A Lean and Modular Two-Stage Network Intrusion Detection System for IoT Traffic

Bruno Brandão Inácio*, Juliana C. Carvalho de Araújo*,
Aldri L. dos Santos†, Michele Nogueira†, R. Hirata Jr.*, Daniel M. Batista*
*Department of Computer Science, University of São Paulo (USP), Brazil
{bruno.inacio,julianaraujo,hirata,batista}@ime.usp.br
‡Department of Computer Science, Federal University of Minas Gerais (UFMG), Brazil
{aldri,michele}@dcc.ufmg.br

2024 IEEE Latin-American Conference on Communications (LATINCOM) | 979-8-3315-2111-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/LATINCOM62985.2024.10770670

***Abstract*—The popularization of the Internet of Things (IoT) has led to cyberattacks targeting interconnected applications. Traditional intrusion detection systems (IDS) struggle to cope with the increasing volume and complexity of IoT data, hindering their ability to identify all threats. In order to address this issue, we propose a modular two-stage Network IDS for IoT, with each stage specialized in a specific set of attacks. This approach allows for independent training and optimization of each stage, improving processing time and classification metrics when compared to single-stage systems. The effectiveness of this design is demonstrated using the CICIoT2023 dataset. Compared to a single model, our proposal obtains better inference time (13 seconds vs. 94 seconds) and an overall enhanced detection rate (0.9055 vs. 0.8370 in terms of F1-Score). An additional contribution of our work is the sharing of all developed code as open-source software, facilitating the reproduction and extension of our proposal.**

***Index Terms*—Intrusion Detection, IoT, Security.**

I. Introduction

The widespread adoption of Internet of Things (IoT) has profoundly reshaped modern life. The omnipresent smart devices, spanning from household appliances to industrial sensors and wearable technology, offer substantial advantages in automation, efficiency, and real-time data acquisition. However, this interconnected ecosystem presents significant challenges to cybersecurity, mainly due to the low processing capacity of the devices and the absence of a well-established strategy for firmware/software updates, making the IoT environment susceptible to DDoS attacks, data theft, and other security threats [10]. For instance, in a domestic setting, smart home devices can provide convenience, accessibility, and more comfort, but at the same time, they can be the entry point for attacks related to privacy invasion [14].

As IoT devices become increasingly integrated into personal environments, industrial control systems, and even critical infrastructure, they present a tempting target for malicious actors seeking to exploit vulnerabilities [7], [11]. The very nature of IoT ecosystems, characterized by a diverse array of devices communicating across various networks, creates a complex security landscape riddled with potential threats and attack vectors. Hence, ensuring the security of IoT has become a concern for governments, businesses, and individuals.

However, designing effective intrusion detection mechanisms for IoT environments, such as NIDS (Network Intrusion Detection Systems), presents its own set of challenges [5], [15]. Traditional signature-based detection methods struggle to keep pace with the ever-evolving arsenal of novel attack vectors and sophisticated malware targeting IoT devices. Furthermore, resource constraints inherent in many IoT devices require the development of lightweight and energy-efficient detection for real-time operation without hindering device performance [2], [6], [7].

The deficiency of traditional NIDS to protect modern environments such as the IoT has motivated employing machine learning techniques to empower anomaly-based NIDS [1]. However, to apply this approach effectively, one must be prepared to deal with unbalanced traffic to avoid high accuracy when detecting most present attacks but low accuracy when detecting less present attacks.

In light of this scenario, this paper presents a modular two-stage NIDS. The main contributions are as follows:

(1) the proposal of a two-stage network intrusion detection model based on decision tree models for enhancing detection accuracy and efficiency while fully classifying large-scale unbalanced data;
(2) the usage of a decision tree multi-class classifier to detect DoS, DDoS, and Mirai attacks in the first stage;
(3) the usage of a random forest classifier in the second stage to handle the remaining attack types;
(4) the sharing of all the developed code as open-source software.

Experimental results with the CICIoT2023 dataset [13], an extensive and brand new network traffic dataset, show that this two-stage intrusion detection model can effectively adapt to large-scale IoT network flow data, requiring fewer resources than a similar single-stage model. For instance, 0.9055 vs. 0.8370 in terms of F1-Score, 13 seconds vs. 94 seconds in terms of prediction time, and 5.07GB vs. 14.07GB in terms of memory usage.

This paper proceeds as as follows. Section II summarizes relevant research on NIDS for IoT. Section III describes our two-stage intrusion detection model. The dataset used in the experiments is detailed in Section IV. Experimental results and discussion analysis are given in Section V. Section VI concludes the paper and presents future works.

II. Related Works

The evolving landscape of IoT security has spurred a surge in research on Intrusion Detection Systems (IDS). Recent studies, many leveraging the comprehensive CICIoT2023 dataset [13], have explored diverse methodologies to enhance identifying and mitigating IoT-specific threats. As demonstrated by Jony et al. [8], deep learning techniques have shown promising results. Their LSTM-based model achieved exceptional accuracy in discerning subtle attack patterns within CICIoT2023, underscoring the potential of sequential data modeling in bolstering IoT security.

Parallel research has investigated the integration of federated learning with established methods like K-Means clustering (Hajj et al. [5]). This innovative approach enhanced detection accuracy and addressed the critical issue of computational efficiency, a key consideration for resource-constrained IoT devices. The rise of hybrid models, exemplified by Yaras and Dener's [17] fusion of CNN and LSTM networks, has further expanded the toolkit for IoT intrusion detection. Their model's remarkable accuracy across multiple datasets showcases the efficacy of combining complementary architectures to capture diverse attack characteristics. Beyond model architectures, the research of Narayan et al. [12] underscores the importance of feature engineering. By strategically selecting informative features and balancing class representation, they achieved significant improvements in detection performance, particularly for previously under-performing attack categories. Furthermore, Khan and Alkhathami [9] demonstrated the applicability of anomaly detection techniques to critical domains like healthcare IoT. Their emphasis on real-time responsiveness and computational efficiency underscores the practical relevance of their work for real-world deployments. Considering real-world deployments, designing fast-protecting systems, such as the one proposed in [3] based on stream learning is essential. The idea of these systems is to evaluate the traffic as it arrives, reducing the reaction time when an attack begins.

Similar to some existing works, our NIDS is oriented to IoT traffic, aims to reduce the prediction time, and is evaluated with the CICIoT2023 dataset. The advancement in the state of the art comes from our decision to separate the classification based on the most and least present attacks in network traffic.

III. Two-Stage Model Approach

We have developed a two-stage model [1] tailored for the analysis and classification of IoT traffic. Specifically, we considered the network traffic within the CICIoT2023 dataset. Fig. 1 presents how the model operates sequentially, with the first stage focused on identifying prevalent attack types related to Denial of Service, while the second stage specializes in detecting more intricate and less frequent threats. Traffic that does not align with any malicious patterns is classified as benign. In case other datasets are used, the specialization of each stage must be reconfigured, considering the most and least prevalent attacks.

[1]Repository:github.com/inaciobr/Multi-Stage-IoT-Intrusion-Detection

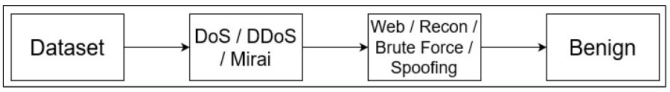

Fig. 1. Two-Stage model approach.

The initial stage of our model is designed aiming to improve efficiency through a simpler model, while keeping a very high accuracy. By concentrating on the detection of common and high-frequency attacks, supported by the already very good results seen in other publications and by the high volume of data available for these attacks, we can significantly reduce the computational burden on the subsequent stage. This strategic filtering process allows for a more focused analysis of the remaining traffic, potentially enhancing the detection of subtle and sophisticated attack vectors that may have been obscured in the initial data.

After the first stage is done acting as a filter for the Denial of Service attacks, the second stage aims to improve the accuracy in identifying the other attacks, which can be done with a more complex model since now we are dealing with a smaller subset of the network flows. We have developed a classifier designed to leverage the power of multiple models in each stage of our two-stage intrusion detection system. The modular nature of the classifier facilitates the substitution of a specific stage by improved models as they become available and even the addition of new modules able to detect specific new attacks.

Each stage of our system is supported by a dedicated model that has been trained on a specific set of labels and is designed to detect only that set of attacks. This stage-specific approach ensures that each model is optimized for the particular task at hand, leading to enhanced detection accuracy. Moreover, the classifier is designed to accommodate diverse feature sets, enabling us to readily swap out models based on performance improvements or changes in the threat landscape. In summary, the modular two-stage design of our NIDS offers distinct advantages in the realm of IoT security:

- **Detection Accuracy**: The layered approach allows each stage to specialize in specific attack vectors, leading to a more granular and accurate identification of threats.
- **Scalability and Adaptability**: Each stage can be independently optimized, updated, or replaced without requiring a complete system overhaul, enabling the NIDS to adapt to new threats and network environments seamlessly.
- **Resource Utilization**: By filtering out common attacks in the first stage, the system reduces computational overhead and ensures optimal utilization of resources for the in-depth analysis in the second stage.

A. Decision trees

Decision trees have emerged as a valuable tool in intrusion detection for IoT environments. Their ability to easily handle both categorical and numerical data, coupled with their interpretable nature, and very lightweight execution, makes

them well-suited for analyzing the diverse and complex data generated by IoT devices.

In the context of IoT intrusion detection, decision trees can be trained on labeled network traffic data to learn patterns and distinguish between normal and malicious behavior. The tree's internal nodes represent decision points based on specific features of the network traffic, such as protocol types, inter arrival time between packets, etc... . Each branch emanating from a node represents a possible outcome of the decision, and the leaf nodes represent the final classification of the traffic as either benign or an specific attack.

The interpretability of decision trees is a significant advantage in intrusion detection. By visualizing the tree structure, security analysts can gain insights into the decision-making process and understand the specific features that contribute to the classification of an attack. This transparency can be invaluable in identifying the root causes of security incidents and developing effective mitigation strategies. Moreover, decision trees can handle the high dimensionality of IoT data.

The first stage of our NIDS is based on decision trees.

B. Random Forest Classifier

The Random Forest Classifier offers several advantages for detecting IoT attacks, making it a valuable tool in cybersecurity. One significant benefit is its ability to handle large-scale datasets with high dimensionality. IoT datasets like CICIoT2023 contain vast information about network traffic, device behavior, and communication patterns. The Random Forest Classifier is particularly well-suited for these large-scale, high-dimensional datasets, effectively leveraging extensive information to detect attacks accurately.

Another advantage of the Random Forest Classifier is its robustness to noise and outliers, which are common in real-world IoT data. IoT environments frequently exhibit network conditions, device configurations, and user behavior variations, leading to data inconsistencies. Despite these challenges, the Random Forest Classifier maintains accurate predictions due to its inherent ability to handle noisy data. This robustness ensures reliable performance even when faced with the irregularities typical of IoT datasets.

The ensemble nature of the Random Forest Classifier, which integrates multiple decision trees, enables it to capture a wide range of attack patterns. The classifier's resilience against overfitting and ability to generalize to new, unseen attacks is particularly valuable in the dynamic scenario of IoT threats, where attackers continuously develop new techniques and strategies. This effective detection of diverse attack types makes the Random Forest Classifier a versatile and powerful tool in cybersecurity.

The second stage of our NIDS is based on random forest.

C. Feature Selection

We have identified several features exhibiting minimal variance or high correlation during an exploratory data analysis (EDA) of the CICIoT2023 dataset. Therefore, we implemented a feature selection process incorporating variance thresholding, Spearman's rank correlation coefficient analysis, and additional checks for redundancy to enhance the efficiency, predictive capability, and the explainability of our intrusion detection model.

Features with variance below a predetermined threshold were removed, as these contribute minimally to model discrimination. Spearman's correlation coefficient was used to identify and eliminate one feature from pairs exceeding a defined correlation threshold, mitigating multicollinearity issues, and ensuring a refined dataset for subsequent analysis. This feature selection process resulted in the elimination of 13 features from the initial set of 46, yielding a final dataset comprising 33 highly informative features [2].

IV. Dataset

The CICIoT2023 dataset is a resource for security analytics in the Internet of Things (IoT) domain, providing a benchmark for evaluating the impact of large-scale attacks on IoT systems. It includes various attack scenarios such as DDoS, DoS, reconnaissance, web-based, brute force, spoofing, and Mirai botnet attacks, simulated within a network of 105 devices [13]. The attacks were carried out on the IoT devices using other IoT devices.

Data is available in pcap files for detailed analysis and feature engineering and CSV files with pre-extracted features for machine learning applications. The dataset contains 46,686,579 samples from the IoT devices, categorized into 34 classes with 46 features per sample. However, it exhibits a significant class imbalance, with the DDoS-ICMP Flood attack class having the most samples and the Uploading Attack class the fewest, creating a 5751:1 ratio. This imbalance can hinder the performance of machine learning models, particularly in detecting minority class instances [4], [16] (Fig. 2 illustrates the class-wise data distribution of the CICIoT2023 dataset).

Our proposal, with each stage specialized in a set of attacks, must cope with this imbalance.

V. Results and Discussion Analysis

The increasing complexity of network environments and the proliferation of IoT devices necessitate robust NIDS. Our approach aims to enhance detection accuracy through careful feature engineering and a staged classification process.

To validate our model, we have used an 80/20 split: 80% of the data was used for model selection, hyperparameter tuning, and training, while the remaining 20% of the data was used solely for performance evaluation. The results were then compared against a random forest baseline derived from the best results reported by the dataset authors [13], which have its results summarized in the Table I (The syntax "$n+1$ attacks" means n attack classes + the benign class).

The performance metrics of the intrusion detection model are related to various categories of network traffic, including benign activities and different types of cyberattacks. Those

[2]The complete list of features resulting from this step can be found at Multi-Stage-IoT-Intrusion-Detection/notebooks/2 - Feature Engineering.ipynb

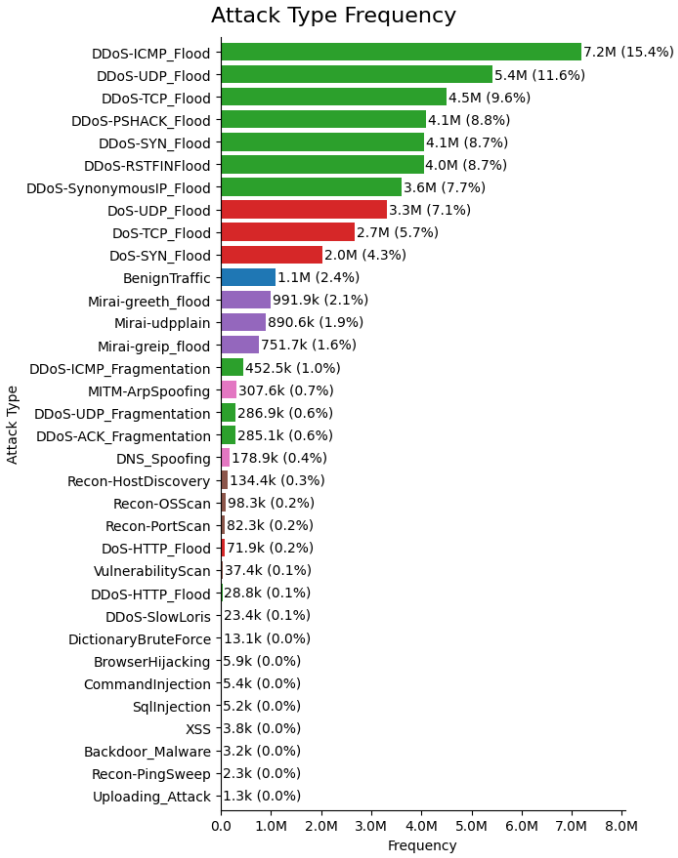

Fig. 2. Class distribution of CICIoT2023 dataset.

metrics include precision, recall, and F1-score for each category. Further, we provide the overall accuracy, macro average, and weighted average of the model's performance.

TABLE I
BASELINE [13] PERFORMANCE METRICS FOR 7+1 ATTACKS

Category	Precision	Recall	F1-Score
Benign	0.8155	0.9462	0.8760
BruteForce	0.9932	0.1468	0.2558
DDoS	0.9992	0.9997	0.9994
DoS	0.9994	0.9972	0.9983
Mirai	0.9989	0.9993	0.9991
Recon	0.7048	0.7078	0.7063
Spoofing	0.8834	0.6239	0.7313
Web	0.9718	0.0381	0.0733
		Accuracy	0.9911
		Macro Avg F1-Score	0.7050
		Weighted Avg F1-Score	0.9911

The dataset author's model effectively identifies DDoS, DoS, and Mirai attacks, with nearly perfect precision, recall, and F1-scores in these categories. However, its performance varies significantly for other attack types. For both Brute Force and Web attacks, the model achieves high precision but very low recall, indicating that it correctly identifies Brute Force attacks when detected, but misses many instances, resulting in many false negatives.

The model's overall accuracy is 0.9911, reflecting its ability to classify most instances correctly. The macro average, which considers each category equally, indicates a balanced but lower performance due to the influence of categories like Brute Force and Web attacks. The weighted average, which accounts for the number of instances in each category, remains high due to the strong performance in high-volume attack categories. This analysis highlights the model's strengths in detecting prevalent and severe attacks while identifying areas for improvement in detecting less frequent or more subtle attacks.

In the initial stage of our IoT NIDS, we opted for a simple Decision Tree model to focus on the identification of Denial of Service attacks. The model choice was driven by its impressive performance in detecting DoS attacks within the CICIoT2023 dataset, as well as its inherent simplicity, efficiency, and straightforward implementation. Here, our primary objective is to effectively filter out DoS attacks, leaving the remaining traffic for more in-depth analysis in the subsequent stage. Therefore, all traffic must be analyzed, but other types of traffic, including benign activities and other attack vectors, are classified as "benign" at this stage. This strategic filtering process reduces the computational burden on the subsequent stage, enabling a more focused and efficient analysis of potentially malicious activities. Tab. II presents the performance metrics of an intrusion detection model in Stage 1 for a validation set in the CICIoT2023.

TABLE II
TWO-STAGE APPROACH - STAGE 1

Category	Precision	Recall	F1-Score
Benign	0.9999	0.9999	0.9999
DDoS	1.0000	0.9999	1.0000
DoS	0.9998	0.9998	0.9998
Mirai	0.9999	1.0000	0.9999
		Accuracy	0.9999
		Macro Avg F1-Score	0.9999
		Weighted Avg F1-Score	0.9999

The model shows its effectiveness across several categories, including DDoS, DoS, and Mirai attacks. It achieves near perfect precision, recall, and F1-score scores for all categories, indicating flawless identification of benign and malicious traffic. Specifically, for benign traffic, DDoS, DoS, and Mirai attacks, the model correctly identifies every instance with very few false positives or false negatives. The overall accuracy of 99.99% confirms that the model classifies the dataset accurately. Both the macro and weighted averages of precision, recall, and F1-scores are also almost perfect, showing uniformly excellent performance across all categories, irrespective of the number of instances. These results highlight the model's robustness and reliability in detecting and categorizing network traffic in the initial stage, pointing out that it can effectively handle various types of network activities.

In the second stage, we excluded the Denial of Service attacks from our focus and train it exclusively on the remaining attacks plus the benign traffic. With fewer than 2 million records in this stage, we could train a model using the

complete dataset. The classes in this dataset are more balanced, which led to improved results compared to the baseline. We observed promising results, especially for Brute Force and Web attacks as shown in Table III.

TABLE III
TWO-STAGE APPROACH - STAGE 2

Category	Precision	Recall	F1-Score
Benign	0.9344	0.9799	0.9566
BruteForce	0.9394	0.6373	0.7594
Recon	0.9176	0.8703	0.8933
Spoofing	0.9320	0.8883	0.9097
Web	0.9332	0.6179	0.7435
		Accuracy	0.9310
		Macro Avg F1-Score	0.8525
		Weighted Avg F1-Score	0.9310

TABLE IV
TWO-STAGE APPROACH (7+1 ATTACKS)

Category	Precision	Recall	F1-Score
Benign	0.9361	0.9804	0.9578
BruteForce	0.9409	0.6416	0.7629
DDoS	1.0000	1.0000	1.0000
DoS	0.9998	0.9998	0.9998
Mirai	0.9999	1.0000	1.0000
Recon	0.9177	0.8720	0.8942
Spoofing	0.9321	0.8915	0.9114
Web	0.9333	0.6155	0.7418
		Accuracy	0.9971
		Macro Avg F1-Score	0.9085
		Weighted Avg F1-Score	0.9971

The results of each stage are reflected in the Two-Staged model as shown in Table IV and Table V. The overall accuracy of 0.9971 for the 7+1 (DoS, DDoS, Mirai, Web Attack, Brute Force, Reconnaissance, Spoofing, and Benign) attacks and 0.9962 for 33+1 (detailed subcategories of the 7 main categories), along with strong macro and weighted averages, indicates that the model is highly effective in classifying most types of network traffic. It reliably identifies benign and specific attack traffic while highlighting areas for further improvement in Brute Force and Web attack detection.

An observation stand out: our initial hypothesis that the first stage would effectively recognize Denial of Service attacks and pass the remaining flows to the second stage was confirmed. Due to the strong performance of the first stage, the metrics for the second stage alone remained largely unchanged in the full two-stage model. This demonstrates the effectiveness of our staged approach in maintaining high detection accuracy across different attack types.

To comprehensively evaluate the performance of our proposed two-stage model, Table VI presents the model metrics for both the 7+1 and 33+1 attack scenarios. Additionally, we provide execution performance metrics, relative to the execution on a Intel Xeon Gold 6148 processor, to demonstrate the computational efficiency of our approach compared to a single model approach.

In both attack scenarios, the two-stage model demonstrates a slight improvement in accuracy and a noticeable increase

TABLE V
TWO-STAGE APPROACH (33+1 ATTACKS)

Category	Precision	Recall	F1-Score
Backdoor_Malware	0.8634	0.5592	0.6788
BenignTraffic	0.9183	0.9880	0.9519
BrowserHijacking	0.9559	0.6373	0.7647
CommandInjection	0.9528	0.6439	0.7685
DDoS-ACK_Fragmentation	0.9999	0.9999	0.9999
DDoS-HTTP_Flood	0.9983	0.9986	0.9984
DDoS-ICMP_Flood	1.0000	1.0000	1.0000
DDoS-ICMP_Fragmentation	0.9998	0.9998	0.9998
DDoS-PSHACK_Flood	1.0000	1.0000	1.0000
DDoS-RSTFINFlood	1.0000	1.0000	1.0000
DDoS-SYN_Flood	0.9999	0.9999	0.9999
DDoS-SlowLoris	0.9973	0.9983	0.9978
DDoS-SynonymousIP_Flood	0.9998	0.9999	0.9999
DDoS-TCP_Flood	1.0000	1.0000	1.0000
DDoS-UDP_Flood	0.9999	0.9998	0.9998
DDoS-UDP_Fragmentation	0.9996	0.9997	0.9996
DNS_Spoofing	0.8663	0.8031	0.8335
DictionaryBruteForce	0.9364	0.6434	0.7628
DoS-HTTP_Flood	0.9992	0.9992	0.9992
DoS-SYN_Flood	0.9998	0.9998	0.9998
DoS-TCP_Flood	0.9999	0.9999	0.9999
DoS-UDP_Flood	0.9997	0.9998	0.9997
MITM-ArpSpoofing	0.9416	0.8748	0.9070
Mirai-greeth_flood	0.9998	0.9999	0.9999
Mirai-greip_flood	0.9997	0.9998	0.9998
Mirai-udpplain	1.0000	1.0000	1.0000
Recon-HostDiscovery	0.8811	0.8934	0.8872
Recon-OSScan	0.8832	0.6808	0.7689
Recon-PingSweep	0.9021	0.5904	0.7137
Recon-PortScan	0.8509	0.7123	0.7755
SqlInjection	0.9038	0.5843	0.7097
VulnerabilityScan	0.9203	0.4739	0.6256
XSS	0.9981	0.9988	0.9985
Uploading_Attack	0.9085	0.5013	0.6461
		Accuracy	0.9962
		Macro Avg F1-Score	0.9055
		Weighted Avg F1-Score	0.9962

TABLE VI
COMPARISON OF TWO-STAGE VS SINGLE-STAGE BASELINE MODELS

Model	Accuracy	Macro F1	Predict Time	Nodes	Size
7+1 Attacks					
Baseline	99.66%	0.8836	27 sec	28.1M	3.35GB
Two-Stage Model	99.71%	0.9085	8 sec	3.5k + 22.1M	2.14GB
33+1 Attacks					
Baseline	99.56%	0.8370	1 min 34 sec	44.9M	14.07GB
Two-Stage Model	99.62%	0.9055	13 sec	5.0k + 29.6M	5.07GB

in the Macro F1 score compared to the baseline model. This suggests that the two-stage model is better at distinguishing between different attack types and has a lower false positive rate. Regarding computational efficiency, the two-stage model shows a significant advantage in prediction time, particularly in the 33+1 attacks scenario where it is considerably faster than the baseline model. Additionally, the two-stage model is substantially smaller in size and uses fewer nodes than the single-stage model, indicating a more efficient use of computational resources.

Overall, the results suggest that the two-stage model outperforms the baseline model in terms of accuracy, F1 score, and computational efficiency, making it a promising approach for intrusion detection in complex network environments with

numerous attack types. We also calculate the feature importance of the models. Inter-Arrival Time (IAT) has high feature importance, contributing 50% to the baseline model and 30% to both stages of our two-stage model, highlighting its critical role in intrusion detection.

VI. Conclusion and Future Research Direction

The proposed two-stage NIDS provides an alternative to traditional single-layer models for protecting IoT networks. Its modular design improves efficiency and accuracy in detecting and classifying malicious activity.

Empirical data shows that the two-stage NIDS performs better in critical metrics. It has a detection rate 90% compared to the single-layer model's 88%. This enhances the identification and mitigation of security threats in IoT environments. Additionally, the modular structure reduces processing and inference times, allowing faster responses to security incidents, vital for real-time IoT applications.

Future research should focus on three areas to enhance IoT ecosystem security.

First, it is essential to understand how attacks occur across different architectural layers. This involves studying how a breach in one layer creates vulnerabilities in others and analyzing attack patterns and signatures at each layer. This knowledge will help develop intrusion detection systems to identify and respond to multi-layered attacks.

Second, improving the generalizability of attack detection models is crucial. Current models rely on tool-specific signatures, which can be bypassed with different tools. Researchers should develop models recognizing broader malicious behaviors by using diverse training data and creating feature representations reflecting the attacks' intent. This will make intrusion detection systems more resilient to new and evolving threats.

Third, testing proposed security approaches with various datasets is essential to ensure their effectiveness across different IoT environments. Using datasets such as Bot-IoT, UNSW-NB15, and CICIDS2017, which include different device types, network setups, and attack scenarios, will evaluate the robustness of security measures. Additionally, comparing these approaches with existing solutions and applying them to real-world data will refine the strategies, ensuring they can handle a range of threats and adapt to changes in IoT architectures.

Acknowledgments

This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq proc. 465446/2014-0, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, FAPESP proc. 14/50937-1, FAPESP proc. 15/24485-9 and FAPESP proc. 22/15304-4. It is also part of the FAPESP proc. 2018/23098-0 and FAPESP proc. 2021/06995-0.

References

[1] Zeeshan Ahmad, Adnan Shahid Khan, Cheah Wai Shiang, Johari Abdullah, and Farhan Ahmad. Network Intrusion Detection System: A Systematic Study of Machine Learning and Deep Learning Approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1):e4150, 2021.

[2] Jahongir Azimjonov and Taehong Kim. Designing Accurate Lightweight Intrusion Detection Systems for IoT Networks Using Fine-Tuned Linear SVM and Feature Selectors. *Computers & Security*, 137:103598, 2024.

[3] João Gabriel Andrade de Araújo Josephik, Yaissa Siqueira, Kétly Gonçalves Machado, Routo Terada, Aldri Luiz dos Santos, Michele Nogueira, and Daniel Macêdo Batista. Applying Hoeffding Tree Algorithms for Effective Stream Learning in IoT DDoS Detection. In *Proc. of the IEEE LATINCOM*, 2023.

[4] K. Ruwani M. Fernando and Chris P. Tsokos. Dynamically Weighted Balanced Loss: Class Imbalanced Learning and Confidence Calibration of Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 33(7):2940–2951, 2022.

[5] Suzan Hajj, Joseph Azar, Jacques Bou Abdo, Jacques Demerjian, Christophe Guyeux, Abdallah Makhoul, and Dominique Ginhac. Cross-Layer Federated Learning for Lightweight IoT Intrusion Detection Systems. *Sensors*, 23(16):7038, 2023.

[6] Mosab Hamdan, Arwa M. Eldhai, Samah Abdelsalam, Kifayat Ullah, Ali Kashif Bashir, MN Marsono, Fabio Kon, and Daniel Macêdo Batista. A Two-Tier Anomaly-based Intrusion Detection Approach for IoT-Enabled Smart Cities. In *Proc. of the IEEE INFOCOM Workshops (INFOCOM WKSHPS)*, 2023.

[7] Safwana Haque, Fadi El-Moussa, Nikos Komninos, and Rajarajan Muttukrishnan. A Systematic Review of Data-Driven Attack Detection Trends in IoT. *Sensors*, 23(16):7191, 2023.

[8] Akinul Islam Jony and Arjun Kumar Bose Arnob. A Long Short-Term Memory Based Approach for Detecting Cyber Attacks in IoT Using CIC-IoT2023 Dataset. *Journal of Edge Computing*, 3:28–42, May 2024.

[9] Maryam Mahsal Khan and Mohammed Alkhathami. Anomaly Detection in IoT-based Healthcare: Machine Learning for Enhanced Security. *Scientific Reports*, 14:5872, 2024.

[10] Constantinos Kolias, Georgios Kambourakis, Angelos Stavrou, and Jeffrey Voas. DDoS in the IoT: Mirai and Other Botnets. *Computer*, 50(7):80–84, 2017.

[11] Wentao Mao, Zongtao Chen, Yanna Zhang, and Zhidan Zhong. Harmony Better than Uniformity: A New Pre-Training Anomaly Detection Method with Tensor Domain Adaptation for Early Fault Evaluation. *Engineering Applications of Artificial Intelligence*, 127:107427, 2024.

[12] Kg Raghavendra Narayan, Srijanee Mookherji, Vanga Odelu, Rajendra Prasath, Anish Chand Turlapaty, and Ashok Kumar Das. IIDS: Design of Intelligent Intrusion Detection System for Internet-of-Things Applications. In *Proc. of the 7th IEEE CICT*, 2023.

[13] Euclides Carlos Pinto Neto, Sajjad Dadkhah, Raphael Ferreira, Alireza Zohourian, Rongxing Lu, and Ali A. Ghorbani. CICIoT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment. *Sensors*, 23(13):5941, June 2023.

[14] Omid Setayeshfar, Karthika Subramani, Xingzi Yuan, Raunak Dey, Dezhi Hong, In Kee Kim, and Kyu Hyung Lee. Privacy Invasion via Smart-Home Hub in Personal Area Networks. *Pervasive and Mobile Computing*, 85:101675, 2022.

[15] Pietro Spadaccino and Francesca Cuomo. Intrusion Detection Systems for IoT: Opportunities and Challenges Offered by Edge Computing. *ITU Journal on Future and Evolving Technologies*, 3(2):408–420, 2022.

[16] Ankit Thakkar and Ritika Lohiya. Attack Classification of Imbalanced Intrusion Data for IoT Network Using Ensemble-Learning-Based Deep Neural Network. *IEEE Internet of Things Journal*, 10(13):11888–11895, 2023.

[17] Sami Yaras and Murat Dener. IoT-Based Intrusion Detection System Using New Hybrid Deep Learning Algorithm. *Electronics*, 13(6):1053, 2024.