Dean&Francis
ISSN 2959-6157

The Advancements of Deep Learning Approaches for Abnormal Network Traffic Detection

Kaiming Chen[1, 2]

[1]College of International Collaboration, Dalian Maritime University, Dalian, China
[2]College of International Collaboration, University of Houston, Houston, America
kchen50@CougarNet.UH.EDU

Abstract:

For network technology, abnormal traffic detection is crucial to ensuring network security. In order to address abnormal traffic, remarkable results have been achieved when deep learning methods are applied to network abnormal traffic detection. This paper first introduces the significance of analyzing network traffic and the impact of abnormal traffic and then focuses on discussing abnormal traffic detection methods based on deep learning, which are divided into three major kinds: Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN). For DNN, it covers Multilayer Perceptron Based DNN, Multi-layer DNN with Adaptive Parameter, and Two-Level Detection System. CNN-related methods include Wavelet-Enhanced CNN and hybrid architectures with RNN or autoencoder. RNN-based approaches involve Hierarchical Hybrid RNN, Session Gate RNN, etc. The discussion section is about the challenges associated with deep learning, including real-time requirements, low interpretability, and the need to address diverse network attacks. It also proposes future directions for deep learning, such as online learning, improving model interpretability, and adversarial training. Finally, the conclusion section states that this article provides a comprehensive review, helping readers better understand the development and future directions of deep learning in abnormal traffic detection.

Keywords: Machine learning; normally traffic detection; deep learning.

1. Introduction

In the field of internet technology, traffic is the volume of data transmitted within a computer network. The index measures the usage and load of the network. By analyzing traffic, network performance can be improved, network architecture can be optimized, network security can be ensured, and network issues

can be diagnosed. Despite the convenience of the internet in people‘s daily lives, security challenges have become more prominent. A key method used in cyberattacks is abnormal traffic, which is a data flow that differs from normal network behavior patterns. Traffic that is abnormal can cause sudden spikes or drops in volume within a short time frame, abnormal surges in bandwidth consumption, and data breaches.

Effectively dealing with security threats arising from the development of the Internet necessitates implementation of a dynamic network management system that effectively manages network traffic, including anomaly traffic detection. Anomaly traffic detection technologies can be classified into four types based on their technical approaches: statistical-based, data mining-based, machine learning-based, and deep learning-based. The most common of these is the traditional machine learning, which has achieved good results in anomaly traffic detection, but it relies too heavily on manual extraction of traffic features, requiring a significant investment of human and time resources. The application of deep learning methods has been beneficial for the network anomaly traffic detection field in dealing with increasingly sophisticated anomaly traffic.

This field has undergone significant progress and innovation. For instance, Salman et al. proposed a method for efficient and secure detection using deep learning and Internet of Thing (IoT) device identification, and anomaly traffic detection frameworks [1]; the use of semi-supervised deep reinforcement learning was proposed by Dong et al. to detect abnormal traffic on networks [2]; Wang et al. proposed a system that uses deep learning hybrid models to detect abnormal traffic in Software-Defined Network (SDN) [3]; Hwang et al. suggested applying an unsupervised deep learning model and D-PACK‘s abnormal traffic detection technique to the early detection of network traffic anomalies [4]. It can be seen from these works that the use of deep learning methods in network anomaly traffic detection has led to significant progress and innovation. Therefore, it is necessary to conduct a comprehensive review of this area.

This review paper consists of the following three parts. The second segment entitled ‚Method‘ will present a summary of some recent advanced methods for anomaly traffic detection that utilize deep learning, their technical details, and innovative aspects. The third part is Discussion, which will focus on some challenges and shortcomings in this field and future prospects. The fourth part is Conclusion, which will summarize the entire paper.

2. Method

2.1 DNN

2.1.1 Multilayer perceptron-based DNN

Ahmad et al. proposed an anomalous traffic detection mechanism based on Deep Neural Network (DNN) for IoT [5]. This approach automatically deep learns data features through multiple layers. This anomalous traffic detection mechanism is based on DNN as the core and able to automatically learn the data features of the output traffic while processing the network traffic data for IoT. In the study, the input layer receives vectors containing features of network traffic, and outputs classification results of the traffic through multiple hidden layers to determine whether the network traffic is normal or abnormal. In this study, network traffic is accurately classified by using binary classification for the target features of normal and abnormal traffic and adjusting the weights and biases through backpropagation so that the DNN model can learn the feature patterns in the data. This approach can greatly improve the accuracy and efficiency of detecting abnormal traffic.

2.1.2 Multi-layer DNN with adaptive parameter

A model for detecting anomaly traffic in IoT was proposed by Reddy et al. using DNN architecture [6]. This DNN architecture has an input layer, a hidden layer, and an output layer. The input layer is used to receive preprocessed IoT network traffic data; the hidden layer consists of multiple neurons, which can learn the feature patterns in the data; the output layer determines whether input network traffic is normal or anomalous by analyzing the output of the hidden layer. The article uses a back propagation learning algorithm to train DNNs, and a loss function is created to measure the variance between the model‘s predicted results and the actual traffic features. The study uses the optimizer to speed up the convergence of the model during the training process and again uses the dropout method to prevent overfitting.

2.1.3 Two-level detection system based DNN and correlation analysis

A two-layer anomaly detection system was proposed by Gao et al. that uses DNN and association analysis [7]. In this learning method, DNN plays a central role in analyzing and classifying data features. The Apriori association analysis algorithm filters the traffic classified by DNN to reduce the false positive rate. To enable DNN to better process network traffic, this study performed targeted preprocessing on the input data, converting the character features of network traffic data into numerical form to align

ISSN 2959-6157

with the input format of the DNN output layer. After the traffic data is input into the DNN module, the DNN can extract the deep feature information of the data, achieving the classification of network traffic data into normal and abnormal categories. The Apriori association rule algorithm is utilized in this study to create association rules for feature matching of abnormal traffic classified by DNN after the DNN module completes the classification process. The system's accuracy and precision are enhanced by combining DNN and association rule algorithms to optimize the detection foundation based on the DNN model.

2.2 CNN

2.2.1 Wavelet-Enhanced CNN for traffic feature extraction

Wang et al.'s proposed abnormal traffic detection system for Software-Defined Network (SDN) uses deep learning hybrid models [3]. This system has a two-stage detection strategy. In the first stage, it uses traffic statistics features based on switch ports to quickly locate abnormal traffic. In the second stage, an MFDLC (MFDLC) is utilized for detection if abnormal traffic is detected in the first stage. The Convolutional Neural Network (CNN) model and Long Short-Term Memory (LSTM) model are two structures in MFDLC. Wang et al. use the CNN model to extract the spatial features of each decomposed time series. Compared to models that rely solely on the CNN model, MFDLC resolves the limitation of the CNN model, which only captures spatial features. To detect traffic time-series anomalies, the LSTM model is infused with spatial features extracted via CNN. The InSDN dataset experiments have shown that MFDLC achieves significantly better detection accuracy than traditional CNN-based methods.

2.2.2 Hybrid architecture-based CNN and RNN

A design called HAST-NAD was proposed by Wei and Wang, which uses CNN and Recurrent Neural Network (RNN) together [8]. To detect anomalous traffic, this method employs two phases of deep learning, with the CNN model learning spatial features. By converting network traffic into two-dimensional images, the CNN model extracts feature of the spatial structure. CNN extracts feature of the spatial structure by converting network traffic into two-dimensional images. Temporal features are the focus of the RNN in the second stage of learning. This stage detects features in time series by processing the spatial features output by the CNN in the first stage. The design is a deep learning model that uses CNN as the foundation and RNN as a supplementary component instead of being solely based on CNN. HAST-NAD enhances detection rate and accuracy compared to traditional machine learning methods through its innovative deep integration of CNN and RNN.

2.2.3 Hybrid architecture of autoencoder and CNN

The D-PACK abnormal traffic detection mechanism was proposed by Hwang et al., which uses a combination of CNN and autoencoder to detect network traffic anomalies [3]. CNN is in charge of automatically learning the traffic characteristics from the raw network traffic data in this study. In processing the data, the mechanism obtains each traffic packet by sampling it and converts it into a one-dimensional vector for input to CNN. The data features of the flows are then extracted using the convolutional and pooling layers of CNN model. After the features are extracted by CNN, the features are used in the training of the autoencoder to determine whether the input traffic is abnormal or not. By innovating this deep learning method based on CNN and autoencoder, Hwang achieves a high detection accuracy and is able to identify abnormal traffic effectively.

2.3 RNN

2.3.1 Hierarchical hybrid RNN

Ullah and Mahmoud proposed a deep learning method based on RNN combined with CNN and other technologies, achieving efficient and accurate detection of abnormal traffic in IoT network traffic [9]. The use of abnormal traffic detection models in this study was based on RNN and its variants, including LSTM, BiLSTM, and Gate Recurrent Unit (GRU). These models can process network traffic data, which is time-series data, accurately identify the characteristics of this data, and the abnormal traffic. In order to improve the model's ability to learn features of data, the study combined RNN with CNN. In the deep learning method constructed in this study, the data stream undergoes feature extraction by CNN, and then these features are input into RNN. The accuracy of anomaly traffic detection can be improved by using this innovative combination to take full advantage of the strengths of both RNN and CNN.

2.3.2 Session gate RNN

Zhao et al. developed a new deep learning technique called ERNN that is based on RNN, resulting in a model that is better than the traditional RNN structure (LSTM) [10]. As the core unit of RNN, LSTM is used in the study and a Session Gate is added to this structure. ERNN's Session Gate is the most innovative component, simulating dynamic interference while processing time series data in abnormal traffic. Four actions make up the Session Gate, and their state updates are influenced by RNN's processing logic in different network phenomena, while

also influencing RNN's ability to read time series data of anomaly traffic. The Session Gate aims to improve the reliability of RNN by reading network traffic time series data against network-induced phenomena and decreasing its reliance on local packets. ERNN has the same process as traditional RNN in handling traffic time series and has a similar model architecture, so this deep learning method represents an improvement to the training process of RNN.

2.3.3 Multi-class classifiers based RNN and CNN

Fotiadou et al. proposed a multi-class classifiers combined a classic variant of RNN: LSTM, with CNN [11]. In this deep learning method, the LSTM module is mainly used to process log data containing network traffic, time series, and text information. It can utilize the time series modeling capabilities of RNN and capture long-term dependencies in time series through memory gate mechanisms, which is a key component in solving the problem of anomaly traffic detection. The study also uses CNN, which extracts spatial features of traffic through convolution and pooling, independent of the temporal dependencies in traffic, that is different from RNN. By forming a multi-class classifier through the integration of LSTM and CNN, this deep learning method enhances accuracy while taking advantage of its strong modeling capabilities on temporal data of traffic.

3. Discussion

3.1 Challenges

3.1.1 Real-time and low-latency requirements

Detecting abnormal network traffic requires complete traffic data, which can be detected by deep learning methods after getting data features from the traffic. Blocking abnormal traffic after extracting traffic features will lead to the expansion of the scope of the impact of abnormal traffic, and it will also lead to problems that are about real-time. On the other hand, although the hierarchical detection of deep learning methods can be accelerated by rough inspection of port information, the extensive inspection still requires processing a significant amount of traffic data, which leads to an increase in latency. Deep learning's ability to handle large-scale traffic is still lacking. In situations where there are a large amount of traffic, it is difficult for deep learning methods to process large amounts of traffic data in real time and with low-latency, which can lead to missed inspections of some parts of the traffic and those may lead to abnormal conditions.

3.1.2 Limited interpretability

Deep learning models have more complex internal mechanisms, which makes it difficult for people to analyze and understand the decision-making process of deep learning methods. This poor interpretability limits the application of deep learning in domains where there is high risk, such as medical diagnosis and autonomous driving. Being less interpretable also means that there is a risk of leading to erroneous results, and in fact many fields have concerns about the application of deep learning models, both because the models cannot give enough information and because there are security issues with deep learning methods. Interpretability is a useful property, and having better interpretable models will have irreplaceable advantages in many application scenarios. Therefore, improving the interpretability of deep learning models is important to facilitate the application of deep learning methods.

3.1.3 Diversification of traffic attacks

Deep learning models strongly depend on training data, so they are easy to suffer adversarial sample or data poisoning attacks, which leads to the failure of deep learning models to detect abnormal traffic. Besides, attackers can circumvent the detection of the deep learning methods by employing methods such as adversarial generative networks. Although the detection of traffic through deep learning methods is effective in extracting data features, the diversity of traffic attack methods can make the deep learning model to misjudge abnormal traffic as normal traffic. After being attacked by data poisoning, the deep learning model will also misjudge normal traffic as abnormal traffic. Therefore, deep learning-based anomalous traffic detection methods need to be more robust and combine more deep learning methods to better defend traffic attack means.

3.2 Future prospects

3.2.1 Online learning and early detection mechanisms

Deep learning models can be adapted to the new changes in network traffic by updating the parameters of the model in real time through online learning or incremental learning algorithms. Computers can update the models as soon as a new batch of traffic data is received, and this method without retraining the entire dataset reduces the detection delay. Designing early detection algorithms based on parts of traffic data can initially determine whether this batch of traffic will be abnormal traffic when a small amount of traffic data is acquired. Through analyzing the preliminary data in real time, the anomalous trends can be detected as soon as they are detected.

3.2.2 Enhancing the understanding of deep learning

Deep learning modeling process can be used to divide

ISSN 2959-6157

methods for improving interpretability into three types. Pre-modeling methods focus on data preprocessing and presentation, using interactive figures and tables for data visualization to display high-dimensional data distributions, providing data for subsequent modeling, and also helping to understand the model's decision-making process. During modeling, models with inherent interpretability can be constructed by designing transparent structures to enhance interpretability. Post-modeling methods target black-box models and include hidden layer analysis, proxy models, and sensitivity analysis.

3.2.3 Model integration and adversarial training

Use multiple different deep learning models for detection and then it can reduce the false alarm rate of a single model by judging their results. This needs to be done by deep learning model integration and fusion. Furthermore, with the aim of improving the robustness of the deep learning method for detecting abnormal traffic, adversarial training techniques can be used to distinguish between normal and abnormal samples and enhance the robustness of the deep learning model against adversarial attacks. Regularly updating the training data of the model to include the latest attack samples can adapt the model to the constantly updated traffic attacks.

4. Conclusion

This research reviews the realization methods related to the detection of abnormal network traffic through various deep learning models. Recent methods to apply deep learning to traffic detection include DNN, CNN and RNN based traffic detection mechanisms. However, in terms of detecting traffic, deep learning has deficiencies in real-time, low-latency requirements, less interpretability, and ways of defending traffic attacks. These shortcomings can be mitigated by incorporating online learning frameworks, enhancing data preprocessing techniques, and integrating hybrid models that combine the strengths of multiple approaches.

References

[1] Salman O, Elhajj IH, Chehab A, et al. A machine learning based framework for IoT device identification and abnormal traffic detection. Transactions on Emerging Telecommunications Technologies, 2022, 33(3): e3743.

[2] Shi D, Xia Y, Peng T. Network abnormal traffic detection model based on semi-supervised deep reinforcement learning. IEEE Transactions on Network and Service Management, 2021, 18(4): 4197–4212.

[3] Wang K, Fu Y, Duan X, et al. Abnormal traffic detection system in SDN based on deep learning hybrid models. Computer Communications, 2024, 216: 183–194.

[4] Hwang RH, Peng MC, Huang CW, et al. An unsupervised deep learning model for early network traffic anomaly detection. IEEE Access, 2020, 8: 30387–30399.

[5] Ahmad Z, Shahid Khan A, Nisar K, et al. Anomaly detection using deep neural network for IoT architecture. Applied Sciences, 2021, 11(15): 7050.

[6] Reddy DKK, Behera HS, Nayak J, Vijayakumar P, Naik B, Singh PK. Deep neural network based anomaly detection in Internet of Things network traffic tracking for the applications of future smart cities. Transactions on Emerging Telecommunications Technologies, 2021, 32: e4121.

[7] Gao M, Ma L, Liu H, et al. Malicious network traffic detection based on deep neural networks and association analysis. Sensors, 2020, 20(5): 1452.

[8] Wei G, Wang Z. Adoption and realization of deep learning in network traffic anomaly detection device design. Soft Computing, 2021, 25(2): 1147–1158.

[9] Ullah I, Mahmoud QH. Design and development of RNN anomaly detection model for IoT networks. IEEE Access, 2022, 10: 62722–62750.

[10] Zhao Z, Li Z, Jiang J, et al. ERNN: Error-resilient RNN for encrypted traffic detection towards network-induced phenomena. IEEE Transactions on Dependable and Secure Computing, 2023.

[11] Fotiadou K, Velivassaki TH, Voulkidis A, et al. Network traffic anomaly detection via deep learning. Information, 2021, 12(5): 215.