Diversity Measure to Tackle the Multiclass Problem in IoT Intrusion Detection Systems

Osama A. Mahdi
School of Information
Technology and Engineering
Melbourne Institute of
Technology
Melbourne, Australia
omahdi@mit.edu.au

Nawfal Ali
Faculty of Information
Technology
Monash University
Melbourne, Australia
nawfal.ali@monash.edu

Ammar Alazab
School of Information
Technology and Engineering
Melbourne Institute of
Technology
Melbourne, Australia
aalazab@mit.edu.au

Savitri Bevinakoppa
School of Information
Technology and Engineering
Melbourne Institute of
Technology
Melbourne, Australia
sbevinakoppa@mit.edu.au

Tahsien Al-Quraishi
School of Information
Technology and Engineering
Melbourne Institute of Technology
Melbourne, Australia
tquraishi@mit.edu.au

Bhagwan Das
School of Information
Technology and Engineering
Melbourne Institute of Technology
Melbourne, Australia
bdas@mit.edu.my

Abstract— The advent of the Internet of Things (IoT) has instigated transformations in various domains, such as healthcare, smart homes, agriculture, transportation, and manufacturing. With the swift proliferation of IoT networks, new security challenges have surfaced, exposing them to a plethora of attacks. To counter these, machine learning-driven intrusion detection strategies have been introduced, which scrutinize the behavior and communication patterns of IoT devices to identify and nullify any suspicious undertakings. While these methodologies demonstrate high accuracy and minimal false alarm rates in static scenarios, their performance stability in dynamic, evolving environments remains undetermined. One critical issue pertains to multiclass problems, wherein the complexity of diverse attack types can significantly affect the efficacy of machine learning-based intrusion detection systems, if not promptly recognized and addressed. This paper introduces an innovative IoT Intrusion Detection approach that incorporates the Diversity measure as a model drift detection method to tackle the multiclass problem in IoT networks. Our proposed approach can detect previously unknown attacks in IoT networks through an advanced drift detection technique.

Keywords— Concept Drift, IoT, Intrusion Detection System, Data Drift, Machine Learning

I. INTRODUCTION

The rise of the Internet of Things (IoT) has revolutionized our interaction with technology, where real-time data gathering and analysis occur across devices and sensors [1] [2]. However, this increased interconnectivity brings new security threats to the forefront. Cyber threats directed at IoT devices can lead to serious consequences, such as loss of crucial data or control over systems [3]. In this context, Intrusion Detection Systems (IDS) serve a key role in detecting and countering cyber-attacks, though they often struggle with system adaptations and recognizing unfamiliar attacks.

Conventional IDS, based on predefined rules and patterns, prove insufficient for detecting emerging types of attacks [4] [5]. Moreover, the dynamic nature of data can lead to changes in its statistical properties, causing the IDS to lose effectiveness over time — a phenomenon referred to as model or concept drift. This drift can lead to false alarms or overlooked threats, resulting in compromised efficiency and escalating system maintenance costs [6].

To address these challenges, our research focuses on incorporating drift detection algorithms into IDS. These algorithms analyze incoming data, identifying changes in data distributions and allowing IDS to update in real time. This constant adaptation enables the system to identify and categorize unusual or potentially harmful activities, leading to improved system performance and security [7].

In addition, the problem of changing the definitions of classes over time decreases the performance (accuracy) of a predictive model which has been trained using old instances. Additionally, processing the multi-class problem is computationally more expensive, particularly in the presence of model drift in IDS where the data is changing over time, and this aggravates the problem of a loss of performance during the process of model drift detection in IoT intrusion detection systems. A solution for these problems is a mechanism which uses continuous diagnostics of model drift. Then, upon detection of concept drift, a process of updating the model to maintain the classification performance is required [1].

To illustrate the phenomenon of model drift in IoT Intrusion Detection Systems lets consider the following scenario. Model drift occurs when a model's predictive accuracy declines due to changes in the underlying data patterns. This phenomenon can be particularly challenging in dynamic environments, such as IoT Intrusion Detection Systems (IDS). Imagine a smart hospital leveraging IoT devices such as patient monitors, infusion pumps, and HVAC systems. An IDS is deployed, classifying network traffic into categories: 'Normal', 'External Intrusion', 'Device Malfunction', and 'Malware Activity'. Initially, the IDS excels in identifying 'External Intrusion' when attackers attempt exploiting known vulnerabilities in specific IoT devices. It's trained on patterns of direct attacks targeting these vulnerabilities.

Over months, attackers innovate. Instead of directly targeting devices, they use spear phishing tactics on hospital staff. When a staff member unknowingly clicks a malicious link, it installs malware on their device, which then seeks to compromise the IoT network from within. Suddenly, the IDS faces challenges. These new, internally-originated threats don't align with the earlier 'External Intrusion' patterns. They might get misclassified as 'Normal' since they emerge from trusted devices, or mistakenly labeled as 'Device Malfunction' due to the anomalous behavior of compromised IoT devices.

This shift in intrusion methodology represents model drift. The threat landscape evolved, but the IDS's model, anchored to its initial training, didn't adapt accordingly. To counteract this, a model drift detection method to tackle this problem in IoT networks, continuous model updates and retraining with current threat data are essential. This scenario emphasizes the critical importance of adaptability in cybersecurity, particularly when guarding against evolving threats in IoT environments.

Motivated by this context, the primary aim of this paper is to introduce an innovative IoT intrusion detection system. This system leverages data drift effectively, thereby responding seamlessly to model drifts within a multiclass problem scenario. In a novel way, our proposed approach analyzes data from IoT devices to detect potential intrusions, identifying shifts in data patterns that may signify a threat within a multiclass problem scenario. Instead of monitoring the error estimates, the proposed detector monitors the diversity of a pair of classifiers.

To surmount this hurdle, we propose a novel approach which merges the disagreement measure [8] [9] used in static learning with the Page-Hinkley test for identifying model drifts. By swiftly detecting drifts with minimal memory usage, the IDS can rapidly adjust to changing network environments and recognize unfamiliar threats. This groundbreaking research effort has the potential to significantly enhance the field of IoT intrusion detection systems by offering an innovative method for model drift detection and bolstering the accuracy and reliability of IDS over time within a multiclass problem scenario.

The structure of this paper is organized as follows. Section 2 introduces the diversity measure for multi-class problem. Section 3 outlines the research question and objective of this study. Subsequently, Section 3 encompasses the discussion, while Section 4 presents the conclusion.

II. Diversity measure for multi-class problem

The diversity measure [10] [11] [12] [13] as a model drift detector for the multiclass problem uses a pair of base learners/classifiers to detect detects attacks in IOT environment. We have already considered a preliminary version of a pair of base learners/classifiers as a drift detector (DMDDM, DMDDM-S), where a disagreement measure has been used with a PH test to detect sudden drifts in the binary classification problem [14] [15] [16]. In this current work, we incorporate a number of modifications, resulting in new contribution. First, it comes with a new formalism that facilitates the way to detect concept drifts in the multi-class problem.

let $X = x_1, . . . , x_n$ be a labelled data set and $y\hat{}_v = [y\hat{}_v (x_1), . . . , y\hat{}_v (x_n)]$ an n-dimensional binary vector that represents the output of a classifier h_v, such that $y\hat{}_v (x_j) = 1$, if h_v correctly predicts the class label and 0 otherwise. Table 1 presents (as it calls oracle outputs) all the possible outcomes for a pair of classifiers h_u and h_v, such that $h_u = h_v$, where N^{ab} is the number of instances $x_j \in X$ for which $y\hat{}_u (x_j) = a$ and $y\hat{}_v (x_j) = b$.

Therefore, all the probabilities of N^{ab} are as follows:

- N^{10} indicates the number of examples where *Classifier_i* predicted class 1 and *Classifier_j* predicted class 0
- N^{01} indicates the number of examples where *Classifier_j* predicted class 1 and *Classifier_i* predicted class 0
- N^{11} indicates the number of examples where *Classifier_i* predicted class 1 and *Classifier_j* predicted class 1
- N^{00} indicates the number of examples where *Classifier_i* predicted class 0 and *Classifier_j* predicted class 0

Table 1. A2 × 2 Of the Relationship Between A Pair Of Classifiers

$h_u = h_v$	h_u *correct (1)*	h_u *incorrect (0)*
h_v *correct (1)*	N^{11}	N^{10}
h_v *incorrect (0)*	N^{01}	N^{00}

In preliminary versions of this work (DMDDM, DMDDM-S) and for the binary classification problem, the disagreement measure $D_{v,u}$ Equation 1 has been used. However, for multi-class classification problems, using the output of Table 1 would be unsuccessful to measure the differences between a pair of classifiers that incorrectly predict the same instance using different labels. Consequently, as the contribution of this work, we propose a new way of tracking the classifiers' exact predictions instead of only the dichotomy correct/incorrect, as has been done in preliminary versions of this work. Therefore, to capture the variation of a pair of classifiers precisely, we construct a table $C_{i,j}$, where each value at the intersection of row i and column j holds the number of instances $x \in X$, where $h_v(x) = i$ and $h_u(x) = j$. Table 2 shows an example of contingency table Ci,j for the k-class problem.

$$D_{u:v} = N^{10} + N^{01} \qquad 1$$

From Table 2, the concomitant decisions of the pair of classifiers are stored in the diagonal in matrix $C_{i,j}$. Thus, in order to weight their similarity, Equation 2 is used to find the summation of its values and divide it by the total number of instances n.

$$\theta = \frac{1}{n}\sum_{i=0}^{K}(C_{i,j}) \qquad 2$$

In addition, in order to signal if there is a drift, we use the PH test from our preliminary work, as shown in Equation 3 and Equation 4.

$$M_T = \sum_{t=0}^{kT} (x_t - x'_T - \delta) \qquad 3$$

$$PH_T = m_T - M_T \qquad 4$$

As a result, when this difference exceeds a specified threshold (λ), a drift is indicated.

III. Research Question

Concept drift has become a growing field of study and advancement in the domain of Intrusion Detection Systems (IDS) for Internet of Things (IoT) networks, focusing on understanding and adapting to the continuously evolving characteristics of data and its underlying concepts. Previous approaches to detect concept drift can be categorized into three primary categories: *Statistical-based Methods* [17] [18] [19] [20], *Window-based Methods* [21] [22] [23], and *Ensemble-based Methods* [24] [14] [25]. Some studies have explored the use of concept drift to develop more effective IDSs.

For example, [26] presented a lightweight framework that utilizes a sliding window technique to identify and adjust to concept drift in IoT data streams. The framework monitors the statistical distribution of the data and employs an online machine learning algorithm to continuously update the model based on new information. Through experiments conducted on two real-world IoT datasets, the authors showcased the efficacy of the proposed framework, exhibiting high accuracy in detecting and adapting to concept drift while ensuring minimal computational complexity and memory usage.

[27] suggested an ensemble-based approach was introduced to integrate multiple online machine learning algorithms in order to create a more precise and resilient intrusion detection system (IDS). The method incorporates feature selection techniques to reduce the dimensionality of the data and improve the performance of the algorithms. Through evaluation using various real-world datasets, the results demonstrated that the proposed approach outperforms individual online machine learning algorithms in terms of accuracy and detection rate. Additionally, the method exhibits effectiveness in identifying intrusions that were previously undetected, highlighting its significance in maintaining network security.

[28] focused on the detection of Botnet cyber-attacks in the context of Internet of Things (IoT) by investigating concept drift. They introduced a method based on dynamic residual projection for efficiently recognizing cyber-attacks. To evaluate the effectiveness of their approach, they conducted experiments using the Bot-IoT dataset, and the results demonstrated enhanced detection accuracy compared to classification models that did not incorporate concept drift analysis.

Table 2: Output of a Pair of Classifiers for the Multi-class Classification Problem

	hu (x) = 0	hv (x) = 0	…	h_u (x) = (k-1)
hu (x) = 0	C_{00}	C_{01}	…	$C_{0(k-1)}$
hu (x) = 1	C_{10}	C_{11}	…	$C_{1(k-1)}$
…	…	…	…	…
hu (x) = (k-1)	$C_{(k-1)0}$	$C_{(k-1)1}$	…	$C_{(k-1)(k-1)}$

[7] proposed a solution to enhance machine learning-based IDSs in dynamic and heterogeneous IoT environments. They introduced a drift detection technique based on principal component analysis (PCA) and an online deep neural network (DNN) with adaptive hidden layers. Experimental results on an IoT-based intrusion detection dataset showed improved performance over static models. However, the proposed solution may require additional computational resources, warranting further research on scalability and real-world applicability.

This research grant proposal seeks to address the following research question:

Research Question: *What is the efficiency of a concept drift detector in detecting model drift in a multiclass problem scenario for IoT Intrusion Detection Systems*? The task is to detect concept drifts in a shorter time and with lower memory consumption while ensuring the accuracy of the data stream models remains constant. The goal is to indicate model drift with reduced time and memory usage while maintaining the accuracy of the IDS model.

Research Objective: The aim is to develop a novel approach that enables the detection of model drifts in multi-class problems with improved accuracy and reduced time and memory consumption compared to other Intrusion Detection Systems (IDSs). This objective will be achieved by conducting empirical comparisons between our proposed IDS and existing ones, using synthetic and real-world data streams. The evaluation will consider various performance measures, including detection delay, true detection, memory usage, and accuracy.

Furthermore, incorporating machine learning techniques, statistical methods, progressive learning algorithms, and drift detection techniques into the development of Intrusion Detection Systems (IDSs) aligns with the objective of utilizing concept and data drift to enhance IDS performance in multiclass problem scenarios within IoT networks. Specifically, adopting progressive learning strategies and continuously updating training data in response to concept drift can improve the IDSs' ability to adapt to network changes and detect new and unfamiliar attacks. Additionally, investigating the potential of hybrid solutions that combine sliding window methods with deep learning-based classifiers for real-time intrusion detection in high-speed network scenarios can enhance the system's capacity to handle multiclass problems. To evaluate the effectiveness of these proposed approaches, genuine IoT network traffic datasets can be employed to assess their efficiency in detecting attacks in dynamic and diverse IoT environments.

IV. Discussions

This paper presents a new approach to tackle the issue of model drift in IoT Intrusion Detection Systems within a multi-class problem environment. We propose the Diversity Measure as a Drift Detection Method for multi-class problems, offering a potential solution for identifying and adapting to changes in the statistical properties of data distribution. By integrating drift detection algorithms into IDS, our method aims to improve the accuracy of anomaly detection in IoT devices over time, reducing both false positives and false negatives. This ultimately leads to enhanced system performance and security in a multi-class environment. Our approach combines the disagreement measure from static learning with the Page-Hinkley test to identify drifts in streaming scenarios. It is designed to be computationally efficient and requires minimal memory usage. We argue that our proposed approach overcomes the limitations of existing drift detection techniques that are either computationally expensive or lack the necessary speed for rapid drift detection. Furthermore, our method has the potential to enhance the adaptability and accuracy of IDS in the dynamic, evolving contexts of IoT networks. The utilization of the Diversity Measure as a drift detection technique can also be extended to other domains and data-driven applications that are susceptible to concept drift.

V. Conclusion

In conclusion, this paper presents a novel approach to address concept drift in IoT Intrusion Detection Systems within a multi-class problem scenario, utilizing drift detection algorithms. The proposed method has the potential to improve the accuracy of anomaly detection in IoT devices over time, reducing false positives and false negatives. By employing the Diversity Measure as a drift detection technique, the IDS can effectively adapt to the dynamic nature of IoT environments, making it more resilient against new and unfamiliar attacks.

Furthermore, the proposed approach can be extended to other data-driven applications that are susceptible to concept drift, such as fraud detection and predictive maintenance. This contributes to the development of more robust and reliable systems. Additionally, providing an open-source implementation of the proposed IDS encourages further research and development in IoT security, promoting collaboration and knowledge-sharing among researchers and practitioners. Ultimately, this research significantly contributes to the IoT security domain by introducing a novel intrusion detection system capable of detecting unknown attacks in IoT networks, thereby enhancing the security of IoT devices, systems, and applications.

Acknowledgment

The authors acknowledge that this research study is ongoing at the Melbourne Institute of Technology (MIT). The authors further acknowledge that this study is funded by an MIT research grant and the grant from the Melbourne Institute of Technology (MIT).

References

[1] A. Alazab, A. Khraisat and S. Singh, "A Review on the Internet of Things (IoT) Forensics: Challenges, Techniques, and Evaluation of Digital Forensic Tools," in *Digital Forensics - Challenges and New Frontiers*, Rijeka , IntechOpen, 2023.

[2] A. Khraisat and A. Alazab, "A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges," *Cybersecurity* , vol. 4, no. 1, pp. 1-27, 2021.

[3] A. Alazab, A. Khraisat, S. Singh, S. Bevinakoppa and O. A. Mahdi, "Routing Attacks Detection in 6LoWPAN-Based Internet," *Electronics*, vol. 12, no. 6, p. 1320, 2023.

[4] O. A. Mahdi, A. Alazab, S. Bevinakoppa, N. Ali and A. Khraisat, "Enhancing IoT Intrusion Detection System Performance with the Diversity Measure as a Novel Drift Detection Method," in *2023 9th International Conference on Information Technology Trends (ITT)*, Dubai, 2023.

[5] A. Alazab, A. Khraisat, M. Alazab and S. Singh, "Detection of Obfuscated Malicious JavaScript Code," *Future Internet*, vol. 14, no. 8, p. 217, 2022.

[6] O. A. Mahdi, E. Pardede, N. Ali and J. Cao, "Fast reaction to sudden concept drift in the absence of class labels," *Applied Sciences*, vol. 10, no. 2, p. 606, 2020.

[7] O. A. Wahab, "Intrusion detection in the iot under data and concept drifts: Online deep learning approach," *IEEE Internet of Things Journal*, vol. 9, no. 20, pp. 19706-19716., 2022.

[8] L. I. Kuncheva, Combining pattern classifiers: methods and algorithms, John Wiley & Sons, 2014.

[9] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 8, pp. 832-844, 1998.

[10] D. D. Margineantu and G. D. Thomas, "Pruning adaptive boosting," *ICML*, vol. 97, pp. 211-218, 1997.

[11] L. L. Minku, A. P. White and . X. Yao, "The impact of diversity on online ensemble learning in the presence of concept drift," *IEEE Transactions on knowledge and Data Engineering*, vol. 22, no. 5, pp. 730-742, 2009.

[12] D. B. Skalak, "The sources of increased accuracy for two proposed boosting algorithms," in *In Proc. American Association for Artificial Intelligence, AAAI-96, Integrating Multiple Learned Models Workshop*, 1996.

[13] J. Gama, R. Sebastiao and P. P. Rodrigues , "On evaluating stream learning algorithms," *Machine learning*, vol. 90, pp. 317-346, 2013.

[14] O. A. Mahdi, E. Pardede and N. Ali , "A hybrid block-based ensemble framework for the multi-class problem to react to different types of drifts," *Cluster Computing*, vol. 24, pp. 2327-2340, 2021.

[15] O. A. Mahdi, "Diversity Measures as New Concept Drift Detection Methods in Data Stream Mining (Doctoral dissertation)," La Trobe, Australia, Melbourne, 2021.

[16] O. A. Mahd, E. Pardede, N. Ali and J. Cao, "Diversity measure as a new drift detection method in data streaming," *Knowledge-Based Systems,* vol. 191, pp. 105-227, 2020.

[17] O. A. Mahdi, E. Pardede and N. Al, "KAPPA as drift detector in data stream mining," in *Procedia Computer Science*, Warsaw, Poland, 2021.

[18] J. Gama, P. Medas, G. Castillo and P. Rodrigues, "Learning with drift detection," in *17th Brazilian Symposium on Artificial Intelligence*, Sao Luis, Maranhao, Brazil,, 2004.

[19] M. Baena-Garc´ıa, J. d. Campo-Avila, R. Fidalgo, A. Bifet, R. Gavalda and R. Morales-Bueno, "Early drift detection method," in *In Fourth international workshop on knowledge discovery from data streams*, 2006.

[20] O. A. Mahdi, E. Pardede and J. Cao, "Combination of information entropy and ensemble classification for detecting concept drift in data stream," in *In Proceedings of the Australasian Computer Science Week Multiconference*, Australia, 2018.

[21] A. Bifet and R. Gavalda, "Learning from time-changing data with adaptive windowing," in *In Proceedings of the 2007 SIAM international conference on data mining*, 2007.

[22] I. Frías-Blanco, J. d. Campo-Ávila, G. Ramos-Jiménez, R. Morales-Bueno, A. Ortiz-Díaz and Y. Caballero-Mota, "Online and non-parametric drift detection methods based on Hoeffding's bounds.," *IEEE Transactions on Knowledge and Data Engineering,* vol. 27, no. 3, pp. 810-823., 2014.

[23] G. J. Ross, N. M. Adams, D. K. Tasoulis and D. J. Hand, "Exponentially weighted moving average charts for detecting concept drift.," vol. 33, no. 2, pp. 191-198., 2012.

[24] H. Wang, W. Fan, P. S. Yu and J. Han, "Mining concept-drifting data streams using ensemble classifiers," in *In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2003.

[25] R. Elwell and R. Polikar, "Incremental learning of concept drift in nonstationary environments," *Transactions on Neural Networks,* vol. 22, no. 10, pp. 1517-1531, 2011.

[26] L. Yang and A. Shami, "A lightweight concept drift detection and adaptation framework for IoT data streams," *IEEE Internet of Things Magazine,* vol. 4, no. 2, pp. 96-101, 2021.

[27] N. Martindale, M. Ismail and D. A. Talbert, "Ensemble-based online machine learning algorithms for network intrusion detection systems using streaming data," *Information,* vol. 11, no. 6, p. 315, 2020.

[28] H. Qiao, B. Novikov and J. O. Blech, "Concept Drift Analysis by Dynamic Residual Projection for effectively Detecting Botnet Cyber-attacks in IoT scenarios," *IEEE Transactions on Industrial Informatics,* vol. 18, no. 6, pp. 3692-3701, 2021.