Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

Designing Effective Intrusion Detection Models for IoT enabled Smart Homes

Muzammil Hassan, Dr Nmarata Bansal, Ms Amanjyoti sethi
Department of ECE, Lingayas Vidyapeeth, Faridabad, India

***Abstract*— This paper presents a comprehensive analysis on the design of an efficient and time-effective intrusion detection system (IDS) for IoT-capable smart home environments using machine learning algorithms. The study categorizes and predicts various network attacks by analyzing network traffic data and real-time sensor inputs from smart IoT environments. Utilizing the open-source, imbalanced 'DS2OS' dataset, the performance of classifiers such as Logistic Regression, Extreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM) and Random Forest is rigorously evaluated. A novel IDS model, termed "LGB-IDS," is introduced, employing the LGBM library utilizing ensemble techniques and a majority voting approach. Key performance indicators, including accuracy, error rate, false-negative rate (FNR), time efficiency, and true-positive rate (TPR), are used to assess the model's efficacy. The primary objective is to propose an IDS design that achieves superior accuracy, enhanced computational efficiency, and minimized false alarm rates in real-world IoT environments.**

Keywords*— *Intrusion Detection System (IDS), Machine Learning, Internet of Things (IoT), Light Gradient Boosting Machine (LGBM), Anomaly Detection.

I. INTRODUCTION

The swift expansion of the Internet of Things (IoT) has transformed smart home environments into increasingly interconnected systems, thereby rendering them prime targets for a spectrum of malicious activities. This expansion has, in turn, led to a surge in cybercrime incidents that ensuring the protection of IoT user data and privacy while also undermining the integrity of the underlying networks [1]. As IoT systems continue to evolve and scale, the imperative to develop more efficient and robust intrusion detection systems (IDS) becomes ever more critical. Given the growing complexity, autonomy, and expansive nature of IoT networks, novel and innovative solutions are required to detect and mitigate cyber threats effectively.

In securing IoT systems, one of the most formidable challenges is the early recognition and timely response to potential security breaches. Owing to the decentralized operation of IoT devices and their diverse interconnected components, early detection of malicious activity is paramount to forestalling breaches. Many IoT devices do not manifest overt signs of compromise, which complicates simultaneous detection of potential attacks [2]. Therefore, developing an effective IDS capable of continuous monitoring of network traffic and system behavior is essential. A well-designed IDS can promptly issue alerts upon detecting anomalous activities, thereby preventing security incidents before they escalate.

Traditional intrusion detection methodologies often falter when confronted with novel or unknown attack types, as their detection capabilities are typically calibrated towards known threat patterns. Due to the substantial increase in the number of complexities connected IoT devices, attacks are becoming more sophisticated and less predictable. This paradigm shift has spurred the demand for advanced IDS frameworks that can accurately identify both known and emergent threats in real time. By continuously monitoring and analyzing network traffic, such systems aim to preemptively identify and neutralize potential cyber threats, thereby fortifying IoT networks against compromise [3].

In today's interconnected digital landscape—where the expansion of IoT has significantly broadened the cyber-attack surface—IDS play an indispensable role in safeguarding network integrity. The persistent evolution of cyber threats has rendered conventional security mechanisms increasingly inadequate. Consequently, there is a growing demand for intelligent and adaptive IDS capable of real-time detection of diverse attack modalities. Machine learning (ML has emerged as a leading approach in addressing owing to its capacity to manage voluminous datasets, adapt to new attack patterns enhancing the overall accuracy of IDS implementations [5].

II. RELATED WORK

An extensive body of research has demonstrated the effectiveness of ML models in managing large, complex datasets for intrusion detection applications. Various methodologies, including Classification and Regression Trees (CART), Support Vector Machines (SVM), and Random Forests (RF), have been investigated to enhance IDS performance. These models are typically evaluated based on their detection accuracy, computational efficiency, and resource utilization. For instance, Htwe et al. (2020) [6] applied the CART algorithm to the 'N-BaIoT' dataset,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

achieving superior accuracy compared to Naïve Bayes classifiers, though their evaluation was limited to accuracy metrics alone. Similarly, Zhou et al. (2020) [7] introduced an IDS framework that integrated feature-selection and ensemble learning, employing the CFS-BA heuristic for dimensionality reduction. Despite promising results on datasets such as NSL-KDD and AWID, time efficiency was not a focal point of their investigation.

Verma et al. (2020) [8] conducted an exhaustive comparative study on ML classifiers for anomaly-based IDS, particularly focusing on Their analysis of both single and ensemble classifiers in IoT systems addresses Denial of Service (DoS) attacks —including CART, MLP, RF, and Gradient Boosted Machines—highlighted the superior detection capabilities of certain models, albeit with limited consideration of execution time. In a related study, Hadem et al. (2021) [9] proposed an SVM-based IDS for software-defined networking, achieving impressive detection performance on the NSL-KDD dataset. Despite their efforts, the study did not adequately address the critical issue of time efficiency.

Despite the strides made in IDS design, achieving an equilibrium between the precision of detection and the efficiency of computation remains a significant challenge. Kumar et al. (2021) [10] introduced a cyber-attack detection system using algorithms such as RF, KNN, and XGBoost, achieving high detection rates but neglecting the influence of time efficiency on decision-making. Similarly, Hục et al. (2021) [11] P\proposed an edge device anomaly detection model, assessed with the DS2OS dataset, yet without sufficient consideration of computational complexity and runtime performance.

Table 1: DS20S dataset fundamental details

Dataset Name	Main Simulation Access Traces (DS2OS network traffic)
Total Features	13
Total No. of Instances	357953
Selected Features	12
Total No. of Classes	7 (attack classes) + 1 (begning class)
Predictes Attacks	DoSattack, dataProbing, maliciousControl, maliciousOperation, Scan, spying, wrongSetUP

In 2022, several studies attempted to mitigate issues related to class imbalance and dimensionality within IDS. Devprasad et al. [12] utilized hierarchy-based chi-square and bat algorithms to develop a context-adaptive classification mechanism; however, their emphasis on accuracy overshadowed the importance of runtime efficiency. Gupta et al. (2022) [13] created a network-based IDS utilizing ensemble algorithm and deep learning to resolve class imbalance issues, yet did not examine execution time. Çetin [14] (2022) also proposed a model targeting imbalanced network attack traces, however, it did not address time efficiency.

In environments where real-time detection is paramount, the development of time-efficient IDS is critical. Xu and Fan (2022) [15] employed logarithmic auto-encoders and XGB to achieve high detection accuracy but did not adequately address computational overhead. Similarly, Saheed et al. (2022) [16] and Le et al. (2022) [17] reported high accuracy using boosting algorithms and multiclass classification approaches, respectively, yet both studies lacked a focus on runtime efficiency. The present study endeavors to bridge this gap by highlighting the dual focus on detecting accurately as well as expediting time in the design of Intrusion Detection Systems (IDS) for the Internet of Things (IoT)-enabled smart homes.

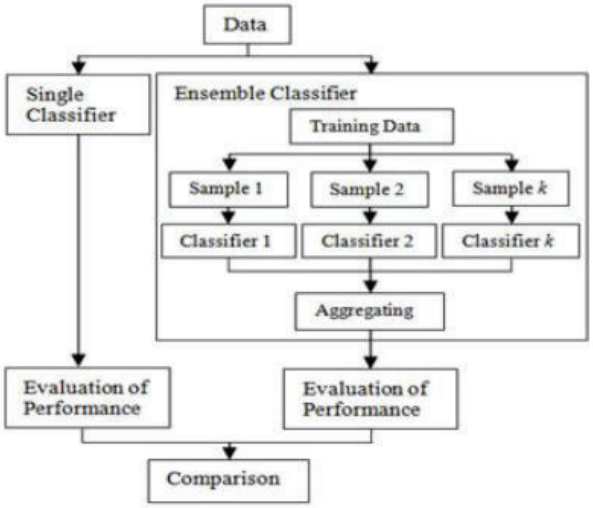

Figure 1: Procedure for single and ensemble classifier models

III. ALGORITHM CLASSIFICATION

In the study, we examine decision-tree-based machine learning techniques—both single and ensemble classifiers—for the prediction of IoT security breaches. Ensemble methods blend multiple models to optimize prediction results, including LGBM Gradient Boosting, Random Forest, and Extreme Gradient Boosting. In contrast, single classifiers such as Logistic Regression (LR) rely on a single model for decision-making. The paper elucidates how sub-classifiers, by aggregating decisions from random samples, compare in performance, drawing methodological inspiration from the systematic analysis conducted by Utami et al [18].

A. SINGLE CLASSIFIER

This section delineates the architecture of a single classifier that utilizes a unified feature set to drive decision-making processes. The model parameters are estimated through linear optimization techniques, and the classifier's performance is rigorously assessed using established statistical methods. This approach offers a baseline for understanding the relative benefits of more complex ensemble strategies.

1) LOGISTIC REGRESSION

Logistic Regression, known as the sigmoid or logit function, is a supervised classification algorithm of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

machine learning primarily employed for predicting categorical or discrete outcomes. Originally developed as a statistical tool for modelling population dynamics, binary logistic regression addresses dichotomous classification problems, while multinomial logistic regression extends the methodology to accommodate multi-class scenarios [19]. The LR model estimates probabilities using the formula:

$$p\,(y = C|X;\, W, b) = \frac{1}{1 + \exp(-W^{inverse}X - b)} \quad (1)$$

Here, X = {X_1, X_2, X_3, ..., X_n} are features, W = {W_1, W_2, W_3, ..., W_n} are weights, b is the bias term, and C represents classes. In this context, LR is applied for multi-class classification using the `multiclass` hyperparameter set to 'multinomial. where X = {X_1, X_2, …, X_n } represents the feature set, W = {W_1, W_2, …, W_n } denotes the corresponding weights, b stands for bias term, and C signifies class labels. In our application, LR is employed for multi-class classification by configuring the multiclass hyperparameter to 'multinomial'.

B. ENSEMBLE CLASSIFIER

Ensemble classifiers offer substantial advantages over single-model approaches by merging predictions from several decision trees to achieve higher accuracy. Bagging techniques are utilized to reduce variance by training models in parallel, whereas boosting methods sequentially minimize bias by iteratively correcting errors [20]. The proposed IDS model capitalizes on these ensemble algorithms to classify network traffic patterns in IoT home environments. By integrating results from a range of classification techniques, validated on 'DS2OS' dataset, system effectively categorizes network traffic based on behavioural patterns.

1) RANDOM FOREST

An ensemble method called Random Forest generates several decision trees from diverse subsets of the dataset to boost predictive accuracy [21]. The algorithm aggregates the predictions of these trees, thereby reducing overfitting and increasing reliability. Although RF exhibits robust performance on large, high-dimensional datasets, its complexity and high computational cost can be significant drawbacks, particularly when interpretability is required. RF classifies traffic as 'normal' or 'anomalous' using a majority voting mechanism derived from the ensemble of trees.

2) GRADIENT BOOSTING

Gradient Boosting, additionally known as Gradient Boosting Decision Tree (GBDT), is a commonly employed decision tree classifier that consolidates multiple weak learners into a cohesive and strong predictive model [22]. By employing a greedy algorithm analogous to gradient descent, GB iteratively refines its predictions based on residual errors from previous iterations. This method enhances feature selection and improves both detection rates and execution speed [23]. However, it has some drawbacks like high-power consumption, extended training times, and potential overfitting necessitate careful regularization of hyperparameters. The additive approximation at each iteration is given by the equation:

$$F^*(X) = F_{l-1}(X) + \rho_l h_l(X) \quad (2)$$

Here, ρl is the weight of lth function hl(X).

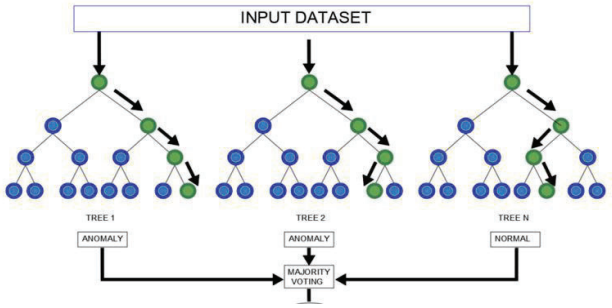

Figure 2: Majority voting classifier

3) EXTREME GRADIENT BOOSTING

Extreme Gradient Boosting (XGB) offers a refined approach to gradient boosting, focusing on maximizing accuracy and reducing false alarms [24]. XGB constructs more robust classification models by balancing the bias-variance trade-off, thereby effectively managing large datasets while mitigating overfitting through rigorous hyperparameter tuning [25]. Despite its accelerated training speed, XGB demands significant computational resources with respect to processing time and memory, and its interpretability remains a challenging aspect. The model is calculated using the formula:

$$F(X, W) = \sum_{l=0}^{L} \alpha_l h_l(X, w_l) = \sum_{l=0}^{L} f_l(X, W_l) \quad (3)$$

where X is the input set, W represents the weights of the respective inputs, and F (X, W) represent the intended model. hl is for single tree, and al is for weight for the L trees. The loss function is minimized in order to optimize the model.

4) LIGHT GRADIENT BOOSTING MACHINE (LGBM)

Light Gradient Boosting Machine, histogram-based decision tree algorithm, has been shown to enhance model efficiency, reduce time of execution, and minimize use of memory. LGBM employs two primary techniques: Gradient-based One-Side Sampling (GOSS), which prioritizes samples with large gradients to improve accuracy, and Exclusive Feature Bundling (EFB), which mitigates the limitations of traditional histogram-based methods. The leaf-wise tree growth of algorithm is depicted in Figure 3, contrasts with conventional level-wise growth by reducing errors and improving performance in classification, regression, and decision-making tasks.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

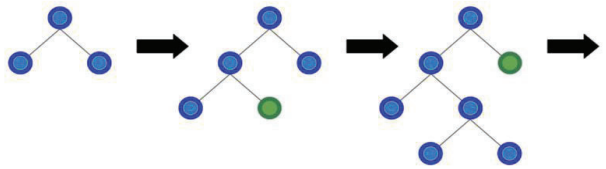

Figure 3: Leaf-oriented tree expansion

5) LGBM

LGBM employs variance gain for feature splitting, in conjunction with GOSS and EFB techniques, to offer several advantages. These include time efficiency, superior performance in comparison to alternative classification algorithms, high training and test accuracy, support for both classification and regression tasks, and the mitigation of overfitting by adjusting the "max_depth" hyperparameter. GOSS enhances gradient boosting by prioritizing samples with larger gradients, thereby accelerating learning and reducing computational complexity. By excluding samples with smaller gradients, it achieves better estimations with smaller data sizes. Given a training dataset with m*m* samples X={X1,X2,X3,…,Xm}$X=\{X1 \ ,X2 \ ,X3 \ ,\ldots,Xm \ \}$, the gradients {g1,g2,g3,…,gm}$\{g1 \ ,g2 \ ,g3 \ ,\ldots,gm \ \}$ are sorted in descending order. The subset T comprises the top (x×100)%(x×100)% samples with the largest gradients, while the subset U includes the remaining samples with smaller gradients. The sample split is based on the estimated variance gain V·J(d)$V{\cdot}J(d)$, calculated using Equation 4.

$$Vector\ V_j * (d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i \in T_a} g_i + \frac{1-t}{u}\sum_{x_i \in U_a} g_i\right)^2}{n_a^j(d)} + \frac{\left(\sum_{x_i \in T_b} g_i + \frac{1-t}{u}\sum_{x_i \in U_b} g_i\right)^2}{n_b^j(d)}\right) \quad (4)$$

EFB bundles sparse, mutually exclusive features to reduce data dimensions without losing accuracy, enhancing training speed. This mechanism, combined with LGBM, simplifies the data processing by reducing histogram complexity and selecting optimal split points, balancing information gain and variance reduction.

EFB is a method used to bundle sparse, mutually exclusive features in high-dimensional data, which is typically sparse. These features are bundled together, reducing dimensionality without significant loss of information. The process changes the complexity of the histogram from O(#data * #features) to O(#data * #bundles), speeding up training without affecting accuracy. Decision trees typically handle discrete data, but when working with continuous data, it needs to be discretized. The challenge lies in finding the optimal split points for features, balancing information gain and overfitting. The LGBM algorithm addresses this by iterating through data points to find the optimal split, It achieves higher information gain with lower variance by partitioning the data into equal-length segments.

IV. METHODOLOGY

In the following sections, the design, analysis, and implementation aspects of the proposed intrusion detection model will be delineated. The process commences with the establishment of the requisite project dependencies and the installation of the necessary libraries. The research methodology is organized into the following key phases:

1. Model Design

The proposed intrusion detection model is designed to distinguish between normal and malicious traffic of network in IoT-enabled environments. This objective is pursued by means of an analysis of data in real time from IoT sensors, devices, and networks. The dataset utilized for this research is publicly accessible on Kaggle[4].The purpsoe of the study is to develop an intrusion detection system able to recognize cyber threats in smart environments. The development of this model is structured in three primary stages:

Preprocessing and Data Preparation – Cleaning, transforming, and structuring the dataset.

Classifier Training – Training the model using techniques of machine learning.

Decision Making – Evaluating the performance of moel and making predictions.

For implementation, the DS2OS dataset has been utilized. This dataset consists of sensor-generated traces from various smart devices configured within a smart environment home. A brief overview of the model's building process is illustrated in figure 4.

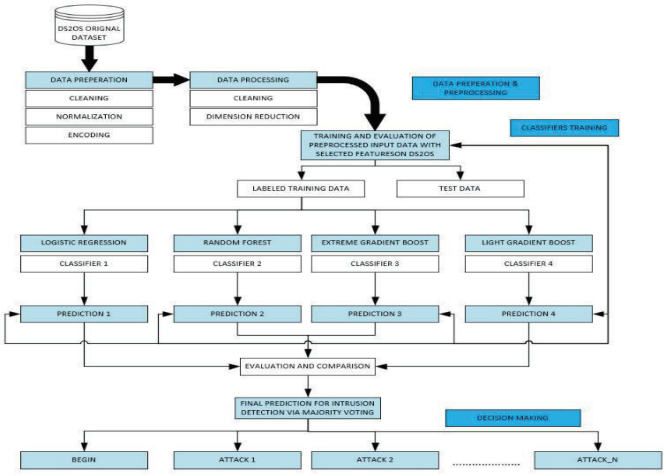

Figure 4: Intrusion detection model- LGB-IDS (Ensemble based)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

2. Model Implementation and Training

Subsequent to the processes of preprocessing and feature extraction, the dataset is segmented into training and testing sets to assess the model's precision and efficacy. Prior to the training phase, the independent variables (X) and the target variable (y) are initialized.The (LGBM) classifier is employed for classification purposes, utilizing the LGBMClassifier method, which operates as an ensemble learning approach.

V. RESULTS

The efficacy of the developed model is evaluated using the DS2OS dataset. In order to ascertain the robustness of the model, statistical classification operations are performed on various portions of the dataset. Initially, a 80:20 ratio is employed to divide the dataset into a training set and a test set. In a subsequent experiment, this ratio is modified to 70:30, although this does not lead to a significant change in the results.The model's effectiveness in detecting intrusions within IoT networks is gauged by its performance metrics, namely accuracy and error scores.

VI. OVERALL PERFORMANCE OF PROPOSED IDS

The Intrusion Detection System (IDS) model was validated using the "DS2OS" dataset, which comprises 357,950 data points, with 71,590 data points designated for training and testing. The model demonstrated an average accuracy that exceeded 99%, thereby substantiating its efficacy in detecting network intrusions. A detailed examination of class-wise performance metrics, including precision, recall, and F1-score, further substantiates the model's efficacy. Specifically, the DoS attack class exhibited a precision of 99%, a recall of 65%, and an F1-score of 77%. Notably, all other classes, including the normal class, attained 100% across these metrics.

Table 2: Detection Rate in Percent

	Dos Attack	Data Probe	Malicius Control	Malicius Opeation	Scan	Spying	Wrong Set Up	Normal
Logistic Regression								
P%	99	100	99	99	99	100	100	100
D%	67	89	100	99	98	100	100	99
Random Forest								
P%	99	99	100	99	99	100	100	100
D%	90	98	100	100	98	100	100	99
Extreme Gradient Boosting								
P%	100	100	100	99	99	100	100	100
D%	98	100	99	97	98	100	100	99
Light Gradient Boosting machine								
P%	100	100	100	99	99	100	100	100
D%	97	100	100	100	98	100	100	99

Table 3: Proposed IDS - Overall Performance

	Precision	Recall	F1-Score	Support
Dos Attack	99	65	79	1110
Data Probe	100	100	100	60
Malicius	100	100	100	151
Malicius	100	100	100	170
Scan	100	100	100	312
Spying	100	100	100	107
Up	100	100	100	31
Normal	99	100	100	69651
Accuracy			99	71590
Average			99	71590

The confusion matrix (Figure 5) further elucidates the performance of model, highlighting both prevalence of errors and the classification accuracy. Comparative analysis reveals that the ensemble-based IDS model outperforms alternative models with respect to accuracy, time efficiency and true-positive rate (TPR). Specifically, the LGB-IDS model, which leverages the Light Gradient Boosting Machine (LGBM) ensemble technique, exhibits high prediction rates coupled with low latency. This significantly enhances both security and operational efficiency in IoT environments, making the proposed IDS model a robust solution for intrusion detection in smart home settings operating in real time.

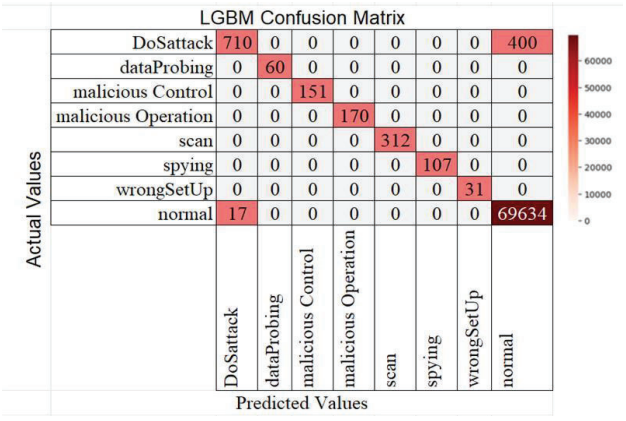

Figure 5: LGBM Confusion matrix

VII. CONCLUSION

This study puts forward a time-efficient and effective intrusion detection system tailored for IoT-enabled environments, the efficacy of which is evaluated using the smart home 'DS2OS' dataset. The proposed model, grounded in ensemble machine learning algorithms and incorporating feature selection techniques, demonstrably reduces prediction latency, computational time complexity, and memory usage. Evaluation metrics such as accuracy, TPR, FPR, and error rates indicate that while no single algorithm is universally superior, the XGB and LGBM models exhibit superior performance. Notably, the LGBM model delivers lower threat prediction latency while maintaining high detection accuracy and reducing overfitting. Consequently,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 05, IIRA 5.0 (2026)

the proposed IDS model is applicable to diverse IoT environments and real-time scenarios, offering significant promise in mitigating cybercrimes and countering security threats in both governmental and private sectors.

REFERENCES

[1] H.F. Atlam, A. Alenezi, M.O. Alassafi, A.A. Alshdadi, and G.B. Wills, "Security, cybercrime and digital forensics for IoT," in *Principles of Internet of Things (IoT) Ecosystem: Insight Paradigm*, vol. 174, Jan. 2020, pp. 551–577.

[2] M. Alkasassbeh and S. A.-H. Baddar, "Intrusion detection systems: A state-of-the-art taxonomy and survey," *Arabian J. Sci. Eng.*, vol. 2022, pp. 1–44, Nov. 2022.

[3] Z. Sun, Z. Lv, H. Wang, Z. Li, F. Jia, and C. Lai, "Sensing cloud computing in Internet of Things: A novel data scheduling optimization algorithm," *IEEE Access*, vol. 8, pp. 42141–42153, 2020.

[4] (2018). *DS2OS Traffic Traces, IoT Traffic Traces Gathered in the DS2OS IoT Environment*. Accessed: Jul. 28, 2022. [Online]. Available: https://www.kaggle.com/francoisxa/ds2ostraffictraces

[5] D. Rani, N. S. Gill, and P. Gulia, "Classification of security issues and cyber attacks in layered Internet of Things," *J. Theor. Appl. Inf. Technol.*, vol. 100, no. 13, pp. 4895–4913, Jul. 2022.

[6] C. S. H. Twe, Y. M. Thant, and M. M. S. Thwin, "Botnets attack detection using machine learning approach for IoT environment," *J. Phys., Conf. Ser.*, vol. 1646, no. 1, pp. 1–8, 2020.

[7] Y. Zhou, G. Cheng, S. Jiang, and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Comput. Netw.*, vol. 174, Jun. 2020, Art. no. 107247.

[8] A. Verma and V. Ranga, "Machine learning based intrusion detection systems for IoT applications," *Wireless Pers. Commun.*, vol. 111, no. 4, pp. 2287–2310, Apr. 2020.

[9] P. Hadem, D. K. Saikia, and S. Moulik, "An SDN-based intrusion detection system using SVM with selective logging for IP traceback," *Comput. Netw.*, vol. 191, May 2021, Art. no. 108015.

[10] P. Kumar, G. P. Gupta, and R. Tripathi, "Toward design of an intelligent cyber attack detection system using hybrid feature reduced approach for IoT networks," *Arabian J. Sci. Eng.*, vol. 46, no. 4, pp. 3749–3778, Apr. 2021.

[11] A. Hục, J. šalej, and M. Trebar, "Analysis of machine learning algorithms for anomaly detection on edge devices," *Sensors*, vol. 21, no. 14, pp. 1–22, 2021.

[12] K. D. Devprasad, S. Ramanujam, and S. B. Rajendran, "Context adaptive ensemble classification mechanism with multi-criteria decision making for network intrusion detection," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 21, pp. 1–12, Jun. 2022.

[13] N. Gupta, V. Jindal, and P. Bedi, "CSE-IDS: Using cost-sensitive deep learning and ensemble algorithms to handle class imbalance in network-based intrusion detection systems," *Comput. Secur.*, vol. 112, Jan. 2022, Art. no. 102499.

[14] G. Çetin, "An effective classifier model for imbalanced network attack data," *Comput., Mater. Continua*, vol. 73, no. 3, pp. 4519–4539, 2022.

[15] W. Xu and Y. Fan, "Intrusion detection systems based on logarithmic autoencoder and XGBoost," *Secur. Commun. Netw.*, vol. 2022, pp. 1–8, Apr. 2022.

[16] Y. K. Saheed, A. I. Abiodun, S. Misra, M. K. Holone, and R. Colomo-Palacios, "A machine learning-based intrusion detection for detecting Internet of Things network attacks," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 9395–9409, Dec. 2022.

[17] T.-T.-H. Le, Y. E. Oktian, and H. Kim, "XGBoost for imbalanced multiclass classification-based industrial Internet of Things intrusion detection systems," *Sustainability*, vol. 14, no. 14, p. 8707, Jul. 2022.

[18] I. T. Utami, B. Sartono, and K. Sadik, "Comparison of single and ensemble classifiers of support vector machine and classification tree," *J. Math. Sci. Appl.*, vol. 2, no. 2, pp. 17–20, 2014.

[19] N. Amraoui and B. Zouari, "Anomalous behavior detection-based approach for authenticating smart home system users," *Int. J. Inf. Secur.*, vol. 21, no. 3, pp. 611–636, Jun. 2022.

[20] S. Khare and M. Totaro, "Ensemble learning for detecting attacks and anomalies in IoT smart home," in *Proc. 3rd Int. Conf. Data Intell. Secur. (ICDIS)*, Padre Island, TX, USA, Jun. 2020, pp. 56–63.

[21] M. Ajdani and H. Ghaffary, "Introduced a new method for enhancement of intrusion detection with random forest and PSO algorithm," *Secur. Privacy*, vol. 4, no. 2, pp. 1–10, Jan. 2021.

[22] A. K. M. M. R. Mazumder, N. M. Kamruzzaman, N. Akter, N. Arbe, and M. M. Rahman, "Network intrusion detection using hybrid machine learning model," in *Proc. Int. Conf. Adv. Electr., Comput., Commun. Sustain. Technol. (ICAECT)*, Bhilai, India, Feb. 2021, pp. 1–8.

[23] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient boosting feature selection with machine learning classifiers for intrusion detection on power grids," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 1, pp. 1104–1116, Mar. 2021.

[24] C. Bentéjac, A. Csörgo, and G. Martínez-Muñoz, "A comparative analysis of XGBoost," 2019, arXiv:1911.01914.

[25] B. S. Bhati, G. Chugh, F. Al-Turjman, and N. S. Bhati, "An improved ensemble based intrusion detection technique using XGBoost," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 6, pp. 1–15, Aug. 2020.

[26] S. Seth, G. Singh, and K. K. Chahal, "Anovel time-efficient learning-based approach for smart intrusion detection system," *J. Big Data*, vol. 8, no. 1, pp. 1–28, Dec. 2021.