Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Internet of Things Security Based on Incremental XGBoost

Omar Shakir Hasan

Directorate of Education in Nineveh , Mosul, Iraq, omarshakir06@gmail.com

ARTICLE INFO

Article history:
Received: 23 /11/2025
Rrevised form: 06 /12/2025
Accepted : 07 /12/2025
Available online: 30 /12/2025

Keywords:

Incremental learning
boosting,
Anomaly Detection
IoT attack.

ABSTRACT

As the Internet of Things (IoT) has become more common in recent years, security vulnerabilities and attacks on the networks associated with it have also multiplied, increasing the urgency for improved attack detection, deterrence, and response methods. As a countermeasure to IoT attacks, this paper proposes incremental XGBoost-learning and studied the case of creating an incrementally-learning model and how to batch-train the model to leverage the data for full analysis. To validate the proposals, three datasets were chosen, the NSL-KDD dataset, the CICIDS2017 dataset, and the BoT-IoT dataset, in which each of these datasets also included different categories in their definitions of Internet-of-Things attacks. Upon obtaining the dataset, the datasets were then subset into training batches that would continuously assess the model's ability to learn and adapt to new data planning without a retraining process from scratch being necessary. Finally, after all batches were trained, the proposed model attained a classification accuracy of 96.5% on the NSL-KDD dataset, 97.3% on the CICIDS2017 dataset, and 96.8% on the BoT-IoT dataset with improvements in training accuracy while still maintaining consistency across the training subset batches.

MSC..

https://doi.org/ 10.29304/jqcsm.2025.17.42577.

1. Introduction

A fresh concept, referred to as Internet of Things (IoT), is based on the connection of devices, from industrial machines to household items [1]. The inter-connectivity, in the fields of smart cities, healthcare, agriculture, and automotive sectors, facilitates data sharing to increase efficiency, flexibility, and capabilities [2] [3]. Conversely, the growing number of IoT devices indicates the risk of security risks in use [4]. Thus, although it is estimated that by 2025 there will be 50 billion IoT devices on the internet [3] [5], one needs to be aware of the expanding possibility into the threat, and how it makes the way for IoT security a sensitive subject. Because of the limitations of their hardware, Internet of Things (IoT) devices are more vulnerable than other hardware-based computing devices to

∗Corresponding author : Omar Shakir Hasan

Email addresses: omarshakir06@gmail.com

Communicated by ‘sub etitor’

malicious activities. The diverse nature of IoT devices adds additional complexity, as security measures taken by manufacturers differ from each other. Thus, there exists an interoperability issue that creates confusion, confusion that makes implementing security protocols difficult. In an IoT environment, major forms of attacks include distributed denial-of-service (DDoS) attacks, where attackers can cripple devices that may one day be able to control infrastructure; and man-in-the-middle (MitM) attacks, which allow an attacker to assume total command and usage of the device interface, enabling them to modify the actual messages being sent. Also, by assuming physical possession of the device, an attacker may potentially harm themselves or invade the privacy of the user by accessing data on the device. The seriousness of the IoT security risk has been demonstrated by incidents such as the Mirai botnet incident of 2016, in which hackers were able to take control of hundreds of thousands of IoT devices using default log-in credentials.

Security solutions that traditional security solutions can provide are inadequate for today's dynamic, evolving and highly constrained internet of things environment [7], this being one of the primary security challenges in this area [6,7]. It is therefore imperative that we find security solutions that will provide an optimal way to address the different types of attacks that are now being presented by IoT [7]. In order to address the security issue facing IoT today, having an effective security solution will require the use of machine learning as a tool in order to increase IoT security [10,11]. The use of ensemble learning as a different type of ML technique with multiple learning algorithms assists in providing the ability to more accurately predict anomalous behaviour or criminal activity within IoT networks [10, 11]. XGBoost is a good example of how ensemble learning with incremental learning is the most advantageous of the proposed strategies for the IoT security environment. Incremental learning allows for continual updates to be made to the model by introducing new data; thus enabling continual updates of the IoT model throughout a dynamic environment [12]. By eliminating the computational resources and time to train/reset the models for newly developed attack patterns that are discovered using batch learning techniques, this method also has a positive effect in enabling the model to become better at predicting . Continuing improvements to the model over time will be achieved by using this way to address the limitations of current batch learning techniques [12].

This study introduces an incremental XGBoost-based intrusion detection approach designed for dynamic IoT environments. The contribution lies in enabling continuous model updating without full retraining, applying a unified preprocessing strategy across three benchmark datasets (NSL-KDD, CICIDS2017, BoT-IoT), and demonstrating improved accuracy compared to baseline and non-incremental models. These findings highlight the practicality and scalability of the proposed method for real-world IoT security. The paper was organized as follows: Section 2 reviews related work on machine learning methods with emphasis on boosting techniques. Section 3 presents the proposed incremental XGBoost methodology, while Section 4 reports the experimental results and comparisons. Finally, Section 5 concludes the paper and outlines directions for future research.

2. Literature Review

Several authors have addressed the use of reinforcement and ensemble learning—particularly XGBoost—in improving IoT security, with reliable detection rates and the ability to accommodate new threats. In [13], the authors presented an interpretable intrusion detection model that leveraged hybrid sampling to address imbalance, redundancy elimination feature selection Recursive Feature Elimination (RFE), and then an improved version of XGBoost. The goal of this work was to create a transparent detection model in interpretable directions, to determine the contribution of any variable to the overall detection decision. The authors concluded that under the interpretability approach, combined with imbalance and feature processing, XGBoost was the most consistent in comparative work.

In [14], the authors proposed five algorithms designed for interpretable ensemble methods for IoT environments. They proposed an IDS that fused ensemble learning and rule inference to provide explanation of the model, and tested Random Forest, AdaBoost, XGBoost, LightGBM, and CatBoost with several public datasets. The method even included procedures for systematic feature selection, and building an explanation with rules to explain decisions made by the classifiers. The authors concluded that, overall, of the five ensemble classification techniques with explainability, the best performance was done with XGBoost.

In [15] the authors proposed a framework to address two highly correlated issues, imbalance, and feature selection. The authors develop a feature selection process in stages, data generation process, LLM based representations refinement process, and eventually made improvements to their LightGBM classification framework. Their results

were more consistent for the recent data, and that the LightGBM, framework was better than each of the competing methods for testing.

In [16], the authors devise a three-tier framework for processing BoT-IoT data that fuses both deep learning and sequential stage structuring. The framework contains three levels of intervention, which consist of data preprocessing, dimensionality reduction, and at least two layers of training models that lead to a final decision. In their comparisons with the previous work internally, they considered their proposed pipeline as the best quality, as it outperformed the classic alternatives that were also tested under the same conditions. In [17], a study engaged with a similar question and extracted representative features from IoT streams with convolutional neural networks (CNNs), and then passed the extracted vectors to XGBoost for classification. The idea was to separate the deep feature extraction from tree-based decision making. The concluding observations of the authors were that combining CNNs (in the feature phase) and XGBoost (in the classification phase) produced better results than both a classifier based on CNNs, and a tree or forest that lacked the deep representations.

The authors in [18] introduced a model called SAPGAN based on a progressively adversarial generator utilizing a self-attention mechanism to solve scarcity and discrepancy before classification. The framework follows a sequential process that includes missing value compensation, selection of features, and generating new samples before the classifier. After thorough testing, the generative pipeline proved to be better than alternative pipelines that excluded the generative step. Ultimately, it was found to be the most effective option of the proposed framework. In [19], the paper sought to review a stacking ensemble based on a combination of CNNs and Long Short-Term Memory (LSTM) and compared it to other models on industrial (IIoT) datasets. CNNs were utilized to account for spatial features and LSTMs were used to account for temporal relationships. In the environments tested, the CNN-LSTM hybrid performed better than other single models or stacking models that did not utilize both dimensions.

Despite the advances in IoT intrusion detection, most existing methods still rely on static, batch-based learning that cannot adapt to continuously changing IoT traffic. Moreover, prior studies have not evaluated an incremental learning approach across diverse datasets using a unified preprocessing strategy.

3. Methodology

The proposed method is based on the incremental version of the XGBoost algorithm, already taking into account training the model in consecutive batches. The method combines multiple functionalities of both incremental learning and reinforcement algorithms, allowing the model to sequentially learn new data while not forgetting data points used in previous training. To show effectiveness, the study employed three datasets for testing: NSL-KDD, CICIDS2017, and BoT-IoT.

3.1 Dataset

Three datasets were used in this study to evaluate the performance of the incremental XGBoost model: NSL-KDD [20], CICIDS2017 [21], and BoT-IoT [22]. Each dataset is unique in their specific environments and issues each was designed to address in our study.

NSL-KDD is a cleaner version of KDD'99 and is often used in research related to intrusion detection. The NSL-KDD dataset has around 125,973 records that are labeled as normal traffic or an attack, with each record having 41 attributes that capture features of a network connection. The four attack types are Denial of Service (DoS), Probe, User to Root (U2R), and Remote to Local (R2L). As an added benefit, the researcher could present their comparison of intrusion detection system (IDS) models to traditional intrusion detection systems models using a robust benchmark without adding redundant benchmark categories like the original KDD'99.

CICIDS2017 Dataset: The CICIDS2017 dataset from the Canadian Institute for Cybersecurity IDS 2017 is regarded as one of the most realistic and comprehensive modern IDS datasets. CICIDS2017 contains over 3 million records with over 80 features that were extracted from the dataset pertaining to network flows. In addition, there is a mixture of benign and several malicious activities, including DDoS, Botnet, Infiltration, Brute Force (SSH and FTP), Web Attacks (SQL Injection, XSS), and Heartbleed. The quantity and quality of the dataset allows for performing evaluations of IDS models under realism conditions.

BoT-IoT Dataset: BoT-IoT is deliberately designed for use in the IoT environment, and is a large dataset (over 70 million, subsets would be commonly used for training and evaluation). BoT-IoT consists of 5 classes: Normal, DoS/DDoS, Reconnaissance, Information Theft, and other IoT-based attacks. The dataset demonstrates realism when examining IoT traffic since many resource-restricted devices often can be subject to large-scale cyberattacks. Table 1 summarizes the number of features and states for each dataset.

Table 1. Datasets description

Dataset	Number of Records	Number of Features	Attack Types Included
NSL-KDD	125,973	41	DoS, Probe, U2R, R2L
CICIDS2017	3,119,345	81	DDoS, Botnet, Infiltration, Brute Force (SSH/FTP), Web Attacks (SQLi, XSS), Heartbleed
BoT-IoT	72,006,791 total	46)	Normal, DoS/DDoS, Reconnaissance, Information Theft, IoT-specific attacks

3.2 Data Preprocessing

Preprocessing was conducted in a consistent manner for the analysis of all three datasets, NSL-KDD, CICIDS2017, and BoT-IoT; based on correlation analysis and feature importance, features were initially filtered, removing any feature with a correlation of greater than 0.99 to another feature, as well as any feature with zero feature importance. After this reduction step, the final number of features varied depending on the dataset. Subsequently, categorical attributes were transformed into numerical values using one-hot encoding, ensuring consistency of representation across the three datasets before model training.

3.3 Incremental Learning with XGBoost

After selecting and preparing the dataset, the XGBoost model is tested in the incremental learning case. The training data is divided into 50 batches. In each batch, the model is updated using the current batch while retaining information from the previous batches. After training each batch, the model makes predictions on the test data and its accuracy is calculated at each stage. After completing training of all batches, a final evaluation of the model is performed on the test dataset and the overall prediction accuracy is calculated.

The XGBoost model is built on the principle of gradient boosting, where the model is trained as an ensemble of trees that are built sequentially, with each new tree correcting the errors of the previous model [23][24]. In the case of incremental learning, the model is updated gradually with each new batch of data without the need for full retraining [25]. XGBoost is based on maximizing the following objective function (1):

$$\mathrm{Obj}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t), \quad \ldots\ldots\ldots\ldots\ldots(1)$$

where $l(\cdot)$ is the pointwise loss (e.g., logistic loss or MSE), y_i is the ground-truth label for sample i, $\hat{y}_i^{(t-1)}$ is the prediction after $(t-1)$ trees, f_t is the new tree added at stage t, and $\Omega(\cdot)$ is a regularization term controlling model complexity.

Using a second-order (Newton) Taylor expansion around $\hat{y}_i^{(t-1)}$, the objective at stage t is approximated by eq (2)

$$\mathrm{Obj}^{(t)} \approx \sum_{i=1}^{n}\left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right) + \Omega(f_t), \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

where

$$g_i = \frac{\partial l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\partial \hat{y}_i^{(t-1)}}, \qquad h_i = \frac{\partial^2 l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\partial\left(\hat{y}_i^{(-1)}\right)^2} \quad \ldots\ldots\ldots(3)$$

are the first- and second-order gradients, respectively.

A standard regularizer used by XGBoost for a tree f_t with T leaves and leaf weights $\{w_j\}_{j=1}^{T}$ is

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \ldots\ldots\ldots (4)$$

where γ and λ are non-negative hyperparameters.

Under incremental learning, data arrive in batches $\{B_k\}_{k=1}^{K}$. At boosting stage t within batch k, the approximated objective becomes

$$Obj^{(t,k)} \approx \sum_{i \in B_k} \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \Omega(f_t), \ldots\ldots\ldots (5)$$

and the prediction is updated batch-by-batch (and stage-by-stage) as

$$\hat{y}_i^{(t,k)} = \hat{y}_i^{(t,k-1)} + f_t(x_i), \qquad i \in B_k. \ldots\ldots\ldots\ldots (6)$$

Thus, the model retains information learned from previous batches while incrementally refining its predictions on new data.

For a fixed tree structure at stage t, let I_j denote the index set of samples that fall into leaf j. The optimal weight for leaf j is

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \ldots\ldots\ldots (7)$$

Notation. x_i is the feature vector of sample i; $y_i \in \{0,1\}$ for binary classification or $\mathbb{R}$ for regression; $f_t \in \mathcal{F}$ where $\mathcal{F}$ is the space of CART trees; n is the number of samples used at the current stage; K is the number of incoming batches.

4. Experimental Results

The evaluation was conducted using NSL-KDD, CICIDS2017, and BoT-IoT. The results compare incremental XGBoost with its discrete counterpart and four baseline models: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NN). Table 2 shows the accuracy rates achieved by the different models on the three datasets.

Table 2 - Accuracy comparison between selected models on datasets

Model	NSL-KDD	CICIDS2017	BoT-IoT
Incremental XGBoost	96.5	97.3	96.8
Discrete XGBoost	93.2	94.6	92.8
Random Forest (RF)	95.1	96.1	95.4
Neural Network (NN)	92.5	95.3	93.6
Support Vector Machine (SVM)	84.7	88.2	83.9
Decision Tree (DT)	81.4	78.6	80.1

Figure 1 shows the performance results of the different models when applied to the NSL-KDD dataset. Figure 2 represents the results for the CICIDS2017 dataset. Figure 3 shows the results for the BoT-IoT dataset. Figure 4 presents a comprehensive comparison between the models across the three datasets combined.

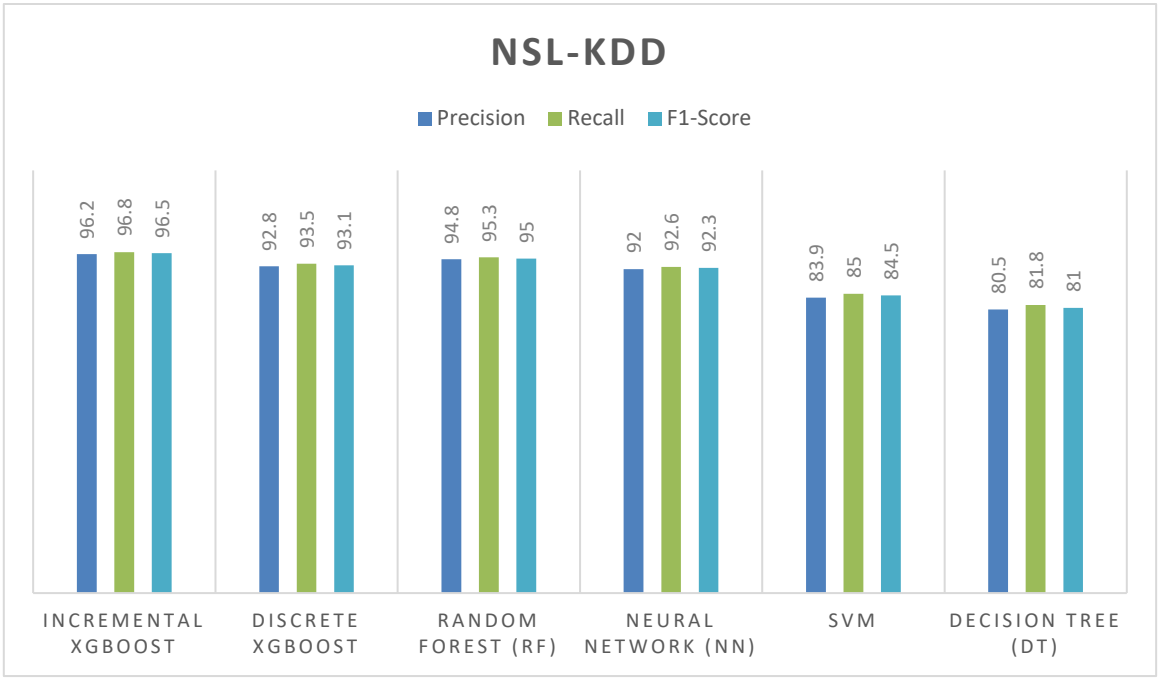

Figure (1) - Model performance on the NSL-KDD dataset

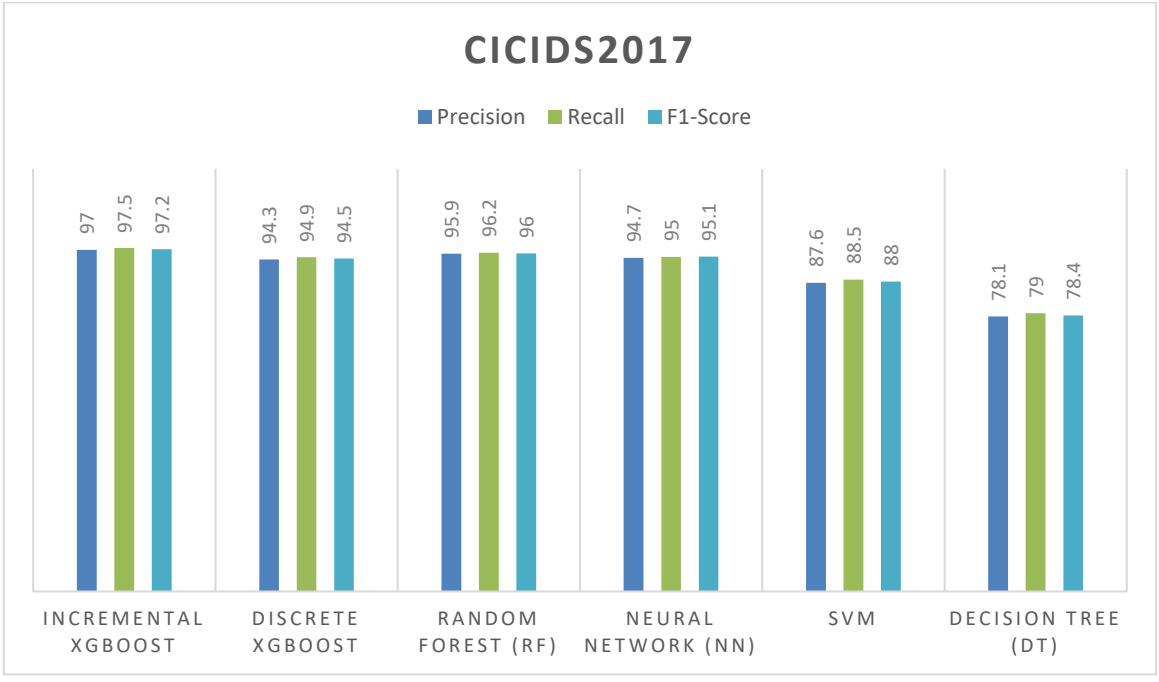

Figure (2):. Model performance on the CICIDS2017 dataset

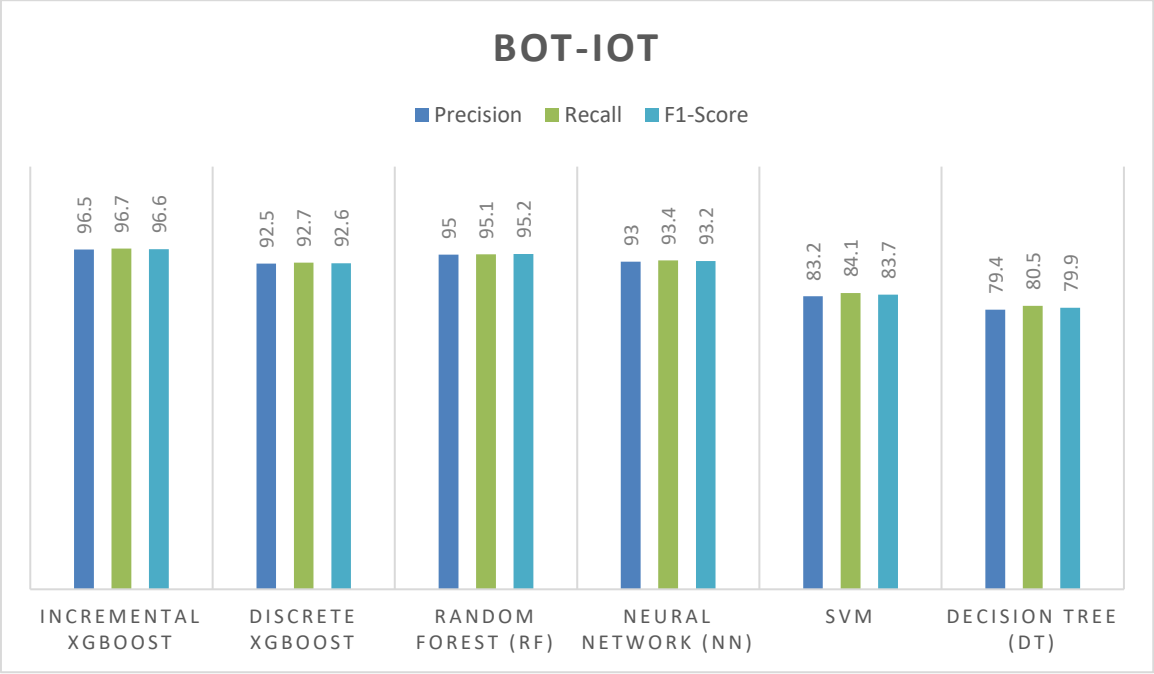

Figure (3):Model performance on the BoT-IoT dataset

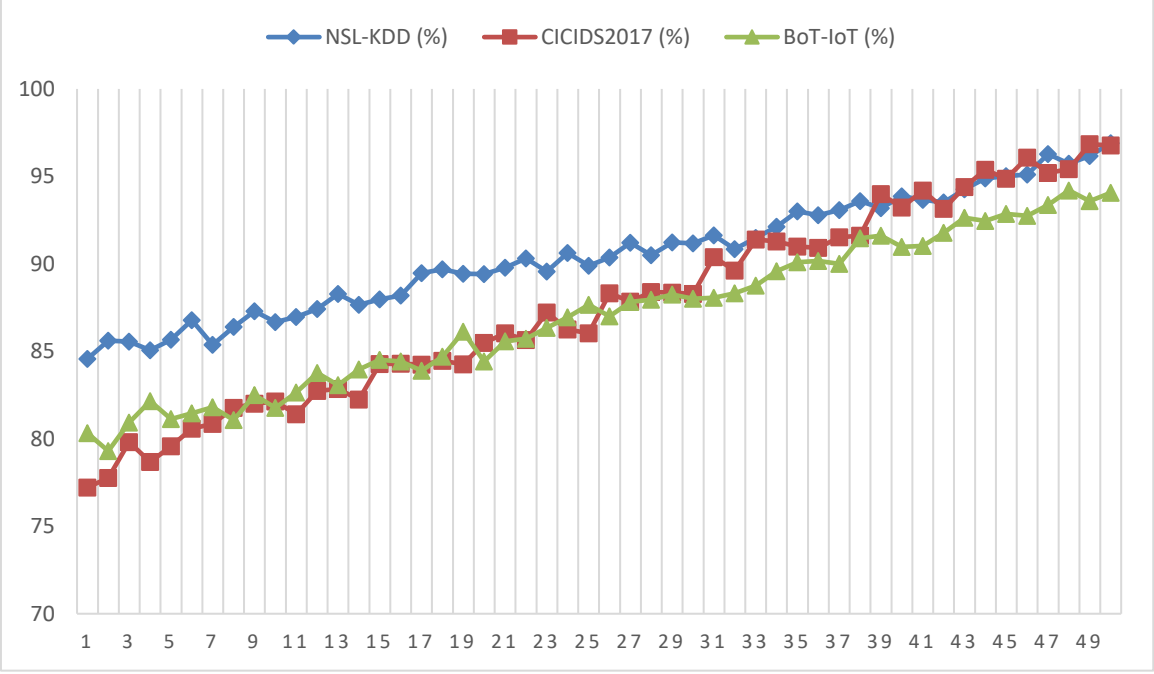

Figure (4): Comprehensive comparison of model performance across the three datasets

The figure shows the progression of accuracy across fifty training batches, with the horizontal axis representing batch numbers from 1 to 50, and the vertical axis representing the percentage accuracy within a range of 70% to 100%. Three main curves are shown in the figure: the first, in blue, represents the NSL-KDD dataset, the second, in red, represents the CICIDS2017 dataset, and the third, in green, represents the BoT-IoT dataset. Each curve starts at a low level in the first batches and gradually increases with the number of batches, stabilizing at levels close to 97% in the final batches.

5. Discussion

The experimental evaluation results for the proposed model using the incremental XGBoost algorithm over three datasets (NSL-KDD, CICIDS2017, and BoT-IoT) showed its ability to obtain high accuracy rates of 96.5%, 97.3%, and 96.8%, respectivelyXGBoost Models were used to demonstrate that they have been proven to be effective at both traditional datasets such as NSL-KDD as well as modern more realistic datasets such as CICIDS2017 and BoT-IoT. When looking at accuracy, the Non-Incremental version of XGBoost ranged from 92.8% - 94.6%. By performing the incremental learning, the model benefitted from keeping all the previous batch knowledge of its training and processing in an incremental manner. Therefore, when looking at the accuracy of the model, we see that the Incremental XGBoost achieved an improvement of 2% - 4% on every dataset tested. While this may seem small in a numerical field, the real-world application of it is significant because the researchers must act quickly and have the most reliable results to effectively detect an IoT-related attack.

The overall performance of Accuracy by Incremental XGBoost is comparatively better by a higher percentage to other methods, such as Random Forest (95.4%), Neural Networks (93.6%), Support Vector Machines (83.9%) and Decision Trees (80.1%) to perform consistently across all dataset categories. For instance, when looking at the CICIDS2017 dataset, Decision Trees dropped below 80% accuracy while Incremental XGBoost maintained a consistent accuracy of over 97%. This could be significant in the real world, as the types of attacks vary depending on the context.

It should also be noted that in each of the comparative graphs (1–3), Incremental XGBoost showed improved results no matter what scenario, while often approaching Random Forest's overall accuracy. The notable difference between Incremental XGBoost to the others is the incremental model's ability to continue to improve the results throughout the training batch cycles, while the others have stopped improving after the initial training iterations. This added benefit is a particularly helpful feature for IoT environments with streaming data that is known to rapidly change while remaining valid data streaming into their denser environments.

Conversely, it should be noted that these results were derived from a testing environment based on standard databases and thus do not apply to all IoT environments in the real-world. While the performance (over 96% in every case) suggests that it could be considered as a real-world practical option, applying this model in the real-world would require further testing on system execution speed and resource efficiency; along with an attack testing framework system on types and attacks not listed in the obtained groups.

6. Conclusion

The proposed model in this paper is based upon the incremental XGBoost Algorithm to enhance the capabilities of intrusion detection systems in the IoT environment. In contrast to standard approaches to learning systems based on full batch training, the proposed iterative model is capable of recognizing changes in the data, incrementally updating the knowledge the model learns from across successive data sets, and reducing the cost of retraining in full, thus providing increased performance. The incremental approach provides an advantageous way to support real-time monitoring of IoT devices, a flexible way of detecting unauthorized access into your network, and in environments where datasets are perpetually changing over time.

The results from our study highlight how incremental learning gives your model the ability to modify itself, as well as maintain a high degree of reliability, even when encountering new types of datasets. While traditional models like Random Forest, Support Vector Machines, Neural Networks, and Decision Trees may be appropriate for detecting intrusions in static datasets, those models are unable to automatically adapt and be updated whenever new attack patterns are identified. In comparison, using the incremental XGBoost approach enables a model to have practical value within a rapidly changing IoT systems that rely upon immediately detecting attacks and responding as quickly as possible, without having to incur the costs associated with re-training the model.

However, it is essential to note that these findings are based upon an experimental design that utilized standardized datasets. Therefore, further validation and refinement should occur before the general applicability of this model can be confirmed. Future research could extend the evaluation of this incremental learning model by combining it with an appropriate deep learning algorithm or employing contemporary optimization strategies, and also expand the evaluation area to include a greater variety of datasets.

Reference

[1] G. Shahinzadeh, H. Shahinzadeh, and S. Tanwar, Security and privacy issues in the Internet of Things: A comprehensive survey of protocols, standards, and the revolutionary role of blockchain, in Proc. 8th Int. Conf. Smart Cities, Internet of Things and Applications (SCIoT). IEEE, (2024).

[2] M. Salb, L. Jovanovic, N. Bacanin, M. Antonijevic, M. Zivkovic, N. Budimirovic, and L. Abualigah, "Enhancing Internet of Things network security using hybrid CNN and XGBoost model tuned via modified reptile search algorithm," Applied Sciences, vol. 13, no. 23, (2023), p. 12687. DOI: 10.3390/app132312687

[3] A. F. Glavan and V. Croitoru, "Incremental learning for edge network intrusion detection," Rev. Roum. Sci. Techn.–Électrotechn. et Énerg., vol. 68, no. 3, (2023), pp. 301–306.

[4] N. U. Prince, M. A. Al Mamun, A. O. Olajide, O. U. Khan, A. B. Akeem, and A. I. Sani, "IEEE standards and deep learning techniques for securing Internet of Things (IoT) devices against cyber attacks," Journal of Computer Analysis and Applications, vol. 33,

[5] P. H. Nguyen, M. Mumtaz, and H. Al-Raweshidy, "Internet of Things (IoT) applications security trends and challenges," Discover Internet of Things, vol. 4, no. 1, (2024), p. 46. DOI: 10.1007/s43926-024-00071-3

[6] S. Gawande, S. Chauhan, and A. V. Deshpande, "A survey on IoT security: Vulnerability detection and protection," Computers & Security, vol. 142, (2024), p. 103393. DOI: 10.1016/j.cose.2023.103393

[7] S. D. Raut and V. M. Thakare, "A survey on privacy and security issues in IoT-based environments," Computers & Security, vol. 136, (2023), p. 103498. DOI: 10.1016/j.cose.2023.103498

[8] J. Li, X. Zhou, and L. Wang, "Intrusion detection method based on active incremental learning," Journal of Internet of Things, vol. 4, no. 2, (2022), pp. 179–190.

[9] A. Amouri, M. M. Al Rahhal, Y. Bazi, I. Butun, and I. Mahgoub, "Enhancing intrusion detection in IoT environments: An advanced ensemble approach using Kolmogorov–Arnold networks," arXiv preprint, arXiv:2408.15886, (2024). Available: https://arxiv.org/abs/2408.15886

[10] T. Lai, F. Farid, A. Bello, and F. Sabrina, "Ensemble learning based anomaly detection for IoT cybersecurity via Bayesian hyperparameters sensitivity analysis," arXiv preprint, arXiv:2307.10596, (2023).

[11] A. Namvar, C. Thapa, and S. S. Kanhere, "Discretization-based ensemble model for robust learning in IoT," arXiv preprint, arXiv:2307.08955, (2023).

[12] S. R. Alve, M. Z. Mahmud, S. Islam, M. A. Chowdhury, and J. Islam, "Smart IoT security: Lightweight machine learning techniques for multi-class attack detection in IoT networks," arXiv preprint, arXiv:2502.04057, (2025).

[13] Y. Hosain and M. Çakmak, "XAI-XGBoost: An innovative explainable intrusion detection approach for securing Internet of Medical Things systems," Scientific Reports, vol. 15, (2025), Art. no. 22278. DOI: 10.1038/s41598-025-07790-0

[14] K. S. Adewole, A. Jacobsson, and P. Davidsson, "Intrusion detection framework for Internet of Things with rule induction for model explanation," Sensors, vol. 25, no. 6, (2025), p. 1845. DOI: 10.3390/s25061845

[15] H. Ma et al., "An IoT intrusion detection framework based on feature selection and large language models fine-tuning," Scientific Reports, (2025).

[16] S. Alosaimi and S. M. Almutairi, "An intrusion detection system using BoT-IoT," Applied Sciences, vol. 13, no. 9, (2023), p. 5427. DOI: 10.3390/app13095427

[17] F. H. Zawaideh, G. Al-Asad, G. Swaneh, S. Batainah, and H. Bakkar, "Intrusion detection system for IoT networks using convolutional neural network (CNN) and XGBoost algorithm," Journal of Theoretical and Applied Information Technology, vol. 102, no. 4, (2024), pp. 1749–1762.

[18] V. Kantharaju et al., "Machine learning-based intrusion detection framework for IoT networks," Scientific Reports, (2024).

[19] P. R. Chithra Rani et al., "Deep learning-based ensemble stacking for enhanced intrusion detection in IIoT networks," SN Applied Sciences, vol. 7, (2025), Art. no. 1477. DOI: 10.1007/s42452-025-01477-0

[20] S. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in Proc. IEEE Symp. Computational Intelligence for Security and Defence Applications (CISDA), Ottawa, Canada, (2009), pp. 1–6.

[21] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in Proc. 4th Int. Conf. Information Systems Security and Privacy (ICISSP), Funchal, Portugal, (2018), pp. 108–116.

[22] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," Future Generation Computer Systems, vol. 100, (2019), pp. 779–796. DOI: 10.1016/j.future.2019.05.024

[23] M. Sun, L. Li, and Z. Li, "Research on improved XGBoost algorithm based on gradient boosting framework," IEEE Access, vol. 9, (2021), pp. 136318–136329. DOI: 10.1109/ACCESS.2021.3108836

Chen, T., and Guestrin, C., "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining (KDD), (2016), pp. 785–794. DOI: 10.1145/2939672.2939785

[24] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining, (2016), pp. 785–794. DOI: 10.1145/2939672.2939785

[25] S. Zhang, J.-w. Liu, and X. Zuo, "Adaptive online incremental learning for evolving data streams," Applied Soft Computing, vol. 105, (2021), Art. no. 107255. DOI: 10.1016/j.asoc.2021.107255

.