RADemics

Benchmarking Techniques and Evaluation Metrics for Machine Learning-Driven Network Intrusion Detection Systems

Ravi Bukya, Sarangam Kodati
NALLA MALLA REDDY ENGINEERING COLLEGE (A),
CVR COLLEGE OF ENGINEERING

Benchmarking Techniques and Evaluation Metrics for Machine Learning-Driven Network Intrusion Detection Systems

[1]Ravi Bukya, Associate Professor, Department of Computer Science Engineering (EEE), Nalla Malla Reddy Engineering College (A), Gatkeshar, Medchal-Malkajgiri, Telangana - 500088, India. ravibhukya.cse@nmrec.edu.in

[2] Sarangam Kodati, Associate Professor, Department of Information Technology, CVR College of Engineering, Hyderabad, Telangana, India.k.sarangam@gmail.com

Abstract

This book chapter provides an in-depth exploration of the benchmarking techniques and evaluation metrics critical for assessing machine learning-driven Network Intrusion Detection Systems (NIDS). With the increasing complexity and sophistication of cyber-attacks, traditional NIDS are facing challenges in terms of efficiency, accuracy, and scalability. The chapter highlights key aspects such as dataset diversity, attack representation, and data augmentation, which are essential for improving NIDS performance. A comprehensive analysis of popular benchmarking datasets, including their preprocessing techniques and ethical concerns, was presented. Emerging trends such as blockchain integration for secure dataset sharing and the application of NIDS in domain-specific areas like smart grids and energy systems are also discussed. This chapter aims to bridge the gap between theoretical frameworks and practical implementation, offering insights into future directions for enhancing NIDS through innovative evaluation approaches. The integration of these techniques will ultimately enhance cybersecurity resilience across various sectors.

Keywords:

Network Intrusion Detection Systems, Machine Learning, Benchmarking, Datasets, Data Augmentation, Smart Grids.

Introduction

The increasing sophistication of cyber threats and attacks has highlighted the need for advanced methods to protect network infrastructures [1]. Network Intrusion Detection Systems (NIDS) play a pivotal role in safeguarding networks by identifying and responding to malicious activities [2]. However, with the growing complexity of attacks, traditional NIDS have proven insufficient in maintaining high detection accuracy and low false positive rates [3]. To address this gap, machine learning-driven NIDS are emerging as powerful tools that can adapt to dynamic and evolving attack strategies [4,5]. These systems offer a promising solution by leveraging data-driven approaches to detect anomalies and intrusions [6]. Nonetheless, evaluating the performance of these machine learning-based models remains a challenging task, requiring comprehensive benchmarking techniques and robust evaluation metrics [7-9].

Effective evaluation of machine learning-driven NIDS was critical for understanding their strengths, limitations, and real-world applicability [10]. Benchmarking provides a standardized method for comparing different NIDS models across diverse datasets, attack types, and operating conditions [11]. This ensures that the chosen models are not only effective in controlled environments but also resilient in unpredictable real-world scenarios [12]. A key aspect of benchmarking was the use of high-quality datasets that represent various network traffic patterns, attack vectors, and operational conditions [14]. The accuracy and reliability of these datasets significantly impact the performance of NIDS and influence the results of their evaluation [14]. Hence, a well-curated set of benchmarking datasets was essential for effective model testing [15].

The challenge of dataset diversity was one of the most significant obstacles in NIDS evaluation [16]. Many existing datasets suffer from limited attack scenarios or lack diversity in benign network traffic patterns [17]. This creates a skewed representation thatnot reflect the full spectrum of cyber threats that NIDS models need to identify [18]. To address this, datasets must be diverse, encompassing a wide range of attack techniques, including both known and novel threats [19]. These datasets should incorporate traffic from various network environments, such as enterprise networks, IoT ecosystems, and cloud infrastructures, to ensure that NIDS can generalize effectively [20-22]. The inclusion of diverse data sources not only improves the robustness of the models but also helps in reducing overfitting, ensuring that models do not perform well only on specific datasets but can generalize to new, unseen data [23,24].

In addition to dataset diversity, data preprocessing plays a crucial role in the effectiveness of machine learning models [25]. Raw network data often contains noise, inconsistencies, and irrelevant features that can reduce the accuracy of NIDS. Effective data preprocessing techniques, including feature selection, normalization, and filtering, are essential to enhance the quality of the data fed into the model.