Detecting Abnormal Traffic Patterns with Attention and Big Step Convolution Techniques

[1] Juttu Harshini, [2] N.Aishu, [3] B.Sreeja, [4] B. Vanaja
[1,2,3]UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Assistant Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
vanajacse@gmail.com

Abstract:

Abnormal traffic detection is critical to network security and quality of service. However, the similarity of features and the single dimension of the detection model cause great difficulties for abnormal traffic detection, and thus a big-step convolutional neural network traffic detection model based on the attention mechanism is proposed. Firstly, the network traffic characteristics are analyzed and the raw traffic is preprocessed and mapped into a two-dimensional grayscale image. Then, multi-channel grayscale images are generated by histogram equalization, and an attention mechanism is introduced to assign different weights to traffic features to enhance local features. Finally, pooling-free convolutional neural networks are combined to extract traffic features of different depths, thus improving the defects such as local feature omission and overfitting in convolutional neural networks. The simulation experiment was carried out in a balanced public data set and an actual data set. Using the commonly used algorithm SVM as a baseline, the proposed model is compared with ANN, CNN, RF, Bayes and two latest models. Experimentally, the accuracy rate with multiple classifications is 99.5%. The proposed model has the best anomaly detection. And the proposed method outperforms other models in precision, recall, and F1. It is demonstrated that the model is not only efficient in detection, but also robust and robust to different complex environments

Key words: Malware, cyberattacks, Machine learning, Deep learning, Optimal Ensemble Learning, data preprocessing, Least Square Support Vector Machine, kernel extreme learning machine, Regularized random vector functional link neural network, hunter-prey optimization.

1.INTRODUCTION

Internet technology is widely used in all walks of life, and has strongly contributed to the development of economy and society. However, as the current mainstream network security and defense technologies still have many shortcomings, the huge application requirements also make the security configuration of the entire network becomes particularly complex, resulting in the entire network facing the threat of extremely vulnerable to attacks. At the same time, due to the openness of the TCP/IP network architecture, computer viruses spread more widely through disguise, which affects the normal operation of the network and causes social and economic downturn. How to take effective methods to analyze data information to predict the current network development, find abnormalities and take appropriate handling measures is of great significance to maintain network security [1]. .

Anomalous traffic detection can be achieved with the help of network traffic classification. According to its core idea there are mainly the following approaches: port-based [2], deep packet detection based [3], and machine learning based [4]. Machine learning consists of traditional machine learning and deep learning. In the early days, when the Internet was small and the traffic types were simple, the first two methods had stable performance and achieved good classification results [5], [6], [7]. However, with the continuous emergence of new Internet applications, traffic types are increasing and traffic components are becoming more complex, which reduces the classification effect. Machine learning improvement methods are proposed to address the limitations of the above methods. Machine learning is to extract statistical features of network traffic and classify them with reliable efficiency and high accuracy. It also has a wide range of application prospects.

The overall process of network traffic classification consists of collecting data sets, generating normalized data, data pre-processing, feature extraction, training models and classification. Traditional machine learning classification is based on different algorithms to select the optimal subset of features that are similar to the full feature results for classification. This approach relies on feature selection, which can directly affect the classification results and cannot cope with the evolution of modern network traffic. And traditional machine learning models cannot represent the complex relationships between individual features. As a result, deep learning becomes the optimal algorithm for solving network traffic classification, which performs high performance in dynamic and challenging traffic classification environments.

In the past few years, deep learning has had several studies working on network traffic classification [8], [9], [10]. These studies provide the performance-enhancing feasibility of deep learning techniques for handling traffic classification tasks, but reveal that deep learning is still in its infancy for network anomaly detection research. In contrast to machine learning, deep learning not only enables classification of network traffic by automatically extracting structured and complex features and feeding them directly into a training classifier, but also represents complex nonlinear relationships between features. In summary, the anomalous traffic detection model for network security defense has improved in terms of improvement and practicality. However, there are still many problems: First, the classification results are poor for traffic information with similar attribute characteristics. Second, the structure of the anomaly network detection model is inflexible and cannot extract features in multiple dimensions and fields of view, which reduces the accuracy of network traffic classification to a certain extent. Third, multiple pooling using convolution neural networks has the potential for information loss, which can make the sequence less relevant.

To overcome the above challenges and difficulties, this paper makes the following contributions:

• In this paper, we propose an Attention and Big Step Convolution Neural Network (ABS-CNN) model based on the attention mechanism [11]. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. To solve the problems such as similar features leading to worse

classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. Experiments show that the model with enhanced features has higher classification accuracy and better robustness.

• In this paper, we use histogram equalization to solve the problem of single model dimensionality. The traffic data is first processed into grayscale images and then the images are histogram equalized. Combined with improved multi-channel convolution to automatically extract and fuse multi-field fine-grained features. The experiments show that the traffic with histogram equalization performed is relatively well-defined, which results in better model detection performance and better robustness.

• To address the reduced correlation of traffic sequences due to pooling, the traffic features are extracted by combining big-step convolution. And big-step convolution is also called stepwise convolution. Stepwise convolution preserves the sequence-related features extracted by the convolution layer and reduces the harm of accuracy loss due to information loss.

This paper is divided into five parts. Section I briefly describes the research background and main contributions of this paper. Section II will describe and summarize the current research development. Section III carries out the model introduction and algorithm implementation process. Detailed experiments and analysis of the results will be carried out in Section IV. Finally, Section V will analyze and summarize the model proposed in this paper and point out some possible future research directions.

2. LITERATURE SURVEY

This literature survey reviews several studies on network traffic anomaly detection and its application in intelligent transportation systems (ITS) and road safety, emphasizing the convergence of digital security techniques with physical infrastructure improvements. A key focus of the studies is the use of advanced machine learning and deep learning models to enhance anomaly detection capabilities, particularly in dynamic environments like transportation systems.

For example, the study by Changpeng Ji and Wei Dai (2024) on "Network Traffic Anomaly Detection Based on Spatiotemporal Feature Extraction and Channel Attention" primarily targets network traffic anomaly detection by leveraging spatiotemporal features and attention mechanisms. However, the paper also mentions the development of a road accident prediction model using the Random Forest algorithm, which appears to be an unrelated application. This misalignment suggests a possible error in associating the correct contribution with the study's title or reflects a broader use of machine learning techniques in road safety, influenced by their focus on feature extraction methods.Similarly, Ling Xing and Shiyu Li (2024) provide a review of anomalous behavior detection in social networks. Their work, which focuses on identifying suspicious activity such as fake accounts or unusual behavior patterns, also discusses the design of an intelligent transportation system aimed at pedestrian safety. This suggests that their network anomaly detection techniques could be adapted for use in transportation systems, particularly in ensuring pedestrian safety by analyzing movement and behavior in public spaces. This shows the potential for cross-domain applications of anomaly detection techniques in enhancing both digital security and real-world safety.The study by Xinjian Zhao and Weiwei Miao (2024) introduces a sparse transformer-based model for abnormal traffic detection, combining feature fusion techniques to improve detection accuracy. This model highlights the convergence of network anomaly detection methods with traffic management strategies, ultimately contributing to safer road environments. Similarly, Shaojun Sheng and Xin Wang (2023) propose using chaotic neural networks for network traffic anomaly detection. They also explore the application of cloud computing algorithms to assess road accident risks, indicating a potential overlap between data-driven anomaly detection and road safety efforts.Liu Wu, Niandong Liao, and Yanqi Song (2023) present a network traffic classification model based on attention mechanisms and spatiotemporal features. Their model aims to detect anomalies in network traffic while also offering a framework for ITS that focuses on real-time vehicle detection and tracking. This model demonstrates how network traffic anomaly detection techniques can be adapted for transportation systems to monitor and manage traffic more effectively, thereby improving road safety.Jie Gao and Keyong Hu (2023) propose a hybrid model combining attention mechanisms and convolutional neural networks (CNNs) to detect abnormal traffic patterns in network traffic. Their approach also incorporates natural language processing (NLP) to analyze traffic accident reports, seeking to identify causal factors behind accidents. This interdisciplinary approach demonstrates how machine learning can not only detect network anomalies but also enhance road safety by extracting actionable insights from accident data to prevent future incidents.

However, despite these advancements, challenges remain. Many models struggle to accurately classify traffic that shares similar attributes, leading to misclassification. Moreover, CNNs, while effective at capturing local patterns, use pooling layers that can reduce the dimensionality of data and result in information loss, particularly in sequential data like network traffic. This information loss can impair the model's ability to detect subtle anomalies. These challenges suggest the need for more sophisticated models that can handle the complexity of modern network traffic and adapt to its evolving nature.

3. PROPOSED METHODOLOGY

• In this paper, we propose an Attention and Big Step Convolutional Neural Network (ABS-CNN) model based on the attention mechanism [11]. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. Experiments show that the model with enhanced features has higher classification accuracy and better robustness.

• In this paper, we use histogram equalization to solve the problem of single model dimensionality. The traffic data is first processed into grayscale images and then the images are histogram equalized. Combined with improved multi-channel convolution to automatically extract and fuse multi-field fine-grained features. The experiments show that the traffic with histogram equalization performed is relatively well-defined, which results in better model detection performance and better robustness.

• To address the reduced correlation of traffic sequences due to pooling, the traffic features are extracted by combining big-step convolution. And big-step convolution is also called stepwise convolution. Stepwise convolution preserves the sequence-related features extracted by the convolution layer and reduces the harm of accuracy loss due to information loss.

Advantages

• An input layer, three convolutional layers, a fully connected layer and an output layer are set in the ABS-CNN model, and a convolutional attention mechanism is introduced to enhance the ability of convolution to extract traffic features.

• In the proposed system, the ablation study is performed by removing each component in turn from the proposed ABS-CNN and comparing it with the ABS-CNN of the complete pair to verify the impact of each component on the model. To examine the effects of attention

mechanism, histogram equalization, and large-step convolution on model performance.

4. EXPERIMENTAL ANALYSIS

Figure 1: Login Page

This figure serves as the initial interaction point for users, emphasizing the authentication process. The presence of separate login options for "SERVICE PROVIDER" and "REGISTER" suggests a system with distinct user roles, each having different levels of access and functionalities. The use of a traffic image as the background reinforces the context of the system. This figure highlights the importance of user authentication and provides a glimpse into the role-based access control within the system.

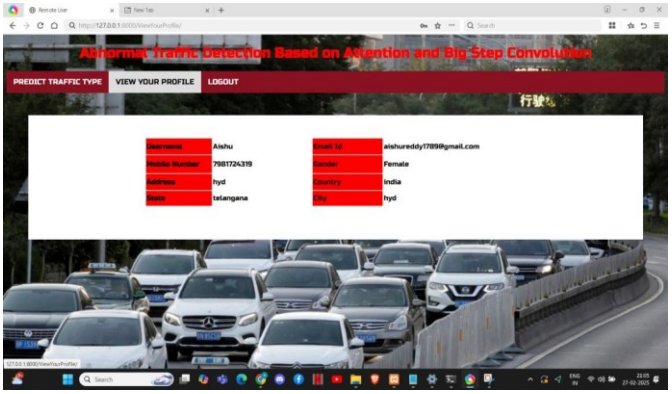

Figure 2: Profile page

This figure displays a user's profile page, which likely allows users to manage their account information. The specific details visible might include the username, contact information, and potentially settings related to traffic data preferences. The presence of a profile page indicates that the system supports user personalization and account management. . management.

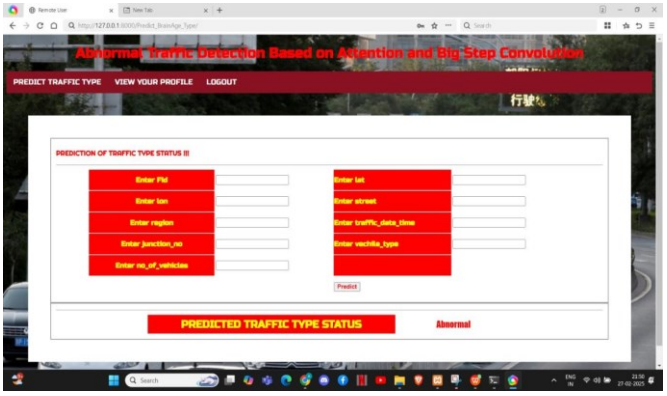

Figure 3: Prediction of traffic type

These figures demonstrate the core functionality of the system – predicting traffic types. They likely show a visualization or textual output indicating whether the analyzed traffic is classified as "normal" or "abnormal." These figures highlight the system's ability to analyze traffic data and provide predictions, which could be crucial in identifying anomalies and potential traffic issues.

Figure 4: Service provider login page

This figure shows the dedicated login page for Service Providers, who likely have administrative access to the system. The login page might have additional security measures or options compared to the regular user login. This distinction emphasizes the different roles within the system and the elevated privileges granted to Service Providers.

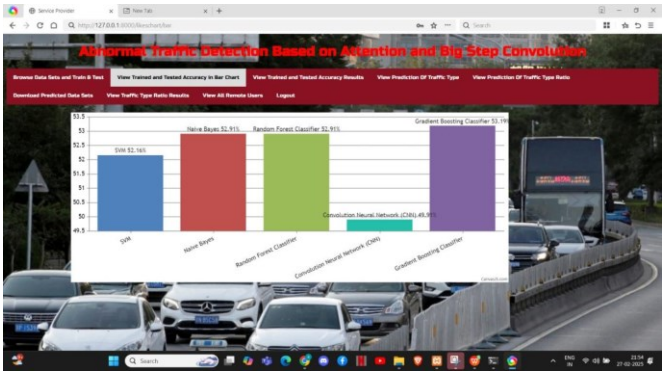

Figure 5: Trained and tested accuracy in bar chart

This figure visualizes the performance of the trained machine learning models. The bar chart likely compares the accuracy of the models on training and testing datasets, providing insights into their generalization ability. This visualization showcases the system's capability to evaluate and visualize model performance, aiding in model selection and optimization.

5. CONCLUSION

This paper proposes an abnormal traffic detection model, ABS-CNN, that leverages attention mechanisms and big-step convolution to overcome challenges posed by similar features and single model structure. Experiments conducted on both publicly available datasets and real-world traffic crawls demonstrate ABS-CNN's superior performance in accuracy, precision, recall, and F1-Score compared to traditional CNN models. ABS-CNN's high sensitivity is evident from its 100% classification accuracy for multiple traffic types, and its efficiency is reflected in shorter training and testing times. The ablation analysis reveals that incorporating attention mechanisms enhances feature differentiation, while histogram equalization and the removal of the pooling layer improve model efficiency and detection performance. The model also shows strong results in detecting encrypted and malicious traffic in real-world environments, indicating its robustness. ABS-CNN successfully adapts to complex environments, showcasing its fine-grained capability to classify encrypted traffic types and detect anomalies effectively. The results confirm that the model not only achieves high accuracy but also maintains operational efficiency, making it a promising solution for practical deployment in abnormal traffic detection systems. Future

research directions include improving data pre-processing tools to avoid sample loss and invalid duplicates, exploring the use of more suitable pre-processing sequences, and investigating spatial and temporal relationships in traffic data to further enhance anomaly detection.

REFERENCES

[1] O. Salman, I. H. Elhajj, A. Kayssi, and A. Chehab, "A review on machine learning–based approaches for internet traffic classification," Ann. Telecommun., vol. 75, nos. 11–12, pp. 673–710, Dec. 2020.

[2] A. Madhukar and C. Williamson, "A longitudinal study of P2P traffic classification," in Proc. 14th IEEE Int. Symp. Modeling, Anal., Simulation, Monterey, CA, USA, Sep. 2006, pp. 179–188, doi: 10.1109/MASCOTS.2006.6.

[3] S. Sen, O. Spatscheck, and D. Wang, "Accurate, scalable in-network identification of P2P P2P traffic using application signatures," in Proc. 13th Int.

Conf. World Wide Web, New York, MY, USA, May 2004, pp. 512–521.

[4] L. Ding, J. Liu, T. Qin, and H. Li, "Internet traffic classification based on expanding vector of flow," Comput. Netw., vol. 129, pp. 178–192, Dec. 2017.

[5] T. Liu, Y. Sun, and L. Guo, "Fast and memory-efficient traffic classification with deep packet inspection in CMP architecture," in Proc. IEEE 5th Int. Conf. Netw., Archit., Storage, Macau, China, Jul. 2010, pp. 208–217, doi: 10.1109/NAS.2010.43.

[6] N. Cascarano, L. Ciminiera, and F. Risso, "Optimizing deep packet inspection for high-speed traffic analysis," J. Netw. Syst. Manage., vol. 19, no. 1, pp. 7–31, Mar. 2011.

[7] G. Aceto, A. Dainotti, W. de Donato, and A. Pescape, "PortLoad: Taking the best of two worlds in traffic classification," in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM), San Diego, CA, USA, Mar. 2010, pp. 1–5, doi: 10.1109/INFCOMW.2010.5466645.

[8] L. Vu, C. T. Bui, and Q. U. Nguyen, "A deep learning based method for handling imbalanced problem in network traffic classification," in Proc. 8th Int. Symp. Inf. Commun. Technol., Dec. 2017, pp. 333–339

[9] P. Wang, F. Ye, X. Chen, and Y. Qian, "Datanet: Deep learning based encrypted network traffic classification in SDN home gateway," IEEE Access, vol. 6, pp. 55380–55391, 2018.

[10] J. H. Shu, J. Jiang, and J. X. Sun, "Network traffic classification based on deep learning," J. Phys., Conf. Ser., vol. 1087, Sep. 2018, Art. no. 062021.

[11] D. Bahdanau, K. H. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, arXiv:1409.0473.

[12] C. Wang, T. Xu, and X. Qin, "Network traffic classification with improved random forest," in Proc. 11th Int. Conf. Comput. Intell. Secur. (CIS), Shenzhen, China, Dec. 2015, pp. 78–81, doi: 10.1109/CIS.2015.27.

[13] Z. Yuan and C. Wang, "An improved network traffic classification algorithm based on Hadoop decision tree," in Proc. IEEE Int. Conf. Online Anal. Comput. Sci. (ICOACS), Chongqing, China, May 2016, pp. 53–56, doi: 10.1109/ICOACS.2016.7563047.

[14] A. V. Phan, M. L. Nguyen, and L. T. Bui, "Feature weighting and SVM parameters optimization based on genetic algorithms for classification problems," Appl. Intell., vol. 46, no. 2, pp. 455–469, Mar. 2017.

[15] B. Schmidt, A. Al-Fuqaha, A. Gupta, and D. Kountanis, "Optimizing an artificial immune system algorithm in support of flow-based internet traffic classification," Appl. Soft Comput., vol. 54, pp. 1–22, May 2017.

[16] S. Dong, "Multi class SVM algorithm with active learning for network traffic classification," Expert Syst. Appl., vol. 176, Aug. 2021, Art. no. 114885.

[17] J. Cao, Z. Fang, G. Qu, H. Sun, and D. Zhang, "An accurate traffic classification model based on support vector machines," Int. J. Netw. Manage., vol. 27, no. 1, Jan. 2017, Art. no. e1962.

[18] D. Md. Farid, N. Harbi, and M. Zahidur Rahman, "Combining Naive Bayes and decision tree for adaptive intrusion detection," 2010, arXiv:1005.4496.

[19] G. D'Angelo and F. Palmieri, "Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial–temporal features extraction," J. Netw. Comput. Appl., vol. 173, Jan. 2021, Art. no. 102890.

[20] X. Ren, H. Gu, and W. Wei, "Tree-RNN: Tree structural recurrent neural network for network traffic classification," Expert Syst. Appl., vol. 167, Apr. 2021, Art. no. 114363..