RADemics

Convolutional and Transformer-Based Deep Learning Architectures for Real-Time Anomaly Detection in Network Traffic

Nalini Poornima Suresh, V. Vallinayagi, S. Nanthini
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY, SRI SARADA COLLEGE FOR WOMEN, SAVEETHA SCHOOL OF ENGINEERING.

Convolutional and Transformer-Based Deep Learning Architectures for Real-Time Anomaly Detection in Network Traffic

[1]Nalini Poornima Suresh, Assistant Professor, CSE, SRM Institute of Science and Technology, Ramapuram, nalinipoornima636@gmail.com

[2]V. Vallinayagi, Associate professor and head, Computer science, Sri Sarada college for women, Ariyakullam, Tirunelveli 627011.Vallinayagimahesh1@gmail.com

[3]S. Nanthini, Professor, Saveetha School of Engineering, SIMATS, Chennai, nanthini27j88@gmail.com

Abstract

The rapid expansion of digital infrastructures has led to an unprecedented increase in cyber threats, necessitating advanced techniques for real-time anomaly detection in network traffic. Traditional rule-based and statistical methods often fail to detect sophisticated attacks due to their reliance on predefined signatures and limited adaptability to evolving threats. Deep learning has emerged as a promising alternative, leveraging data-driven approaches to enhance detection accuracy. Convolutional Neural Networks (CNNs) have demonstrated efficiency in extracting spatial-temporal patterns from network traffic, while Transformer-based architectures excel in capturing long-range dependencies and sequential anomalies. However, existing solutions face challenges related to scalability, computational overhead, imbalanced datasets, and adversarial robustness. This chapter provides a comprehensive analysis of CNN and Transformer-based deep learning architectures for real-time anomaly detection, highlighting their strengths, limitations, and practical deployment challenges. A hybrid approach integrating CNNs and Transformers is explored to enhance detection performance by combining local feature extraction with global sequence modeling. the role of synthetic data augmentation, adaptive learning techniques, and adversarial defense mechanisms in improving model generalization and resilience is examined. Future research directions focus on explainable AI, lightweight models for real-time applications, and self-supervised learning for mitigating data scarcity. The insights presented in this chapter contribute to the advancement of AI-driven cybersecurity solutions, enabling proactive threat detection and risk mitigation in dynamic network environments.

Keywords: Deep Learning, Anomaly Detection, Network Security, Convolutional Neural Networks, Transformers, Cyber Threats

Introduction

The exponential growth of networked systems and digital infrastructures has significantly increased the attack surface for cyber threats, making real-time anomaly detection a crucial aspect of modern cybersecurity. Traditional intrusion detection systems (IDS) rely on rule-based or statistical anomaly detection methods, which often struggle to detect novel and sophisticated

attacks. These conventional approaches are inherently limited by their dependence on predefined signatures and manually crafted feature sets, making them ineffective against zero-day attacks, polymorphic malware, and adversarial exploits. As cyber threats continue to evolve in complexity, security frameworks must transition toward data-driven methodologies that can autonomously learn patterns from vast volumes of network traffic. Deep learning, with its ability to extract meaningful representations from high-dimensional data, has emerged as a promising solution to address these limitations in network anomaly detection.

Convolutional Neural Networks (CNNs) have demonstrated strong capabilities in analyzing structured network traffic by capturing spatial and temporal correlations within network flow data. By leveraging hierarchical feature extraction, CNNs can detect localized patterns associated with malicious behaviors, making them highly effective for anomaly detection in network security. However, CNNs have inherent limitations when dealing with long-range dependencies, as they primarily focus on local feature extraction. This drawback restricts their ability to detect complex attack patterns that unfold over extended time sequences, such as multi-stage intrusions or advanced persistent threats (APTs). To overcome this challenge, researchers have explored alternative deep learning architectures that can capture both short-term and long-term dependencies within network traffic.

Transformer-based architectures, originally developed for natural language processing (NLP), have recently gained prominence in cybersecurity due to their self-attention mechanisms, which enable them to model long-range dependencies more effectively than traditional recurrent or convolutional models. Unlike CNNs, which process data in a fixed receptive field, Transformers dynamically weigh the importance of different network traffic segments, allowing them to learn intricate relationships between different network flows. This ability to capture complex dependencies makes Transformers particularly suitable for identifying sophisticated cyber threats, including low-and-slow attacks and coordinated distributed denial-of-service (DDoS) campaigns. Moreover, Transformer-based models have shown superior performance in handling sequential data, making them a viable alternative to traditional deep learning techniques for real-time anomaly detection.