Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 10, No. 3, 2024

Network Anomaly Traffic Analysis

Kaibin Lu[1, *]

[1]Faculty of Mathematics and Computational Sciences, School of Guilin University of Electronic Technology, Guilin, 541004, China
[*]Corresponding author's email: aaron011005@163.com

Abstract: This paper rigorously analyzes two principal methodologies in network traffic anomaly detection: feature detection and anomaly detection. Each methodology exhibits distinct strengths and confronts specific challenges. The study elucidates how the integration of deep learning with artificial immune systems could potentially transform feature detection. Moreover, it illustrates the enhancement of anomaly detection through the synthesis of machine learning techniques with traditional methods. Looking ahead, the paper delineates research trajectories that concentrate on merging deep learning, artificial intelligence, and behavioral analysis. This integration aims to augment the precision, efficiency, and adaptability of network anomaly traffic monitoring systems. Proposed future strategies include advanced methods in data preprocessing, model development, pattern recognition, and adaptive adjustments. These strategies are directed towards fortifying network defenses in response to the dynamically changing spectrum of cyber threats.

Keywords: Network traffic anomaly detection; Feature detection; Anomaly detection; Deep learning; Pattern recognition; Behavioral analysis.

1. Introduction

With the rapid development of the internet and escalating network security threats, network anomaly traffic analysis has become a crucial area in network security and management. This technique focuses on detecting and understanding abnormal traffic patterns, which may arise from various sources, such as cyberattacks, system failures, or unusual user behavior. This paper explores the importance and application contexts of network anomaly traffic analysis, which is essential in identifying and defending against cyber threats. In today's network environment, characterized by increasingly sophisticated attack methods ranging from Distributed Denial of Service (DDoS) attacks to malware proliferation and various intrusion attempts, effective traffic analysis is key to promptly detecting these threats, thereby significantly enhancing network security.

Moreover, monitoring and optimizing network performance are vital applications of network anomaly traffic analysis. Issues like traffic congestion and configuration errors can create anomalies in network traffic. Analyzing these anomalies allows network administrators to quickly identify and resolve these issues, improving network stability and efficiency. Additionally, compliance with regulatory requirements drives the evolution of network anomaly traffic analysis, with various industries and countries imposing stringent regulations and compliance mandates regarding network security. Continuous monitoring and analysis of network traffic help organizations ensure their network activities comply with these requirements.

2. Challenges

Network Anomaly Detection remains a challenging task, aiming to detect anomalous network traffic for security purposes. Typically, network traffic data are large-scale and imbalanced, often containing noisy labels. This paper addresses these challenges, utilizing the million-scale and highly imbalanced ZYELL dataset. The proposed approach involves training deep neural networks with class weight optimization to learn complex patterns from rare anomalies in the traffic data. A novel model fusion combines two deep neural networks, including a binary normal/attack classifier and a multi-attacks classifier. This solution can detect various network attacks, such as Distributed Denial of Service (DDoS), IP probing, PORT probing, and Network Mapper (NMAP) probing. Experiments conducted on a real-world ZYELL dataset demonstrate promising performance, with the proposed approach outperforming the baseline model in terms of average macro Fβ score and false alarm rate by 17% and 5.3%, respectively.

3. Detective Ways

In network traffic anomaly detection, feature detection and anomaly detection are two primary methodologies, each with distinct strengths and limitations. Feature detection is known for its high precision and rapid response in identifying known attack patterns, especially effective in dealing with predefined anomalies with a low false positive rate. However, it struggles against novel or unidentified attacks and incurs high maintenance costs due to the need for regular updates to the feature database. Conversely, anomaly detection is valued for its robust adaptability and ability to detect unknown attacks, adjusting its detection strategies in response to real-time changes in network traffic, thereby comprehensively monitoring various potential threats. Yet, this method faces challenges, including a higher false positive rate and significant computational resource requirements. Setting appropriate thresholds and parameters for anomaly detection also requires specialized knowledge. As machine learning and deep learning evolve, these approaches are constantly being redefined and refined, offering both challenges and opportunities in network security.[1-3]

3.1. Feature Detection

Chastikova and Mitugov's research introduces a novel method for network attack detection that combines deep learning with artificial immune systems, offering a new approach for efficiently and accurately detecting anomalies in

complex network environments. This method integrates deep learning, particularly convolutional neural networks (CNNs), with artificial immune systems (AIS) to create a unique neuroimmune model that enhances the adaptability and feature-learning capabilities of network security systems. The study employs an AIS based on the clonal selection algorithm, combined with an improved genetic dueling algorithm, thereby increasing overall system efficiency. Moreover, CNNs serve as a critical tool in this model for feature extraction and pattern recognition, used in training antibody-detector images based on AIS. The system's testing and training on the "Intrusion Detection Evaluation Dataset" CIC-IDS2017 demonstrated that the neuroimmune approach exhibits greater accuracy and efficiency in network attack detection compared to traditional techniques like CNNs, perceptrons, and HAIS.[1]

Chastikova and Mitugov's research not only provides fresh insights into network traffic anomalies but also contributes significantly to the advancement of future network security technologies. The potential advantages of this method in handling complex and dynamic network environments, coupled with its ability to identify attack patterns and optimize the detection process, warrant further exploration and development. Their work serves as a vital reference in comprehensive articles analyzing network traffic anomalies.

3.1.1. Challenges Faced by Feature Detection and Optimization Directions

Feature detection is crucial in network anomaly traffic monitoring, as it entails accurately extracting characteristics from extensive network data that reflect network behavior. Contemporary research leverages machine learning and deep learning techniques to automate the identification and extraction of key features. These technologies adeptly learn patterns in multidimensional data, thereby precisely identifying features correlated with anomalous behavior.

3.1.2. Optimization Directions

The key to optimizing feature detection lies in developing dynamic adaptive algorithms capable of automatically adjusting feature extraction strategies in response to changes in network behavior. Given the high dimensionality of network data, dimensionality reduction techniques such as Principal Component Analysis (PCA) or autoencoders are employed to simplify the feature space while preserving essential information. Additionally, enhancing the ability to analyze large-scale datasets in real-time represents a significant research direction in the field of feature detection.[4,5]

3.2. Anomaly Detection

In their study, Ma, Sun, and Cui introduce the SVM-C model, which synergizes the strengths of Support Vector Machines (SVM) and clustering techniques for anomaly detection in network traffic. The model converts URLs from network traffic logs into feature vectors using statistical rules and linear projection, classifying them as normal or anomalous with an SVM classifier. The researchers also developed an optimization model for training the feature extraction method and the traffic classifier parameters. In numerical tests, SVM-C consistently outperformed existing methods across tested datasets, particularly in terms of accuracy and efficiency in detecting network anomalies. Upon concluding the study, the authors suggest further optimization of the feature extraction and traffic classification parameters of the model. They also propose integrating additional machine learning techniques with SVM-C to enhance its adaptability and effectiveness in various network environments.[6,7]

Du and colleagues developed a novel method for detecting network traffic anomalies based on wavelet analysis. This method involves wavelet analysis of pcap file data to extract waveform features for classification. Specifically, the team parsed packet lengths from pcap files to create a sequence, which underwent wavelet analysis for feature extraction. A Support Vector Machine (SVM) classifier was then used to differentiate between normal and anomalous traffic. The model effectively classified traffic anomalies in pcap files, excelling in distinguishing between two types of delay injections. Both qualitative and quantitative results demonstrated the model's superior accuracy in detecting network traffic anomalies compared to traditional methods. The researchers recommend further optimizing the feature extraction process to enhance the model's ability to discern subtle differences in traffic patterns. They also suggest integrating advanced machine learning techniques with wavelet analysis to improve the model's adaptability and efficacy in various network environments. This study presents an innovative and effective solution for detecting network traffic anomalies, highlighting the potential of wavelet analysis in enhancing detection accuracy and efficiency.[8,9]

A primary challenge in anomaly detection is enhancing detection accuracy and reducing false positive rates. Traditional rule-based and signature-based methods often struggle to address novel and unknown attack patterns. Achieving efficient real-time anomaly detection is challenging in the context of large-scale network traffic. Current research focuses on employing deep learning technologies to improve detection capabilities for complex attack patterns, as well as augmenting unknown attack detection through behavior analysis techniques. Adaptive learning models, adjusting detection strategies based on real-time changes in network traffic, represent another critical research area. Additionally, integrating rule-based methods with machine learning approaches in ensemble techniques provides more comprehensive protection for network security.

4. Consider Optimization Directions

In the domain of network anomaly traffic monitoring, cutting-edge methodologies predominantly focus on the application of deep learning and machine learning technologies. These approaches enhance the accuracy and efficiency of anomaly traffic detection by automating the learning and analysis of network data. Techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have proven particularly effective in handling high-dimensional and complex data structures. Concurrently, behavior-based analysis techniques are increasingly being applied to anomaly detection, identifying irregularities through the statistical characteristics of network behavior. Moreover, ensemble learning methods, integrating various algorithms and technologies, offer comprehensive protection for network security.[7-9]

4.1. Research Directions and Methods

4.1.1. Integration of Deep Learning and Artificial Intelligence

Contemporary research is predominantly focused on integrating deep learning and artificial intelligence technologies to enhance the accuracy and efficiency of

network anomaly traffic monitoring. This integration involves utilizing autoencoders for data preprocessing and feature extraction, employing Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs) for complex pattern recognition, and applying machine learning algorithms like reinforcement learning for adaptive adjustments.[10-12]

4.1.2. Behavioral Analysis and Pattern Recognition

Behavioral analysis and pattern recognition techniques, such as clustering algorithms, enable researchers to identify anomalous patterns within vast datasets, thereby improving the detection capabilities for unknown attacks and advanced persistent threats (APTs).

4.2. Multi-Model and Ensemble Learning

Furthermore, research emphasizes the combination of diverse machine learning models, including traditional rule-based methods and contemporary deep learning models, to create a comprehensive and more robust detection system. Our future proposed research trajectory aims to implement a systematic methodology that merges the intricacies of deep learning and artificial intelligence, significantly enhancing the efficiency of anomaly detection systems in network traffic analysis.

4.2.1. Data Preprocessing and Feature Extraction

We start with comprehensive data collection, capturing a wide range of network traffic, both standard and non-standard. ur approach includes employing data cleaning techniques to resolve issues with missing values and outliers, followed by normalization to ensure data consistency. We propose the use of autoencoders for efficient data dimensionality reduction, facilitating the extraction of key features indicative of network behavior.

4.2.2. Model Building

The selection of deep learning models is based on specific data characteristics. For instance, Convolutional Neural Networks (CNNs) are chosen for processing spatial features, while Long Short-Term Memory Networks (LSTMs) are used for analyzing sequential data. These models are to be rigorously trained using a substantial corpus of historical network traffic data, enabling them to differentiate between normal and anomalous traffic patterns effectively.

4.2.3. Behavioral Analysis and Pattern Recognition

Advanced machine learning algorithms like clustering will be implemented to perform detailed behavioral analysis of network traffic, aiming to identify patterns that deviate from typical behavior.

The trained deep learning models will then be utilized to undertake complex pattern recognition tasks within the network traffic.

4.2.4. Adaptive Adjustment

We plan to integrate feedback mechanisms to continuously refine and optimize the model, based on its performance and the outcomes of detection. The incorporation of reinforcement learning algorithms will facilitate dynamic adaptation, particularly in addressing new and evolving network threats.

4.2.5. System Integration

Our methodology includes combining deep learning models with other machine learning approaches, aiming to create a robust and multi-dimensional anomaly detection system. We emphasize the establishment of real-time monitoring mechanisms to ensure prompt detection and response to network anomalies.

4.2.6. Expected Outcomes

We anticipate significant improvements in detection accuracy, leveraging the advanced feature learning capabilities of deep learning models, coupled with precise behavioral analysis. The adaptability of the system is expected to be enhanced, enabling it to efficiently address emerging and novel attack patterns, and thereby reducing the incidence of false positives and negatives. Our approach aims to achieve real-time processing and responsiveness in anomaly detection, which will substantially strengthen the network's defensive mechanisms against potential threats. This integrated approach, combining deep learning and AI, promises to bring about a revolution in network traffic anomaly analysis, paving the way for more resilient and intelligent cybersecurity infrastructures.

5. Conclusion

In conclusion, this comprehensive review underscores the critical importance of network anomaly traffic analysis in the contemporary landscape of network security. The evolving complexity of cyber threats, ranging from DDoS attacks to advanced persistent threats, underscores the need for robust and sophisticated detection methods. The exploration of both feature detection and anomaly detection methodologies reveals a dynamic field that balances the strengths and weaknesses of various approaches.

Feature detection, with its emphasis on precision and efficiency in identifying known attack patterns, remains integral in addressing predefined anomalies. However, its limitations in confronting novel or unidentified attacks highlight the necessity for continual evolution and integration of more adaptive technologies. On the other hand, anomaly detection, lauded for its versatility in identifying unknown threats, presents challenges in terms of higher false positive rates and computational demands. The future of network anomaly traffic analysis lies in the amalgamation of machine learning, deep learning, and artificial intelligence, which promises enhanced accuracy, efficiency, and adaptability in detection systems.

The proposed research directions, focusing on the integration of deep learning with AI, behavioral analysis, and multi-model ensemble learning, aim to address the existing challenges and optimize detection capabilities. Through the systematic methodology of data preprocessing, model building, behavioral analysis, and adaptive adjustments, we anticipate significant advancements in the field. The expected outcomes include improved detection accuracy, adaptability to emerging threats, and real-time responsiveness, contributing to the fortification of network defenses against a wide spectrum of cyber threats.

As we look to the future, it is evident that the field of network traffic anomaly analysis will continue to evolve, driven by the relentless pace of technological advancement and the changing face of cyber threats. The approaches and methodologies discussed in this paper lay a foundation for further research and development, with the potential to create more robust, intelligent, and resilient cybersecurity infrastructures.

References

[1] Chastikova, V.A., Mitugov, A. (2021). The method for detecting network attacks based on the neuroimmune approach. https://doi.org/10.1088/1742-6596/2094/3/032035.

[2] Du, Z., Ma, L., Li, H., et al. (2018) Network Traffic Anomaly Detection Based on Wavelet Analysis. In: 2018 IEEE 16th International Conference on Software Engineering Research, Management and Applications (SERA). Kunmming. pp. 94-101.

[3] Ma, Q., Sun, C., Cui, B. (2021). A Novel Model for Anomaly Detection in Network Traffic Based on Support Vector Machine and Clustering. Security and Communication Networks. https://doi.org/10.1155/2021/2170788.

[4] Li, M., Han, D., Yin, X., et al. (2021). Design and Implementation of an Anomaly Network Traffic Detection Model Integrating Temporal and Spatial Features. Secur.Commun Networks, 7045823:1-7045823:15.

[5] Wang, W., Sheng, Y., Wang, J., et al. (2018). HAST-IDS: Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection. IEEE Access, 6:1792-1806.

[6] Xu, W., Jang-Jaccard, J., Singh, A., et al. (2021). Improving Performance of Autoencoder-Based Network Anomaly Detection on NSL-KDD Dataset. IEEE Access, 9: 140136-140146.

[7] Su, T., Sun, H., Zhu, J., et al. (2020). BAT: Deep Learning Methods on Network Intrusion Detection Using NSL-KDD Dataset. IEEE Access, 8:29575-29585.

[8] Ji, S., Sun, T., Ye, K., et al. (2019). DAFL: Deep Adaptive Feature Learning for Network Anomaly Detection ,Network and Parallel Computing,11783: 350-354.

[9] Félix Iglesias, Zseby, T. (2014). Analysis of network traffic features for anomaly detection. Machine Learning, 101:59-84.

[10] Zhong, Y., Chen, W., Wang, Z., et al. (2020). HELAD: A novel network anomaly detection model based on heterogeneous ensemble learning. Computer Networks, 107049:169.

[11] Xiao, Q., Liu, J., Wang, Q., et al. (2020). Towards Network Anomaly Detection Using Graph Embedding. Computational Science – ICCS 2020, 12140:156 - 169.

[12] Vartouni, A., Kashi, S., Teshnehlab, M. (2018). An anomaly detection method to detect web attacks using Stacked Auto-Encoder. 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), Kerman, pp.131-134.