actual rating sample, the correct rating proportion model is as follows:

\[
\text{RESJ} = \frac{TA}{TA + FN}
\]

(17)

4. Algorithm Evaluation Results and Data Analysis

4.1. Evaluation Results

| Downstream bytes | Upstream bytes | Maximum number of RRC connections | Unlimited utilization | Data load level |
|------------------|----------------|----------------------------------|-----------------------|----------------|
| 3234234          | 644543         | 143                              | 87                    | High load 2    |
| 2342424          | 353622         | 127                              | 56                    | Low load 2     |
| 4597284          | 342425         | 667                              | 94                    | High load 2    |
| 4532423          | 526234         | 343                              | 89                    | Low load 2     |
| 5325424          | 245242         | 644                              | 43                    | Low load 1     |

As shown in the part of the training data set listed in Table 1, after literature research, in view of the poor stability and slow convergence speed in the application of BP neural network, a monitoring model based on cluster analysis and neural network coupling is proposed. Improve the stability and learning speed of neural networks. First, the method of cluster review is used to reduce multi-
dimensional independent variables and extract principal components to solve the problem of multiple correlations between variables. By combining partial cluster analysis and BP neural network to process the data, the principal component scores and the weights of the original independent variables and dependent variables are obtained. Then, the number of hidden nodes of the neural network, the initial weights of the input layer and hidden layer, and the weights of the hidden layer and output layer are determined by the principal component scores and weights, and used in practical applications, the actual effect is very good.

![Figure 1. Data mining efficiency of the improved BP neural network algorithm](image)

In the previous article, we use the computer network operator to directly influence the data mining algorithm in the work efficiency of the inspection center in the network data. Therefore, this article has made a comprehensive improvement and analysis of the traditional data mining algorithm. Now we have conducted experiments to compare the traditional data mining algorithm and the improved data mining algorithm. Of course, we focus on the efficiency of the system after using the new data mining algorithm. Next, through corresponding classification of 5546 links, several different data attributes are selected for corresponding data mining based on improved BP neural network algorithm. After data mining through the improved algorithm, different support and confidence levels are obtained. The results of data mining are shown in Figure 1.

![Figure 2. The processability of network communication security data based on the improved BP neural network](image)

Analyzing the data shown in Table 1 and Figure 2, in the implementation of intelligent analysis of the processing degree of network communication security data based on the improved BP neural network, data mining processing on it will lead to the possibility of data intermittent situations, the consequences of this situation are more serious. On the other hand, in a large number of normal operating conditions and abnormal operating conditions, managers are mainly concerned about...
abnormal operating conditions, and the rest of the normal operating conditions are actually out of everyone's consideration. In supervised learning, this kind of unlabeled samples that are not part of the predictions play a very important role, and they are usually included in the data set for risk assessment, analysis and processing; and in semi-supervised learning, this kind of sample describes the data distribution, the construction of learning models is of great benefit. Therefore, by designing experiments, this paper conducts supervised learning and semi-supervised learning on the same data set with the same labeled part, uses consistent evaluation indicators to evaluate the performance of the model, and finally selects the best performing model for the risk of computer network communication. The status of the rating makes accurate predictions.

5. Conclusion
Based on the above-mentioned research, the rapid development of computer network communication technology has brought tremendous pressure to network security protection. The rapid development of intrusion and attack technology, and more and more forms of intrusion, various methods are also strange, due to The hazards can quickly spread to the intelligent and efficient early warning system in a short period of time. Therefore, the experimental research results show that if certain intelligent means and technologies can be used to realize the design of the network security early warning system, compared with the traditional prevention, the traditional efficiency can be improved by about 32.7%. Through the function of the system will be more perfect, it can simultaneously analyze the problems in various networks at the same time, analyze the consequences of the intrusion at any time, and then a more comprehensive security solution: use security management methods to solve network security problems. Carrying out a comprehensive management, this not only improves the user's access speed, but also enables the operator to increase the security of the data transmitted by the user by 63%, and fully increases the security of the computer network communication data transmission.

References
[1] Celik A , Tetzner J , Sinha K , et al. 5G device-to-device communication security and multipath routing solutions[J]. Applied Network Science, 2019, 4(1):1-24.
[2] Lu W , Zheng X , Xu J , et al. Improving physical layer security and efficiency in D2D underlay communication[J]. Wireless Networks, 2019, 25(8):4569-4584.
[3] Cai Z . Security and confidentiality of network communication using chaotic encryption technology[J]. Journal of Computational Methods in Sciences and Engineering, 2019(5):1-10.
[4] Arfaoui M A , Soltani M D , Tavakkolnia I , et al. Physical Layer Security for Visible Light Communication Systems: A Survey[J]. IEEE Communications Surveys & Tutorials, 2020, PP(99):1-1.
[5] Helma C , Cramer T , Kramer S , et al. Data mining and machine learning techniques for the identification of mutagenicity inducing substrucitures and structure activity relationships of noncongeneric compounds[J]. J Chem Inf Comput, 2018, 35(4):1402-1411.
[6] Oldham, Daniel. Data Mining and Machine Learning to Improve Northern Florida's Foster Care System[J]. Beyond: Undergraduate Research Journal, 2019, 3(1):3-3.
[7] Amin D M , Garg A . Performance Analysis of Data Mining Algorithms[J]. Journal of Computational and Theoretical Nanoence, 2019, 16(9):3849-3853.
[8] He F , Zhang L . Mold breakout prediction in slab continuous casting based on combined method of GA-BP neural network and logic rules[J]. International Journal of Advanced Manufacturing Technology, 2018, 95(9-12):4081-4089.
[9] Li D J , Li Y Y , Li J X , et al. Gesture Recognition Based on BP Neural Network Improved by Chaotic Genetic Algorithm[J]. International Journal of Automation and Computing, 2018, 15(03):1-10.
[10] Sopian, Abdul, Wahab, et al. Cluster Analysis Evaluating PM2.5, Occupation Risk and Mode of Transportation as Surrogates for Air-pollution and the Impact on Lung Cancer Diagnosis
and 1-Year Mortality.[J]. Asian Pacific journal of cancer prevention : APJCP, 2019, 20(7):1959-1965.