Modular ML-Based IDS: Layered Attack Detection via Header & Payload Analysis

Mohammad Saghali
Politecnico di Torino, Italy
mohammad.saghali@studenti.polito.it
Nokia Bell Labs, France
mohammad.saghali@nokia.com

Mahdi Naderibeni
Politecnico di Torino, Italy
mahdi.naderibeni@studenti.polito.it
Nokia Bell Labs, France
mahdi.naderibeni@nokia.com

Vikramajeet Khatri
Nokia Bell Labs, Finland
vikramajeet.khatri@nokia-bell-labs.com

Siwar Kriaa
Nokia Bell Labs, France
siwar.kriaa@nokia-bell-labs.com

Serge Papillon
Nokia Bell Labs, France
serge.papillon@nokia-bell-labs.com

Mehrnoosh Monshizadeh
Nokia Bell Labs, France
mehrnoosh.monshizadeh@nokia-bell-labs.com
Department of Information and Communication Engineering, Aalto University, Espoo, Finland
mehrnoosh.monshizadeh@aalto.fi

***Abstract*—We present a demonstration of Hybrid Anomaly Detection Model (HADM), a machine learning based architecture that effectively identifies and filters malicious network activities. HADM comprises of a protocol analyzer as well as several classification and clustering algorithms. Through interactive visualizations and real-time analysis, we demonstrate the platform's effectiveness using various classification and clustering metrics, including precision, recall, and silhouette score. The demonstration highlights HADM's robust scalability in handling datasets of varying sizes and its adaptability to diverse attack patterns, validating it as a comprehensive solution for modern network security threats.**

Keywords— Intrusion Detection System, Machine Learning, Classification, Clustering, Header Analysis, Payload Analysis.

I. Introduction

Intrusion detection systems are considered well-known tools for monitoring and detecting malicious traffic in communication networks. However, traditional intrusion detection systems rely on known signatures [1] and lack the ability to detect novel attacks. Therefore, machine learning techniques are introduced to complement intrusion detection and to dynamically identify the relevant data of interest and intelligently find out the security threats.

On the other hand, the 5G+ and 6G networks are expected to deliver massive connectivity to numerous IoT/IoE devices [2], where a huge amount of data needs to be analyzed by artificial intelligence enabled mechanisms. Consequently, a mature and scalable architecture must be considered as a mandatory objective in machine learning based intrusion detection systems.

This paper presents a machine-learning based architecture that solves the mentioned issues in the cyber-security domain. The paper proposes an intelligent, modular, robust, and scalable security solution to dynamically detect known and unknown cyber-attacks. The proposed architecture enhances the intrusion detection with a hybrid machine learning based mechanism named Hybrid Anomaly Detection Model (HADM) that consists of a protocol analyzer and several supervised and unsupervised techniques to filter network traffic and identify malicious activity in network traffic [3] [4] [5].

II. System Architecture

As shown in Figure 1, the HADM compromises two main parts where each one independently increases the efficiency of attack detection based on the metrics such as computation time, precision, recall and so on. Overall, the proposed model utilizes the protocol analyzer and a combination of classification and clustering algorithms with supervised and unsupervised learning processes for network traffic filtering. In part 1, the protocol analyzer classifies and filters vulnerable protocols to avoid unnecessary computation load. The classifiers detect known cyber-attacks, while clustering algorithms use these attributes and features to cluster unknown traffic.

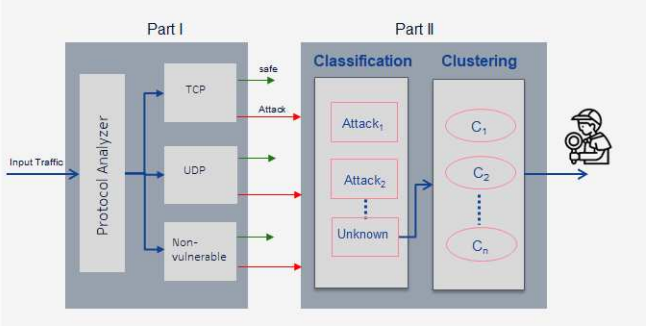

Fig. 1. Hybrid Anaomly Detection Model

After passing through protocol analyzer in Part I, the traffic carried on the Transmission Control Protocol (TCP), User Datagram Protocol (UDP) and other protocols will be separately analyzed by respective classifier algorithms (TCP, UDP and Non-vulnerable respectively). These algorithms are binary classifiers and predict whether the traffic is attack or benign. This structure not only helps to reduce the load of the model, but also helps the analysis process with having specific algorithms according to the nature of the protocols and traffic carried over it.

The traffic classified as attack in Part I will be forwarded to Part II. This part consists of classification and clustering algorithms. The classification algorithm classifies attack into multiple classes, and if the detection probability is low or uncertain, it classifies the attack as unknown. The unknown class is forwarded to the clustering algorithm, which creates clusters to categorize unknown traffic. Such clustering helps

ISBN 978-3-903176-72-0 © 2025 IFIP

security investigator in analyzing and defining new attacks based on the cluster composition [6] [7].

In cyber-security domain, IDS can be applied on network traffic packet or flow. In this paper, packet-based analysis is applied since this approach allows to have full information about network activities and makes real-time analysis possible. Furthermore, the network packet consists of the header and the payload. The header is a structured data with distinct information (features) about packets. On the contrary, the payload is the actual intended data in unstructured format and with unspecific features.

The datasets CIC-IDS-2017 [8] and UNSW-15 [9] are used for header and payload analysis in this paper. As it is shown in Table I, the current experiment applies a combination of mentioned datasets, which is a common approach in the field of machine learning [6] [10] [11].

TABLE I. DATASETS

Type of Traffic	Dataset	Class	Number of Packets
Header	CIC-IDS-2017	Normal	9074
		Unknown	10000
		FTP/SSH-Patator	2328
		DoS	32749
		Heartbleed	439
		Web Attack	339
		File Download	2143
		Botnet	85
		Port Scan	2843
	UNSW-15	Normal	58014
		Unknown	20000
		DoS	1423
		Exploits	2876
		Fuzzers	1387
		Reconnaissance	1094
		Generic	1185
		Shellcode	1013
		Backdoor	1002
		Worms	1004
		Analysis	1002
Payload	CIC-IDS-2017	Normal	105373
		Unknown	10000
		Dos	32267
		Brute Force	7264
		Exploit	33056
		Botnet	31785

Since CIC-IDS-2017 dataset does not contain Unknown class, some of the packets from this dataset are separated and considered as Unknown. On the other hand, while the labels of these packets are known to us, algorithms are not trained with them, so the packets are labelled as unknown. This approach has been applied for header and payload analysis of this dataset.

III. DEMO DESCRIPTION

The HADM dashboard leverages pre-trained machine learning models for binary classification and multiclass classification, as detailed in Part I and Part II respectively as shown in Fig 2. Additionally, it incorporates a novel density based clustering algorithm to categorize unknown packets in distinct clusters. This demonstration showcases HADM's capabilities using various types of network traffic, including both header and payload traffic.

A. Settings Tab

In order to demonstrate, the robustness of HADM different datasets are presented here. Furthermore, in this tab header and payload analysis are considered.

B. Overall Architecture Tab

This tab illustrates the operational workflow of HADM with the selected dataset. Data is streamed to HADM in a pipeline manner, where packets are processed at regular interval.

The proposed pipeline consists of a binary classifier followed by a multi classifier in order to reduce the overall false positive rate. The architecture recovers some of false positives that were generated in the first layer of detection (binary classifier). The last module of this pipeline is a clustering algorithm used to categorize the Unknown class.

At each interval, updated plots are displayed, showing the classification results for binary and multi classification, as well as clustering outcomes. The number of classified attacks is updated in real-time on the plots. Pre-trained models are used for classification and clustering.

C. Performance Tab

This tab is divided into several sections: overall statistics, performance progress, clusters, and cluster details.

- **Overall Statistics**: Displays precision and recall metrics for each type of attack classified by the multi classifier. The plots are updated in real-time as new packets are received. Users can toggle between false positive/false negative rates and precision-recall metrics using a radio button, which displays the corresponding plot.
- **Performance Progress**: Allows users to select between precision and recall metrics. The selected metric's progress over time is visualized, showing how these metrics evolve as new packets are processed.
- **Clusters**: Demonstrates the clustering of unknown packets. Dimension reduction is applied to these packets before clustering, and a 3D visualization of the clusters is provided.
- **Cluster Details**: Contains three plots to evaluate clustering performance. The first plot shows the distribution of attack types across clusters, updated in real-time. The second plot presents the silhouette analysis, which quantifies the degree of similarity between each data point and its assigned cluster in comparison to other neighbouring clusters. The third plot visualizes the distribution of attacks within clusters, providing insights into how different attack types are spread across clusters.

Fig. 2. The overall architecture of the HADM dashboard

D. Alert Tab

This tab lists all packets classified as malicious. Users can select individual packets to view detailed information about the identified attacks e.g., source, destination and nature of malicious packet, which can be used to enhance network security.

IV. Conclusion

In this paper, we introduced a comprehensive ML-based intrusion detection platform, HADM, which integrates multiple components including a protocol analyzer, classification and clustering algorithms. Each component serves a distinct purpose within the security pipeline: the protocol analyzer efficiently segregates network traffic based on protocols, classification algorithms identify known attack patterns, and clustering algorithms detect novel, unknown threats. The demonstrator's capabilities are showcased through extensive performance evaluations, utilizing respected metrics for each algorithm. A key highlight of our demonstration is the real-time 3D visualization of streaming unknown network traffic, specifically designed for zero-day attack detection. This dynamic visualization effectively illustrates evolving clusters in multi-dimensional space, providing immediate visual identification of unidentified attack traffic. Furthermore, the demonstrator incorporates an alert system that provides immediate notifications for detected malicious packets, ensuring prompt response to potential threats. This comprehensive approach validates HADM's effectiveness as a robust, scalable solution for modern network security threats.

References

[1] C. Iyanu-Oluwa Onietan, I. Martins, T. Owoseni, E. C. Omonedo and C. P. Eze, “A Preliminary Study on the Application of Hybrid Machine Learning Techniques in Network Intrusion Detection Systems,” in 2023 International Conference on Science, Engineering and Business for Sustainable Development Goals (SEB-SDG), 2023.

[2] M. Banafaa, I. Shayea, J. Din, M. Hadri Azmi, A. Alashbi, Y. Ibrahim Daradkeh and A. Alhammadi, “6G Mobile Communication Technology: Requirements, Targets, Applications, Challenges, Advantages, and Opportunities,” Alexandria Engineering Journal, vol. 64, pp. 245-274, 2023.

[3] M. Monshizadeh, V. Khatri, B. Atli and R. Kantola, “An Intelligent Defense and Filtration Platform for Network Traffic,” in Wired/Wireless Internet Communications, 2018.

[4] M. Monshizadeh, V. Khatri, A. Buse Gul, R. Kantola and Z. Yan, “Performance Evaluation of a Combined Anomaly Detection Platform,” IEEE Access, vol. 7, pp. 100964-100978, 2019.

[5] M. Monshizadeh, “Machine Learning Techniques to Detect Known and Novel Cyber-attacks,” Aalto University, 2023.

[6] M. Monshizadeh, V. Khatri, M. Gamdou, R. Kantola and Z. Yan, “Improving Data Generalization With Variational Autoencoders for Network Traffic Anomaly Detection,” IEEE Access, vol. 9, pp. 56893-56907, 2021.

[7] M. Monshizadeh, V. Khatri, R. Kantola and Z. Yan, “A deep density based and self-determining clustering approach to label unknown traffic,” Journal of Network and Computer Applications, vol. 207, p. 103513, 2022.

[8] I. Sharafaldin, A. H. Lashkari and A. A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization,” in 4th International Conference on Information Systems Security and Privacy (ICISSP), 2018.

[9] “UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” in 2015 Military Communications and Information Systems Conference (MilCIS), 2015.

[10] M. Essid and F. Jemili, “Combining intrusion detection datasets using MapReduce,” in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.

[11] M. Elayni and F. Jemili, “Using MongoDB Databases for Training and Combining Intrusion Detection Datasets,” in Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Springer, 2018, pp. 17-