Automated Threat Detection Using Machine Learning in Python

ABSTRACT

The rapid evolution of cyberattacks such as ransomware, Distributed Denial-of-Service (DDoS), and advanced persistent threats highlights the limitations of traditional rule-based and signature-based intrusion detection systems, which struggle to detect real-time threats in the face of increasing network complexity. In order to provide a reliable method for detecting both known and unknown cyberthreats, this study presents an automated threat detection system that uses machine learning to examine network packet capture (PCAP) data. To achieve high accuracy and adaptability, the suggested framework blends the supervised learning algorithm Random Forest with the unsupervised anomaly detection technique Isolation Forest. While Isolation Forest minimizes false positives in dynamic contexts by isolating unique abnormalities, Random Forest is excellent at categorizing existing assault patterns.The system trains the hybrid model for real-time deployment, guaranteeing scalability across cloud infrastructures and enterprise networks, after preprocessing raw PCAP logs to remove noise and extract parameters like packet size and flow duration. Tested on the CICIDS2017 dataset, which contains a variety of attack scenarios like port scanning and DDoS, the system achieves remarkable precision (0.99) and recall (0.98), as well as 98% accuracy through Random Forest and a 20% false positive rate. Isolation Forest improves detection even more by identifying zero-day anomalies that signature-based techniques miss.Using fresh threat data, a dynamic feedback loop continuously improves the model's resilience against changing attack methods. It is perfect for Security Operations Centers (SOCs), cloud platforms, and real-time monitoring systems due to its excellent accuracy, scalability, and adaptability, even with a moderate 20% false positive rate. By bridging the gap between state-of-the-art machine learning and real-world cybersecurity requirements, this research provides actionable intelligence for proactive threat mitigation and lays the groundwork for future advancements in automated detection systems.

Keywords: Machine Learning, Cybersecurity, Network Traffic Analysis, Anomaly Detection, Real-Time Threat Detection.