An Improved Network Intrusion Detection Based on Deep Neural Network

An Improved Network Intrusion Detection Based on Deep Neural Network

Lin Zhang, Meng Li, Xiaoming Wang, Yan Huang

Chengdu Chengdian Electric Power Engineering Design Co. Ltd, Chengdu 610000, China

Lin Zhang, psplayer@126.com

Abstract. Network intrusion detection is of great significance for network security in Local Area Network (LAN). Traditional methods such as firewalls do not completely protect against attacks on the LAN due to lack of continuous learning. Recently, the ability of convolutional neural networks (CNN) to extract features in the field of computer vision has received extensive attention. CNN can automatically extract effective complex features to adapt to constantly changing environments, which is especially important in network intrusion detection. In this paper, we focus on network security in the LAN. We propose an approach based on CNN to implement intrusion detection in LAN. This approach can effectively identify network attacks and has an accuracy of 98.34% on the KDD99 dataset. The experimental results show that the proposed approach based on the CNN has high accuracy in intrusion detection.

1. Introduction

With the rapid development of the information industry, LAN plays a quite important role in the management of various industries (e.g., enterprise, schools, government). Unfortunately, LAN faces complex and serious security risks. Many security problems in LAN are caused by internal personnel and leaks, not caused by external hackers and viruses. On the other hand, the majority of previous works on network information security focus on a public network or Wide Area Network (WAN). Up to now, far too little attention has been interested specifically targeting security about the LAN. Hence, there is an urgent need to address security issues caused by network abnormal behavior in the LAN.

The usual method of LAN security is physical isolation (e.g., Network Firewall), that is, the internal network is not connected to the public network or WAN. Therefore, in this way, the internal network is not affected by interference and attacks from the external network to a certain extent. However, the method of physical isolation technology can not completely guarantee the security of the LAN, because it is difficult to avoid the internal security risks after resisting the external invasion. A large number of researchers are committed to establishing effective solutions to detect network attacks, malicious network behavior and abnormal network traffic for security monitoring purposes[1–3].

Traditionally, intrusion detection techniques use the rule-based approach[4–7]. The rule is a signature of a malicious program or a description of the malicious behavior, and the program code or network behavior that matches the rule is detected as an attack. From the practical application, rule-based intrusion detection technology can effectively defend against known attacks, but it is helpless for new unknown attacks. Moreover, many cybersecurity researchers have shifted their focus to anomaly-based intrusion detection technology[8–11]. The detection technology mainly uses the statistical machine learning method. The idea is collecting normal program and network behavior data, extracting features, and training the machine learning model (support vector machine, Bayesian network, etc). During the detection phase, program code or network behavior that deviates from the normal value beyond the tolerance is considered a malicious code or network attack behavior. However, the practice shows that the pros and cons of anomaly detection models mainly depend on feature extraction. In the existing research, the feature extraction work is mainly done manually by relevant domain experts, which makes the link rely heavily on expert experience and lacks adaptability in different application scenarios.

However, as the intensity and frequency of current cyber threats continue to rise[12], existing defense mechanisms and approaches are no longer able to protect the network, including LAN, from cyber threats in novel threats[13]. In addition, the traditional network intrusion detection method can only process a small amount of data, and cannot fully utilize a large amount of data for training, leading to a decrease in the accuracy of the inspection.

The neural networks, which is inspired by neurons in the brain, is considered to be a new branch of machine learning. Deep neural networks (DNN) can make full use of large amounts of data, and have good data representation and feature extraction capabilities. Because of its ability to extract high-level features, the neural networks are very suitable for network intrusion detection[14]. The small mutations and even newly developed attacks can be detected by this method[15].

In this paper, the work focuses on network intrusion detection about LAN. More specifically, using CNN to perform intrusion detection on the LAN to defend against internal attacks. Our contributions are as follows:

- We focus on internal security in the LAN, utilizing network intrusion detection method. As far as we know, far too little attention has been paid to the internal security in the LAN.
- The network intrusion detection model based on CNN is proposed in this paper to enhance the ability of feature extraction and improve accuracy.
- In order to better extract features through CNN and make full use of raw data, we reshape the data into image format.

The rest of this paper is scheduled as follows. In section 2, we review previous related works. We elaborate in detail our method with CNN-based intrusion detection in Section 3. Section 4 shows the results of the experiment. Finally, we conclude in Section 5.

2. Related work

Network intrusion detection has been studied by many researchers, which is discussed next. Current methods can be classified as machine learning-based and deep learning based. We also review some related work on deep convolutional neural networks.

2.1 Network intrusion detection

Intrusion detection technology was first proposed by Anderson[16] in 1980 and has been the focus of research in the field of network security since then. With the development of machine learning, researchers have found that utilizing machine learning methods to train a large number of intrusion detection data can more effectively improve the accuracy of network intrusion detection. Xian et al.[17] and Tang et al.[18] utilized the SVM-based intrusion detection technology to achieve higher accuracy with less prior knowledge and less training time. Li et al.[19] utilized the improved KNN algorithm to train data that can be effectively detected with a small amount of data. The method has a high detection rate and an abnormality of a low false alarm rate. Tsai et al.[20] combined k-means and KNN machine learning algorithms for intrusion detection. In this method, cluster centers of attacks are performed by k-means, and finally, the intrusion detection is performed by the KNN classifier. Experimental results show that this method is superior to SVM-based and KNN-based models.

Traditional machine learning-based methods have achieved good results in network intrusion detection, but they also have limitations: 1) The performance depends on feature engineering; 2) Unable to process large amounts of data; 3) The ability to learn independently is lacked. In order to solve these problems, neural network-based intrusion detection technology has received extensive attention. Compared with manual pre-designed statistical features, deep learning neural networks can automatically extract features from the raw data[21]. In addition, deep learning neural network algorithm has a better effect than traditional machine learning method in dealing with big data. Gao et al.[22] applied deep belief networks to network intrusion detection, achieving better performance than using other machine learning methods. Staudemeyer[23] firstly applied long short-term memory recurrent neural networks to network intrusion detection, which is very suitable for classifying high-frequency attacks. In addition, experiments by Kim et al.[24] and Elsherif[25] Tom et al. and Jam show that LSTM and RNN are very effective for network intrusion detection. Raman et al.[26] utilized the features selected by the hypergraph-based feature selection technique to train a probabilistic neural network in network intrusion detection. The experimental results show that this method increases the detection rate of fewer frequency attacks.

2.2 Convolution neural networks

The neural network consists of three main categories: the input layer, the hidden layer, and the output layer. However, the hidden layer of the CNN is formed by a series of convolutional, activation (non-linear), pooling (downsampling), and fully connected layers. Because of this structure, CNN is different from DNN and RNN. The main differences are: 1) local connections. The layers are not fully connected but partially connected; 2) weight sharing. The weights of the connections between a subset of neurons in the same layer are shared. 3) Subsampling. A convergence layer is periodically inserted between successive convolutional layers. Therefore, the number of weight for CNN is relatively less than other neural networks, which is more suitable for network intrusion detection.

The research by Vinayakumar et al.[27] shows that CNN and it's variant architecture are better than classic machine learning classifiers in network intrusion detection. Experimental results show that one-dimensional convolution in CNN has high accuracy in network intrusion detection. Moreover, experiments by Liu et al.[28] show that the intrusion detection model based on CNN has a high detection rate and accuracy. It also proves the feasibility of applying CNN in highly-intrusion detection. CNN–based approaches show its powerful feature extraction ability in network intrusion detection and our work is also under this powerful framework.

3. Our approach

In this paper, CNN-based intrusion detection in the LAN is proposed. Due to the powerful feature extraction capabilities of CNN, we apply it to network intrusion detection. CNN can effectively extract the spatial features of image data. Therefore, in order to make full use of the characteristics of CNN, we finally convert raw data into image data. After that, the convolutional neural network can be used to classify the image to detect the input data. The architecture is shown in **Figure 1.**

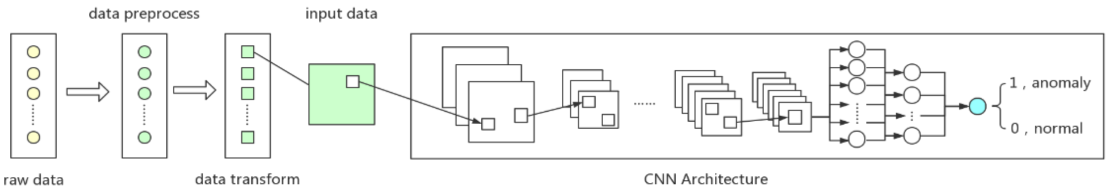

Figure 1. Proposed architecture.

Our approach is briefly elucidated as follows. First, the raw data is preprocessed to obtain standard data. Then, the standard data is converted to image data. Finally, the image data is classified by the CNN model. The details of each step are as follows.

3.1 Data preprocess

KDD99[29] is one of the most widely used intrusion detection datasets. The dataset consists of five million records, including one normal type and four types of attacks (e.g., DoS, Probe, U2R, R2L). Each record in the KDD99 dataset identifies a network connection, represented by 41 features. Among these 41 features in KDD99 dataset, there are 38 numerical features and 3 categorical features.

For numerical features, the size of the numerical features varies greatly due to the difference in dimensions. Therefore, normalization is required to eliminate the effects of different dimensions. For each numerical feature $X=\{x_1,\ x_2,\ \cdots,\ x_n\}$ in KDD99 shall be given as:

$$x_i=\frac{x_i-\min(X)}{\max(X)-\min(X)},\ x_i\in X,\ i=1,2,\cdots \tag{1}$$

Where, x_i is the original input of feature X , $\max(X)$ and $\min(X)$ are the maximum and minimum of the feature, respectively.

For categorical features, we consider digitizing them. The one-hot encoding and dummy variable encoding are adopted for KDD99. If one-hot encoding is adopted for all the category features, one record will be 122 dimensions in KDD99. For convenience in data transform, we consider using dummy variable encoding for one (e.g., protocol_type) of the categorical features. After the pre-processing step, the number of dimensions of the feature has been expanded from 41 to 121.

3.2 Data transform

The convolutional neural network has a powerful capability of feature extraction in the field of computer vision. For computers, the essence of images is the array of pixel values. Based on this inspiration, we transform each of the data-preprocessed records of KDD99 with a 121-dimensional record into a 11x11x1 array. The operation is shown in **Figure 2.**

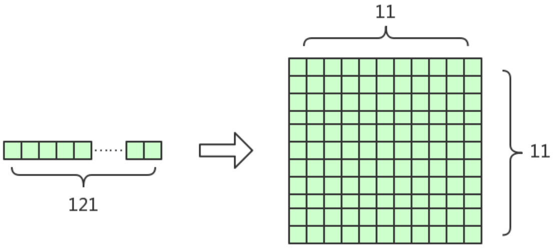

Figure 2. Data transform.

Since the pixel values of the image range from 0 to 255, the attribute of each record shall be given as:

$$p_i=x_i\times 255 \tag{2}$$

Where, p_i is the element of the array. Then, the result will be used as input to a CNN model.

Figure 3. shows that a normal input and an anomaly input in KDD99 in the data. After data transform, normal data and abnormal data are different.

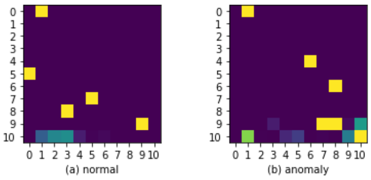

Figure 3. Normal and anomaly data inputs.

3.3 CNN Architecture

After the data is preprocessed and transformed, the raw data becomes an image with 1 channel. Then, we use VGG19 neural network[30] for training. We modify the last few layers of the VGG19 neural network for adapting to our task. Steps as follows. First, flatten the output after the last pool layer of VGG19. Then pass the result of the previous step through a dense layer with the rule activation function. Its output is a 128-dimensional vector. Finally, the previous result is passed through a dense layer with the sigmoid activation function. The one-dimensional result is output.

Since this is a two-classification task, we use binary cross entropy is given as the loss function C:

$$C = -\frac{1}{n}\sum_{X}[y\ln a + (1-y)\ln(1-a)] \tag{3}$$

$$a = \sigma(z) = \frac{1}{1+e^{-z}} \tag{4}$$

$$z = \sum_{i=1}^{n} w_i x_i + b \tag{5}$$

Where, n is the number of training data. And y is the expected output. $\sigma(z)$ is the sigmoid function. In addition, the neural network optimizer uses the RMSprop algorithm.

4. Experiment results

In our experiments, the hyper-parameter settings are as follows: batch-size=512, epoch=800, learning-rate =0.001, decay-factor=0.9. We observed the change in accuracy by setting the different number of epoch in our experiments. It can be seen from Figure 4. that the accuracy rate increases as the number of epoch increases. When epoch reaches around 800, the accuracy rate stops rising. The evaluation result in the test set we have divided from KDD99 shows that the accuracy rate is 98.34%. The recall is 90.64%. The precision is 99.20%.

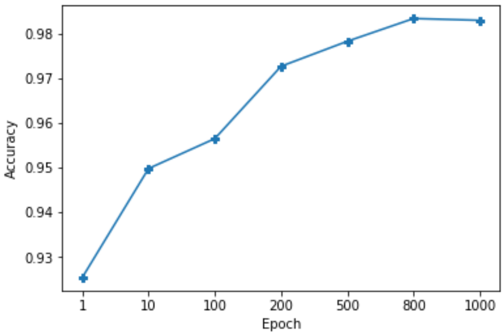

Figure 4. The accuracy of our method in different epoch.

With the condition of using the same dataset, **Figure 5.** Shows that the comparison experiments between our proposed method and the traditional machine learning algorithms including multilayer perceptron (MLP), nearest neighbor (NN), decision tree (DT). It can be seen that the accuracy of the method we proposed is higher than that based on traditional machine learning algorithms.

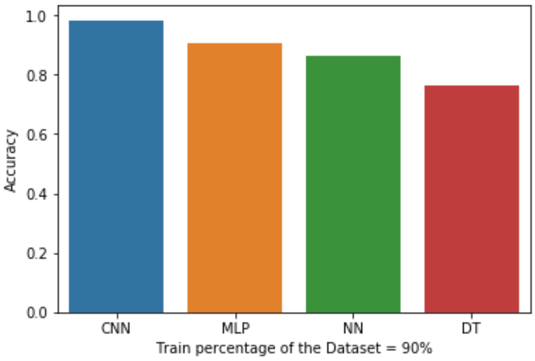

Figure 5. The accuracy of CNN-based and ML-based.

We use different scales on the KDD99 dataset as a training set and the rest as a test set. In **Figure 6.**, as the training set increases, the accuracy of CNN-based methods continues to rise, while the accuracy of machine-based learning methods does not. Compared with traditional machine learning methods, CNN can make full use of big data.

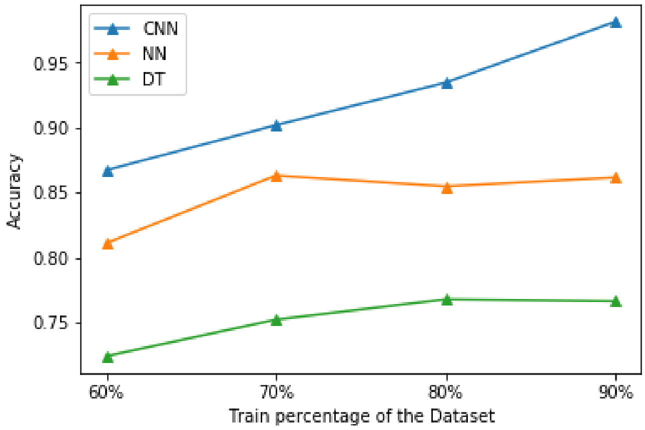

Figure 6. Comparison of different sizes in the training data.

The above experimental results show that the convolutional neural network is superior to the machine learning methods in terms of feature extraction ability. The accuracy of traditional machine learning methods is lower than that based on convolutional neural networks because the representation of normal and abnormal data after feature extraction is similar at low levels. Convolutional neural networks can extract complex high-level features from these similar low-level features. **Figure 7.** shows the visualization of the inter-layer features in the neural network after the data transform.

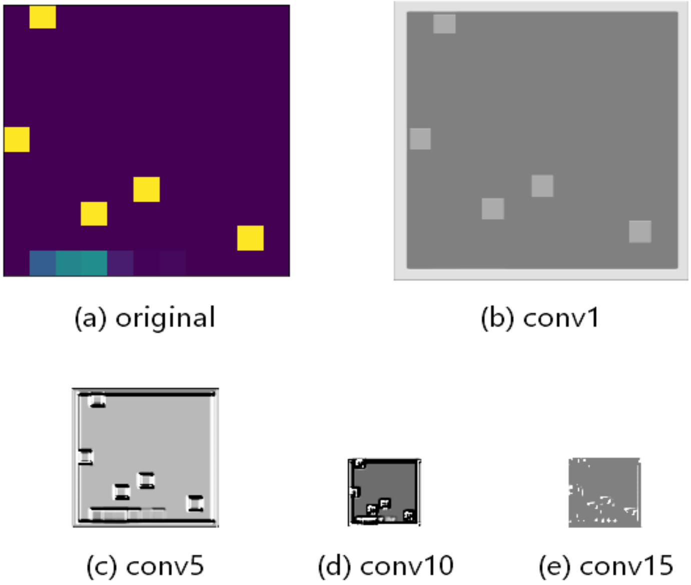

Figure 7. Visualization of inter-layer features in VGG19.

This is a feature extraction scheme that does not require manual design features. Therefore, CNN can continuously adapt to the intrusion detection of the changed network environment in the LAN.

5. Conclusion

In this paper, we focus on network security on the LAN. However, traditional intrusion detection methods have limitations in the LAN. The performance of current machine learning algorithms depends on feature engineering. Machine learning algorithms cannot take advantage of large amounts of data. In addition, machine learning algorithms lack the ability to adapt to independent learning. In order to overcome these limitations, this paper proposes a CNN-based intrusion detection method in LAN. In order to be better extracted for feature by the CNN, the result is converted into a 11x11x1 array as the input of the neural network after we preprocessed the data. Finally, the experimental results show that the proposed CNN-based method is superior to the machine learning-based method and has good effects in the field of network intrusion detection. Therefore, a solution with great possibilities is provided to protect the security of the network in the LAN.

References

[1] Chandola V, Banerjee A and Kumar V 2009 Anomaly Detection: A Survey *ACM Comput. Surv.* **41** 15:1–15:58

[2] Bhuyan M H, Bhattacharyya D K and Kalita J K 2014 Network Anomaly Detection: Methods, Systems and Tools *IEEE Communications Surveys & Tutorials* **16** 303–36

[3] Ahmed M, Naser Mahmood A and Hu J 2016 A survey of network anomaly detection techniques *Journal of Network and Computer Applications* **60** 19–31

[4] Denning D E 1987 An Intrusion-Detection Model *IEEE Transactions on Software Engineering* **SE-13** 222–32

[5] Wenke Lee, Stolfo S J and Mok K W 1999 A data mining framework for building intrusion detection models *Proceedings of the 1999 IEEE Symposium on Security and Privacy (Cat. No.99CB36344)* 1999 IEEE Symposium on Security and Privacy (Oakland, CA, USA: IEEE Comput. Soc) pp 120–32

[6] Tiwari M, Arya K V, Choudhari R and Choudhary K S 2009 Designing Intrusion Detection to Detect Black Hole and Selective Forwarding Attack in WSN Based on Local Information *2009 Fourth International Conference on Computer Sciences and Convergence Information Technology* 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology (Seoul, Korea: IEEE) pp 824–8

[7] Littler T, Wang H F, Yang Y, McLaughlin K and Sezer S 2013 Rule-based intrusion detection system for SCADA networks *2nd IET Renewable Power Generation Conference (RPG 2013)* 2nd IET Renewable Power Generation Conference (RPG 2013) (Beijing, China: Institution of Engineering and Technology) pp 1.05-1.05

[8] Hu W, Liao Y and Vemuri V R Robust Anomaly Detection Using Support Vector Machines 7

[9] Heller K A, Svore K M, Keromytis A D and Stolfo S J One Class Support Vector Machines for Detecting Anomalous Windows Registry Accesses 8

[10] Kruegel C, Mutz D, Robertson W and Valeur F 2003 Bayesian event classification for intrusion detection *19th Annual Computer Security Applications Conference, 2003. Proceedings.* 19th Annual Computer Security Applications Conference, 2003. (Las Vegas, Nevada, USA: IEEE) pp 14–23

[11] García-Teodoro P, Díaz-Verdejo J, Maciá-Fernández G and Vázquez E 2009 Anomaly-based network intrusion detection: Techniques, systems and challenges *Computers & Security* **28** 18–28

[12] Gauthama Raman M R, Somu N, Kirthivasan K, Liscano R and Shankar Sriram V S 2017 An efficient intrusion detection system based on hypergraph - Genetic algorithm for parameter optimization and feature selection in support vector machine *Knowledge-Based Systems* **134** 1–12

[13] B S and K M 2019 Firefly algorithm based feature selection for network intrusion detection *Computers & Security* **81** 148–55

[14] Vinayakumar R, Soman K P and Poornachandran P 2017 Evaluating effectiveness of shallow and deep networks to intrusion detection system *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* pp 1282–9

[15] Kang M-J and Kang J-W 2016 Intrusion Detection System Using Deep Neural Network for In-Vehicle Network Security ed T Tang *PLOS ONE* **11** e0155781

[16] Anderson J P Computer Security Threat Monitoring and Surveillance 56

[17] Xian R, Chunxi D and Shaoquan Y 2003 An Intrusion Detection System Based on Support Vector Machine *Journal of Software*

[18] Tang C, Xiang Y, Wang Y, Qian J and Qiang B 2016 Detection and classification of anomaly intrusion using hierarchy clustering and SVM: Detection and classification of anomaly intrusion *Security and Communication Networks* **9** 3401–11

[19] Li Y and Guo L 2007 An active learning based TCM-KNN algorithm for supervised network intrusion detection *Computers & Security* **26** 459–67

[20] Tsai C-F and Lin C-Y 2010 A triangle area based nearest neighbors approach to intrusion detection *Pattern Recognition* **43** 222–9

[21] Dahiya P and Srivastava D K 2018 Network Intrusion Detection in Big Dataset Using Spark *Procedia Computer Science* **132** 253–62

[22] Gao N, Gao L, Gao Q and Wang H 2014 An Intrusion Detection Model Based on Deep Belief Networks *2014 Second International Conference on Advanced Cloud and Big Data* pp 247–52

[23] Staudemeyer R C 2015 Applying long short-term memory recurrent neural networks to intrusion detection *South African Computer Journal* **56**

[24] Kim J, Kim J, Thu H L T and Kim H 2016 Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection *2016 International Conference on Platform Technology and Service (PlatCon)* pp 1–5

[25] Elsherif A 2018 Automatic Intrusion Detection System Using Deep Recurrent Neural Network Paradigm *Journal of Information Security and Cybercrimes Research*

[26] Raman M R G, Somu N, Kirthivasan K and Sriram V S S 2017 A Hypergraph and Arithmetic Residue-based Probabilistic Neural Network for classification in Intrusion Detection Systems *Neural Networks* **92** 89–97

[27] Vinayakumar R, Soman K P and Poornachandran P 2017 Applying convolutional neural network for network intrusion detection *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* pp 1222–8

[28] Liu Y, Liu S and Zhao X 2018 Intrusion Detection Algorithm Based on Convolutional Neural Network *DEStech Transactions on Engineering and Technology Research*

[29] Lee W and Stolfo S J 1998 Data Mining Approaches for Intrusion Detection *Proceedings of the 7th Conference on USENIX Security Symposium - Volume 7* SSYM'98 (Berkeley, CA, USA: USENIX Association) pp 6–6

[30] Simonyan K and Zisserman A 2014 Very Deep Convolutional Networks for Large-Scale Image Recognition *arXiv:1409.1556 [cs]*