Enhancing IoT Security: Intrusion Detection System Using Machine Learning for Cross-Environment Attack Identification

Sagor Sikdar
Computer Science
Stony Brook University
Stony Brook, USA
sagor.sikdar@stonybrook.edu

Priscilla Kyei Danso
Computer Science
Stony Brook University
Stony Brook, USA
priscilla.danso@stonybrook.edu

P.A.C. Abisheka
Computer Science
Stony Brook University
Stony Brook, USA
Abisheka.Pitumpe@stonybrook.edu

G M Tasnim Alam
Computer Science
Stony Brook University
Stony Brook, USA
gmtasnim.alam@stonybrook.edu

***Abstract*—The heterogeneity of connectivity and communication protocols, coupled with the sheer volume of data generated by Internet of Things (IoT) devices, creates a complex landscape that is challenging to secure. The diversity of devices and communication protocols within IoT networks complicates the task of identifying compromised devices, particularly as the number of connected devices increases. While similar research has been conducted in the past, this study aims to address the challenge of misclassifications when training and testing IDS models in different environments. Our primary objective is to generate robust features that can excel in training and testing environments distinct from one another. By doing so, we aim to enhance the security of IoT ecosystems and contribute to the development of more reliable intrusion detection mechanisms. The CICIoT2023 and MedIoTBot datasets will be used in the evaluation of our study. Our findings indicate that .....**

***Index Terms*—Intrusion Detection, Transferability, Machine Learning, Cross-Environment, IoT Security**

I. Introduction

Internet of Things devices have become a part of our daily lives. Prompting Alexa to narrate what the weather will be like tomorrow to locking smart doors from our office when we forget to lock our homes that iRobot independently vacuuming our homes to our automobile equipped with built-in sensors to alert us when tire pressure is low [7]. It is evident IoT devices have come to make our lives easier and better. With such comfort and ease comes with somewhat a price to pay in terms of security and privacy. The Internet of Things popularly defined by the Internet Engineering Task Force (IETF) as the, *"network of physical objects, or 'things' embedded with electronics, software, sensors, actuators, and connectivity to enable objects to exchange data with the manufacturer, operator, and/or other connected devices"* [18]. This interconnectedness has become so invasive that according to Statista reports estimates between 2021 and 2025, the global deployment of IoT devices is predicted to reach 30.9 billion units, up from 13.8 billion deployed in 2021 [6]. This exponential growth is making the globe more linked than ever. The diverse range of linked IoT devices means there are numerous vulnerabilities and potential attack methods that can impair their functionality. Notable cyberattacks on susceptible IoT devices have inflicted significant harm. One prominent example is the attack on the Dyn DNS provider. In October 2016, the Mirai botnet assault still considered the most substantial Distributed Denial of Service (DDoS) event targeting IoT devices, brought down Dyn's DNS servers [2]. This led to the Mirai cyberattack, which transformed devices from smart homes into "botnets" or "zombie armies" [5]. As a result, many high-profile sites like Amazon, CNN, Github, Netflix, Paypal, and Twitter were inaccessible for several hours [8]. The German government's Federal Office for Information Security revealed an incident in which an attacker infiltrated a steel factory in December 2014. They first entered the company's network via a spear phishing email before moving on to the system at the facility. Once inside, they caused several system components to fail. As a result, critical processes were uncontrolled, resulting in considerable physical harm [16]. Most research tackles the aforementioned issues using an Intrusion Detection System (IDS) specifically the anomaly-based IDS. The limitation of the rule-based IDS where only known attacks are predicted in the era of zero days attacks makes them unsuitable for anomaly detection. Anomaly-based IDS, however, *observes a sequence of incoming events while the system is not under attack and builds a model of the system's normal behavior and uses the trained model to detect anomalies based on a similarity index between normal and abnormal observations* [9]. Anomaly detection in the Internet of Things (IoT) using machine learning

is a well-researched topic. A less research area is the challenge of training and testing models in different environments, often referred to as domain adaptation or transfer learning, which is particularly relevant in IoT scenarios. This is because, in real-world settings, IoT devices might be deployed in diverse environments, and the data distribution during testing may not always mirror the training data distribution.

This study proposes an Intrusion Detection System (IDS) utilizing Machine Learning to uniquely identify various attack types within IoT networks and develop a methodology for distinguishing network traffic of the trained model in one domain and using the trained model for anomaly detection in another domain. Our primary objective is to demonstrate and mitigate the issue highlighted by Haque et al. [13] by generating robust features that can excel in training and testing environments distinct from one another. By doing so, we aim to enhance the security of IoT ecosystems and contribute to the development of more reliable intrusion detection mechanisms. To the best of our knowledge, this is the first research to tackle anomaly detection with the underlying concept of domain adaption. The main contributions of this article can be summarized as follows.

- We present a domain-adaptative anomaly-based IDS able to differentiate the behavior of incoming IoT traffic and make a prediction of the type of anomaly.
- We extract strong and robust features to be employed in our study. Due to the overly unique approach of our study, the only way our model is going to excel in an environment where the model makes a prediction in our target domain based on the performance of the model in the source domain requires efficient features.
- We evaluate our approach in two publicly available datasets. With the CICIoT acting as our source domain and the MedBIoT acting as our target domain.
- We explore different Machine Learning approaches in our study to achieve the best results in our domain-adaptative anomaly-based IDS predictions.

II. Literature Review

A. Machine Learning for Anomaly Detection

In the IoT anomaly detection field, challenges revolve around the scarcity of labeled data due to irregular anomalies and the complexity of obtaining real system data. There is a significant gap in formalizing the process from data collection to model development and validation in real-world settings. Existing methods [10] mainly involve training models on normal behavior and identifying anomalies, but more precise techniques are required for complex, real-world datasets. The availability of suitable, regularly updated datasets with a variety of normal and abnormal behaviors is crucial for training and validating real-time anomaly detection methods. Current datasets often suffer from labeling errors and are not well-suited for real-time detection. To address this, new datasets should replicate realistic scenarios and include ground truth data to enhance their credibility for testing new anomaly detection systems. Developing a model for anomaly detection is a challenging task due to data complexity, which includes issues like imbalanced datasets, unexpected noises, and data redundancy. Effective approaches for data curation are necessary to gather valuable information and knowledge from such datasets.

Numerous machine learning anomaly detection techniques for data streams have been proposed in various fields, such as manufacturing, finance, military, healthcare, and the Internet of Things. These techniques aim to identify anomalies in real-time data streams. Several methods have been introduced, such as Cumulative Local Outlier Factor (C-LOF), AutoCloud, and a multi-kernel approach. They address different aspects of anomaly detection but have limitations, including high time complexity and issues related to evolving data.

Other approaches, like xStream, focus on feature evolution in data streams and offer promising results. However, some challenges, like evolving clusters, remain unaddressed. Additionally, methods for anomaly detection in specific domains, such as air traffic control and road safety, have been developed with varying degrees of success.

Hierarchical Temporary Memory (HTM) algorithms have also been employed for real-time anomaly detection in resource utilization and network metrics. While these methods show promise in catching anomalies, accuracy and latency improvements are needed.

In the context of evolving multi-density data streams, the MuDi-Stream approach has been proposed, which maintains summary information and uses density-based clustering algorithms. However, its performance is affected by the number of empty grids and high-dimensional data.

Fast anomaly detection algorithms using Extreme Learning Machine (ELM) have been developed for aviation datasets but lack comparison to state-of-the-art techniques. Another approach called AMAD focuses on dynamic anomaly detection in evolving attributed networks and performs attribute selection. However, it does not address high dimensionality-related challenges.

Recent advancements in anomaly detection leverage deep learning techniques. Notable methods include KPI-TSAD for addressing imbalanced classification in non-stationary time series, Streaming Autoencoder (SA) for evolving data streams, OFAT for accurate long-range forecasting in high-dimensional, feature-evolving time series, e-SREBOM for real-time malware detection, ISTL for real-time video surveillance, and an e-SREBOM variant for efficient malware classification. While these methods show promise in different contexts, some may have limitations such as reliance on labeled data, scalability issues, or a lack of real-time interactivity.

The choice of an anomaly detection technique depends on the nature of the data under analysis, particularly when dealing with data streams, a common source of which is the Internet of Things (IoT). Data streams are characterized by continuous and unlimited data flow with evolving patterns over time. Detecting anomalies in data streams is crucial across various fields due to its practical applications and the need for

precision and real-time responses. Numerous learning algorithms have been proposed for data stream anomaly detection, including methods like C-LOF [22], AutoCloud [4], TEDA Clustering, evolving spiking neural networks, KPI-TSAD [20], HTM [22], ensembles neural networks, multiple kernel learning, xStream [17], CEDAS, MuDi-Stream [1], Long Short-Term Memory, and density-based clustering, among others.

Supervised anomaly detection involves creating grouping rules for predicting anomalies. It can be classified into classification and regression approaches. Examples include the use of SVM for real-time alerts in smart transportation and Naive Bayes in smart homes for intrusion detection. The K-Nearest Neighbor (k-NN) algorithm is applied to cyber-physical attack detection in industrial systems. Regression-based models, such as the Decision Tree approach, are used in intelligent transportation to identify correlations between accidents and resource measurements.

Semi-supervised anomaly detection classifies anomalies within standard data. Methods include Bayesian networks for monitoring maternal sleep in healthcare and reinforcement learning in intelligent transportation for detecting motor outliers. Temporal and spatial-temporal techniques are suggested in the smart object domain.

Unsupervised anomaly detection deals with unlabeled information and doesn't require separate training and testing. Clustering-based detection methods, like K-means clustering, Gaussian mixture models (GMM), and hidden Markov models (HMM), are used in various domains such as gas turbine engine anomaly detection in transportation, detecting attacks in smart cities' IoT architecture, and real-time sleep anomaly analysis in healthcare. These methods help identify normal data patterns in diverse datasets.

B. Domain adaptation or transfer learning for Anomaly Detection

Recent advancements in securing the Industrial Internet of Things (IIoT) and Network Intrusion Detection (NID) have seen a surge in research focused on innovative machine learning techniques. One prominent avenue is the exploration of Malware traffic classification (MTC) methods, as evidenced by the introduction of ConvLaddernet and KT-ConvLaddernet [15]. These models, though effective, exhibit high complexity, demanding careful consideration of their practical implementation. Additionally, the transition to deeper architectures, like VGG networks, underscores the need to balance complexity with computational efficiency.

Addressing the challenge of scarce labeled samples in NID, the use of transfer learning has become prevalent. The proposal of an adversarial Domain Adaptation (DA) [3] approach stands out in this regard. This approach facilitates the training of Deep Learning (DL) classification models for NID with minimal labeled data, offering organizations the ability to identify new attack families by leveraging existing datasets. Experimental results showcase the superiority of the adversarial DA approach over traditional methods, including transfer learning using fine-tuning, in both homogeneous and heterogeneous DA scenarios. Looking ahead, the literature suggests avenues for future research. The exploration of semi-supervised and unsupervised DA scenarios, where labeled data in the target dataset varies, presents intriguing possibilities. Additionally, the efficacy of adversarial DA approaches in multi-class classification, distinguishing specific attack categories, warrants investigation. Evaluating more sophisticated Generative Adversarial Network (GAN) [14] architectures, such as Wasserstein GANs, can further enhance the robustness of these methods.

In parallel, the literature emphasizes the need for pragmatic considerations in the deployment of these models. Complex models like VGG networks may demand substantial computing resources, necessitating thoughtful simplification without compromising accuracy. Moreover, the reliance on pretraining models, a common theme in transfer learning and DA, poses challenges in practical applications. Privacy concerns arising from the sharing of source datasets prompt the call for privacy-preserving DA techniques, particularly in scenarios resembling federated learning.

In summary, while the proposed MTC methods and adversarial DA approaches present promising strides in enhancing IIoT security and NID, the literature highlights the importance of addressing challenges in model complexity, practical implementation, and privacy preservation. Future research endeavors should strive for solutions that strike a balance between sophistication and applicability, advancing the field toward secure, efficient, and widely applicable solutions. With our approach we plan to mitigate these problems and use the publicly available datasets to produce strong and robust features.

In a paper entitled "A Systematic Review of Data-Driven Attack Detection Trends in IoT" [12], the authors emphasized the significance of identifying distinctive features within IoT traffic that set it apart from conventional networks. They also highlighted the challenges that come with datasets made using a small number of IoT devices which include issues such as overfitting which could lead to inaccurate results. The authors emphasized the importance of exploring the connections between attacks occurring at various architectural layers of IoT systems which can significantly inform the design of more accurate and targeted Intrusion Detection Systems (IDSs).

Presently available datasets except for MedBIoT and CICIDS2023 which are considered large, can lead models to classify benign traffic as harmful and vice versa. It is therefore a main concern in developing IDS models that has enough data it could rely on to provide a correct classification.

III. Domain adaptation or transfer learning

IV. The Overall Proposed Framework

V. The Experimental System Flow of the Proposed Framework

A. Data Preprocessing

B. Feature Engineering

C. Feature Selection

D. Model Training

E. Transfer IDS

VI. Dataset Description

A. CICIDS2023 [19]

B. MedIoTBoT [11]

VII. Experimental Results and Analysis

A. Experiments

This section evaluates the data collection, data sampling or reduction technique used in this research work as well as the feature selection technique. Additionally, we describe our approach in the training and testing of the machine learning models and the performance metrics used on the training dataset.

1) Data collection: CIC IoT dataset 2022 [19] functions as the MedBIoT [11] The CIC IoT dataset 2023 [1] contains the feature set that makes up the generated CSV files, as well as descriptions and definitions of each feature. The collection of features created for the source domain and the target domain are identical.

2) Data sampling: An unbalanced dataset has a detrimental effect on the precision of class predictions in most classification problems [21]. This skewness was addressed in this study by randomly selecting an equal number of records for each class to guarantee that each distribution class had an equal probability of being chosen.

3) Feature selection:

4) Training and Testing of the ML methods: Scikit learn[2], a Python library, is used in the implementation of the IDS. To determine which machine learning technique performs the best, six machine learning models are utilized to train and test the dataset for both when the data is from the same distribution or a different distribution. When using data from the same distribution, we first cross-validate each of our six machine learning classifiers using a 7:3 split, which involves training on 70% of the data that was randomly chosen and testing on the remaining 30%.

[1]https://www.unb.ca/cic/datasets/iotdataset-2023.html

[2]https://scikit-learn.org/stable/

5) Performance measuring metrics: In this section, we outline five benchmark performance indicators for assessing how well our classifiers performed on the dataset with the training and test data taken from the set. Macro-averaging is employed for the overall performance of precision, recall, and F1-score. For datasets with training and testing data drawn from a different distribution (lab), a different metric is used to evaluate the device identification performance.

Confusion matrix is a visual depiction of the model's performance in predicting the various classes.

Accuracy is the proportion of accurate predictions to all observations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision is calculated as the number of correctly identified device types by the number of correctly or incorrectly predicted device types by the classifier.

$$Precision = \frac{TP}{TP + FP}$$

Detection Rate shows how many actual device types the classifier classified properly.

$$Recall = \frac{TP}{TP + FN}$$

F1-score is a geometric average of precision and recall.

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

B. Results and Analysis

1) CIC IoT dataset 2023:

2) MedBIoT Dataset:

C. Comparison or Discussions

VIII. Conclusion and Future Work

Acknowledgment

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks . . .". Instead, try "R. B. G. thanks. . .". Put sponsor acknowledgments in the unnumbered footnote on the first page.

References

[1] Amineh Amini, Hadi Saboohi, Tutut Herawan, and Teh Ying Wah. Mudi-stream: A multi density clustering algorithm for evolving data stream. *Journal of Network and Computer Applications*, 59:370–385, 2016.

[2] Kishore Angrishi. Turning internet of things (iot) into internet of vulnerabilities (iov): Iot botnets. *arXiv preprint arXiv:1702.03681*, 2017.

[3] Elisa Bertino Ankush Singla and Dinesh Verma. Preparing network intrusion detection deep learning models with minimal data using adversarial domain adaptation. In *ACM Conference on Computer and Communications Security ASIA*, 2020.

[4] Clauber Gomes Bezerra, Bruno Sielly Jales Costa, Luiz Affonso Guedes, and Plamen Parvanov Angelov. An evolving approach to data streams clustering based on typicality and eccentricity data analytics. *Information Sciences*, 518:13–28, 2020.

[5] Nadia Chaabouni, Mohamed Mosbah, Akka Zemmari, Cyrille Sauvignac, and Parvez Faruki. Network intrusion detection for iot security based on learning techniques. *IEEE Communications Surveys Tutorials*, 21(3):2671–2701, 2019.

[6] Statista Research Department. Internet of things (iot) and non-iot active device connections worldwide from 2010 to 2025 (in billions) [graph]. 2020.

[7] Nomusa Nomhle Dlamini and Kevin Johnston. The use, benefits and challenges of using the internet of things (iot) in retail businesses: A literature review. In *2016 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pages 430–436, 2016.

[8] Rohan Doshi, Noah Apthorpe, and Nick Feamster. Machine learning ddos detection for consumer internet of things devices. In *2018 IEEE Security and Privacy Workshops (SPW)*, pages 29–35, 2018.

[9] Mojtaba Eskandari, Zaffar Haider Janjua, Massimo Vecchio, and Fabio Antonelli. Passban ids: An intelligent anomaly-based intrusion detection system for iot edge devices. *IEEE Internet of Things Journal*, 7(8):6882–6897, 2020.

[10] Muhammad Fahim and Alberto Sillitti. Anomaly detection, analysis and prediction techniques in iot environment: A systematic literature review. *IEEE Access*, 7:81664–81681, 2019.

[11] Alejandro Guerra-Manzanares, Jorge Medina-Galindo, Hayretdin Bahsi, and Sven Nõmm. Medbiot: Generation of an iot botnet dataset in a medium-sized iot network. In *International Conference on Information Systems Security and Privacy*, 2020.

[12] S. Haque, F. El-Moussa, N. Komninos, and R. Muttukrishnan. A systematic review of data-driven attack detection trends in iot. 23(16):7191, 2023.

[13] Safwana Haque, Fadi El-Moussa, Nikos Komninos, and Rajarajan Muttukrishnan. A systematic review of data-driven attack detection trends in iot. *Sensors*, 23(16), 2023.

[14] Mehdi Mirza Bing Xu David Warde-Farley Sherjil Ozair Aaron Courville Yoshua Bengio Ian Goodfellow, Jean Pouget-Abadie. Generative adversarial nets. 2014.

[15] Bamidele Adebisi Jinhui Ning, Yu Wang and Haris Gacanin. Malware traffic classification using domain adaptation and ladder network for secure industrial internet of things. 9(18), 2022.

[16] Robert M Lee, Michael J Assante, and Tim Conway. German steel mill cyber attack. *Industrial Control Systems*, 30(62), 2014.

[17] Emaad Manzoor, Hemank Lamba, and Leman Akoglu. xstream: Outlier detection in feature-evolving data streams. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1963–1972, 2018.

[18] Roberto Minerva, Abyi Biru, and Domenico Rotondi. Towards a definition of the internet of things (iot). *IEEE Internet Initiative*, 1(1):1–86, 2015.

[19] 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), 2023.

[20] Juan Qiu, Qingfeng Du, and Chongshu Qian. Kpi-tsad: A time-series anomaly detector for kpi monitoring in cloud applications. *Symmetry*, 11(11):1350, 2019.

[21] Shivani Tyagi and Sangeeta Mittal. Sampling approaches for imbalanced data classification problem in machine learning. In *Proceedings of ICRIC 2019*, pages 209–221. Springer, 2020.

[22] Kangqing Yu, Wei Shi, and Nicola Santoro. Designing a streaming algorithm for outlier detection in data mining—an incremental approach. *Sensors*, 20(5):1261, 2020.