Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Lightweight Intrusion Detection Approaches for IoT Networks: A Systematic Literature Review on Machine and Deep Learning Techniques

Dileepa Mabulage

dileepa.20212025@iit.ac.lk

Informatics Institute of Technology (IIT)

Banuka Athuraliya

Informatics Institute of Technology (IIT)

Systematic Review

Keywords: lightweight intrusion detection, machine learning, deep learning, systematic literature review, internet of things

Posted Date: May 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9622258/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Lightweight Intrusion Detection Approaches for IoT Networks: A Systematic Literature Review on Machine and Deep Learning Techniques

1st Dileepa Mabulage
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
dileepa.20212025@iit.ac.lk

2nd Banuka Athuraliya
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
banu.a@iit.ac.lk

***Abstract*—The rapid expansion of the Internet of Things (IoT) has connected a wide range of resource-constrained devices, including sensors, healthcare equipment, home appliances, and industrial systems, to the Internet. Effective intrusion detection systems are crucial, as IoT networks are increasingly targeted by cyberattacks. Although intrusion detection systems (IDS) have been built utilizing machine learning (ML) and deep learning (DL) techniques, their usability in IoT environments is limited by their high computational and memory requirements. Consequently, the design of lightweight intrusion detection systems has gained increasing research attention. This paper presents a systematic literature review conducted using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework to analyze existing lightweight IDS approaches for IoT devices. Relevant studies were systematically retrieved from five major databases, including Google Scholar, ScienceDirect, IEEE Xplore, Scopus, and SpringerLink. From an initial set of identified records, a final group of selected studies was obtained based on predefined research questions and inclusion and exclusion criteria. The selected studies were analyzed with respect to ML and DL techniques, datasets, and evaluation metrics. In addition, key limitations, research challenges, and future directions are identified to support the development of efficient and practical intrusion detection solutions for resource-constrained IoT environments.**

***Index Terms*—lightweight intrusion detection, machine learning, deep learning, systematic literature review, internet of things**

I. Introduction

Originally a theoretical concept, IoT has quickly expanded into a widespread technology that supports many important areas, such as smart cities, healthcare systems, industrial automation, and agriculture [1]. According to [2], billions of interconnected devices, ranging from sensors, actuators, and smart appliances, communicate with each other with minimal input from humans, enhancing automation, productivity, and decision-making. This vast connectivity has many social and economic advantages, although it comes with severe security risks. The growing number of IoT deployments have significantly increased the attack surface, leaving critical infrastructures and daily services open to cyberattacks [3].

IoT environments are much more challenging to secure than traditional IT networks. IoT ecosystems are highly diverse, relying on a range of hardware platforms, communication protocols, and deployment scenarios [4]. The IoT architecture is vulnerable to attacks at several levels, from physical-layer threats such as node capture and jamming to network-layer attacks, including Man-in-the-Middle and routing manipulation [5]. As a result, IDS have become a critical security mechanism for monitoring network behavior and identifying malicious activities [6].

However, conventional IDS solutions are often unsuitable for IoT environments due to their high computational and memory requirements [1]. Deploying resource-intensive security mechanisms on low-power IoT devices can lead to excessive energy consumption, reduced device lifetime, and degraded system performance [7]. This has driven growing research interest in developing lightweight IDS that deliver reliable detection accuracy while maintaining minimal resource overhead [8].

Recent research has looked into incorporating ML and DL methods into lightweight IDS designs in order to overcome this difficulty. [9] have observed that models like Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and ensemble learning approaches have shown great potential in detecting intricate, dynamic attack patterns. These models are frequently paired with effective feature engineering and dimensionality reduction strategies, such as Chi-square, Pearson correlation, and Information Gain, to guarantee effectiveness in resource-constrained environments. By eliminating redundant or irrelevant features, these approaches significantly reduce computational complexity while preserving detection performance [10]. This paper examines the development of lightweight, intelligent intrusion detection frameworks for IoT systems, with a particular focus on optimizing feature selection and model architectures to achieve effective and practical security solutions.

This study focuses on lightweight IDS for IoT environments, examining how existing detection techniques address the stringent computational, memory, and energy constraints

of resource-constrained devices. To achieve a balance between detection accuracy and system efficiency, the integration of machine learning techniques with lightweight feature engineering approaches is emphasized. The following research questions and objectives were formulated to guide the analysis and address the identified research gaps.

II. Background

As the number of IoT networks continues to rise, they have also become significant attack targets for cyberattacks. Due to the resource constraints and lack of security defense systems, IoT devices are vulnerable to cyberattacks such as Distributed Denial of Service (DDoS), botnets (Mirai), ransomware, and spoofing attacks [11], [12]. Traditional security solutions such as complex encryption and traditional firewalls are resource-intensive or ineffective in such settings [1]. As a result, Intrusion Detection Systems (IDS) have become an essential second line of defense [6].

An IDS is a software or hardware solution developed to monitor network traffic or system activity for malicious behavior or policy violations [7]. Upon detecting an intrusion, the system produces an alert signal for administrators to respond accordingly [13].

A. IDS Detection Techniques in IoT

1) Signature-based: Also referred to as misuse-based detection systems. These systems identify attacks by matching network traffic with a database of known attack signatures (patterns). Although successful at identifying known threats with low false positives, they are not very effective at identifying new, unknown (zero-day) attacks [10].

2) Anomaly-based: These systems build a profile of normal network behavior and flag deviations from this baseline as potential attacks. They are capable of detecting unknown attacks but often suffer from higher false positive rates [7].

3) Specification-based: These rely on a set of manually defined rules or specifications to define legitimate behavior. A behavior is flagged if it deviates from the list of allowed behaviors. Because a human expert sets all the criteria, the false-positive and -negative rates are much lower [10].

B. IDS Intrusion Data Sources

1) Host-based IDS (HIDS): Installed on individual devices to monitor internal logs and system files. It records the behavior and notifies the administrator or user to take the necessary action. An HIDS is an agent that determines whether an external or internal entity has breached the security policy of the system [7].

2) Network-based IDS (NIDS): Placed at network points (e.g., gateways/routers) to monitor traffic flowing through the entire network [7] and match them with a list of stored attack signatures in its system to detect suspicious activities. Apart from that, all incoming packets are examined by NIDS, which looks for any strange patterns. Depending on the severity level of the attack, the system can alert the administrators.

Given the computational constraints of IoT devices, there is a pressing need for lightweight IDS solutions that minimize CPU and memory usage [1]. In order to automate detection, increase accuracy, and efficiently handle high-dimensional traffic data, current research focuses on ML and DL approaches [14]. To lower model complexity for deployment on edge devices like Raspberry Pis, methods like feature selection and model quantization are used [15].

III. Methodology of the Systematic Literature Review

A. Research Questions

- RQ1: What ML and DL techniques are used for developing lightweight IDS systems for IoT systems?
- RQ2: What datasets are used to evaluate the existing lightweight IDS models in IoT systems?
- RQ3: What metrics are used to evaluate the performance of lightweight IDS in IoT systems?
- RQ4: What key challenges, gaps and future research directions were identified in lightweight IDS models in IoT devices?

This research utilizes the PRISMA framework for a transparent, reproducible, and rigorous synthesis of existing research [16]. The review process consisted of four sequential phases: identification, screening, eligibility assessment, and final study inclusion.

B. Search Strategy

The literature was systematically selected through a structured search of academic databases including Google Scholar, IEEE Xplore, SpringerLink, ScienceDirect, and Scopus. Publications from January 2022 to January 2026 were the main focus of the review. Electronic links to digital databases are used for the literature review.

Predefined Boolean search string is used to narrow down the research: ((“Lightweight IDS” OR “Lightweight Intrusion Detection System”) AND (“Internet of Things” OR IoT)). A total of 2641 records were initially identified in all databases. After removing 173 duplicate records, 40 inaccessible or non-English publications, and 4 retracted publications, 2424 unique studies were retained for screening.

C. Screening

The retrieved studies were screened during the screening phase using their titles and abstracts to filter out unnecessary articles, especially those unrelated to IDS or lightweight security solutions for IoT contexts. 744 articles remain for full-text review after 1680 records were eliminated at this point due to their lack of relevance. 724 papers proceeded to the eligibility assessment step after 20 of these studies were inaccessible due to paywall restrictions or unavailability of full texts.

D. Eligibility and Inclusion

Full-text articles were evaluated using predefined inclusion and exclusion criteria. Studies were included if they satisfied all the following conditions:

- Explicitly addressed IDS within IoT environments

- Employed ML and/or DL techniques with an emphasis on lightweight or resource-efficient designs
- Were peer-reviewed, written in English, and presented empirical results or comparative performance analyses

Studies were excluded if they did not focus on IoT-based intrusion detection on ML or DL components (577), were irrelevant to the scope of the research (96), or were survey or review papers (17). After applying these criteria, 34 studies met all the inclusion criteria and were selected for the final qualitative analysis. Review articles were excluded from the set of primary studies. Nevertheless, they were referenced to provide background and contextual support. The PRISMA flow diagram, which summarizes the number of records identified, screened, excluded, and incorporated at each stage, is shown in Fig. 1.

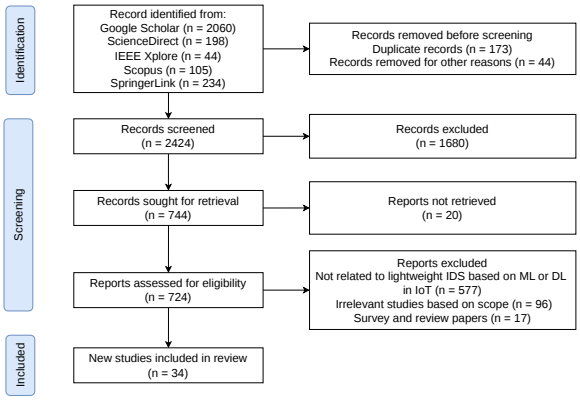


Fig. 1. PRISMA flow diagram.

E. Data Extraction and Analysis

Data from the 34 selected studies were systematically extracted to support a structured and objective evaluation of their contributions. The extracted information included authorship, year of publication, datasets used, detection techniques, feature selection or dimensionality reduction methods and reported performance metrics, along with the stated strengths and limitations of each approach.

A thematic analysis was conducted to identify recurring patterns and group the studies into thematic domains, such as lightweight model design and feature engineering strategies. These themes guided the critical analysis and research trends, limitations, and challenges in developing lightweight IDS for IoT environments.

IV. Results and Discussion

A. RQ1: What ML and DL techniques are used for developing lightweight IDS systems for IoT systems?

Recent studies that examined resource efficiency, computational overhead, suitability for IoT devices with limitations, real-time detection capability, robustness to errors, and changing attack patterns are summarized in this review of the literature. These elements are crucial for designing lightweight intrusion detection systems, which are constrained by limited processing power, memory, and energy availability. The key characteristics, techniques, and primary conclusions of representative lightweight IDS studies for IoT environments are summarized in Table II and discussed below.

[1] proposed IPCA-SAMKNN, a novel IDS designed for resource-constrained devices like the Raspberry Pi. They developed a lightweight machine learning model that can be trained on-device by combining Self-Adjusting Memory K-Nearest Neighbors (SAM-KNN) for classification with Incremental Principal Component Analysis (IPCA) for dimensionality reduction. Using the UNSW-NB15 dataset, the model achieved an accuracy of 98.91% while maintaining minimal resource overheads (using only 1.6% CPU and 1.4% memory).

[17] introduced Realguard, a lightweight Network Intrusion Detection System (NIDS) powered by a Deep Neural Network (DNN) intended for direct operation on IoT gateways. The framework incorporates an efficient feature extraction module based on a damped incremental statistic algorithm to process traffic in real-time. Experimental results on the CICIDS2017 dataset demonstrate that the system can detect ten types of attacks with an average accuracy of 99.57% while maintaining a high packet processing rate of over 10,000 packets per second on limited hardware.

[18] developed a novel IDS based on a political optimizer with a cascade forward neural network (PO-CFNN) specifically for IoT environments. Following the three main processes: preprocessing, classification and parameter optimization. The research contributes to the development of a unified IDS paradigm that can protect IoT networks from a variety of threats, such as denial-of-service (DoS) and distributed denial-of-service (DDoS)

[19] proposed a lightweight intrusion detection system utilizing a novel hybrid feature selection algorithm called MI2G, which combines Mutual Information and Information Gain. This method focuses on selecting features with high statistical dependence and entropy reduction to lower computational costs. The study validates the approach on the CICIDS2018 dataset, showing that the feature selection technique reduced training time by 27–63% while maintaining high performance metrics (¿99% accuracy) across various machine learning classifiers.

[20] study, investigated ensemble learning approaches for multiclass classification and threat detection in IoT systems. ML techniques such as bagging and boosting ensemble decision trees are applied in the process. The results show that earlier intrusion detection systems prioritized accuracy over execution speed, which is essential for maximizing model performance.

[21] designed a Lightweight Neural Network (LNN) that utilizes Principal Component Analysis (PCA) for feature dimensionality reduction to avoid high model complexity. The classifier employs expansion and compression structures, inverse residual structures, and channel shuffle operations to achieve effective feature extraction with low computational cost. Additionally, they introduced a specialized NID Loss

function to handle class imbalance, demonstrating excellent classification performance on the Bot-IoT and UNSW-NB15 datasets with a small model size.

[22] suggested a lightweight IDS model based on Deep Learning using a Multi-Layer Perceptron (MLP) network specifically for the MQTT protocol. The model focuses on high speed and accuracy by dealing with a reduced set of features relevant to MQTT communication. Validated on the MQTTset dataset, the proposed method reduced the number of features by approximately 50% while achieving an accuracy of 96.97%, proving its efficiency for constrained IoT environments.

[23] proposed a lightweight and efficient intrusion detection method based on feature grouping and session merging to reduce data size and computational resources. They designed a fast protocol parsing method to generate semantic-level features from raw packet capture files, avoiding complex flow-level calculations. The method achieved more than 99.5% classification accuracy on three public IoT datasets, requiring significantly fewer computational resources for protocol parsing and model training compared to baseline methods.

In the [24] study five machine learning (ML) methods including Random Forest, K-Nearest Neighbors, Decision Tree, Gradient Boosting, and AdaBoost are used to improve performance on datasets that contain both attacks and typical scenarios. Five-fold cross-validation is part of the process to help reduce bias and variance problems and guarantee consistent model performance on unseen data.

[25] proposed a lightweight method based on two-stage feature selection and Bayesian optimization. They first used the Spearman Correlation Coefficient (SCC) to remove redundant features, followed by the Salp Swarm Algorithm (SSA) to find the optimal feature subset. Finally, they constructed a LightGBM classifier optimized via Bayesian methods. The output model achieved a compact size of only 1.77 MB on the UNSW-NB15 dataset with 96.08% accuracy, making it highly suitable for deployment on IoT terminal devices.

In [15] study, authors developed a quantized hybrid deep learning model DNN-BiLSTMQ, which incorporates feature engineering methods such as two-stage Incremental Principal Component Analysis (IPCA) and Redundancy-Adjusted Logistic Mutual Information Feature Selection (RAL-MIFS). The work effectively trimmed the model size to just 25.6 KB using advanced quantization techniques such as Post-Training Dynamic Quantization (PTDQ) and Quantization Aware Training (QAT). This method achieves 99.73% accuracy in the CIC-IDS2017, addressing the rigid memory constraints of IoT devices.

[6] introduced a lightweight IDS utilizing a mean thresholding-based Fast Correlation-Based Filter (mFCBF) for feature selection combined with LightGBM and XGBoost classifiers. The aim of the study was to use SVM-SMOTE to address data imbalance and increase efficiency while preserving detection accuracy. With an accuracy of 99.75% and an average CPU consumption of only 2.03%, experimental results on the BoT-IoT dataset showed that the system is both lightweight and resilient to new threats.

In [14], authors presented a lightweight hybrid approach combining Bidirectional Long Short-Term Memory (BiLSTM) networks for temporal dependencies and Convolutional Neural Networks (CNN) for spatial feature extraction. They employed a Chi-square feature selection method to significantly reduce input complexity by selecting the 20 most relevant features. Benchmarking against the UNSW-NB15 dataset, the model achieved 97.90% accuracy with a rapid prediction time, validating its suitability for real-time deployment on IoT devices.

[26] introduced a lightweight feature aggregation mechanism in the pre-processing phase to model fine-grained temporal characteristics of network traffic. The approach integrates session-based, flow-based, and sliding-window aggregation schemes within a multi-layered processing pipeline. The findings indicate that this aggregation strategy enhances the accuracy of subsequent learning models while lowering computational costs.

TABLE I
ABBREVIATIONS

Abbreviation	Notation
CNN	Convolutional Neural Network
DNN	Deep Neural Network
PO-CFNN	Political Optimizer with Cascade Forward NN
DT	Decision Tree
RF	Random Forest
LR	Logistic Regression
LDA	Linear Discriminant Analysis
NB	Naïve Bayes
SVM	Support Vector Machine
GBM	Gradient Boosting Machine
PCA	Principal Component Analysis
LNN	Lightweight Neural Network
MLP	Multi-Layer Perceptron
KNN	K-Nearest Neighbors
SCC	Spearman Correlation Coefficient
SAM-KNN	Self-Adjusting Memory KNN
SSA	Salp Swarm Algorithm
IPCA	Incremental Principal Component Analysis
LSTM	Long Short-Term Memory
BiLSTM	Bidirectional Long Short-Term Memory
FFNN	Feed-Forward Neural Network

B. RQ2: What datasets are used to evaluate the existing lightweight IDS models in IoT systems?

Based on the selected studies, the four major datasets are predominantly utilized for evaluating IDS in IoT environments. Those are UNSW-NB15 followed by BoT-IoT, CICIDS2017 and ToN-IoT. While older datasets such as KDD99 and NSL-KDD are frequently cited, they are less relevant as outdated and lacking the low-footprint attack scenarios that are frequent in modern IoT networks [7], [10].

1) UNSW-NB15: Developed by the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS), this dataset represents a hybrid of real modern normal activities and synthetic contemporary attack behaviors. It contains approximately 2.54 million feature vectors with 49 features, categorized into nine distinct attack families including Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and

TABLE II
RELATED WORK

No	Ref	Feature Engineering Techniques	Classification / Detection Mechanism	Limitations	Datasets
1	[1]	IPCA	SAM-KNN	Reliance on static datasets, resource constraints	UNSW-NB15
2	[17]	Damped incremental statistics	DNN, CNN	Increased hidden nodes reduce detection speed	CICIDS-2017
3	[18]	N/A	PO-CFNN	N/A	NSL-KDD++
4	[19]	Mutual Information + Information Gain	DT, RF, LR, LDA, NB, SVM, GBM	High training time for SVM/GBM, low NB accuracy	CICIDS-2018, UNSW-NB15
5	[20]	Scaling, transformation	Bagging, XGBoost, LightGBM, CatBoost	High execution time for single classifiers	ToN-IoT
6	[21]	PCA	LNN	N/A	UNSW-NB15, BoT-IoT
7	[22]	Normalization	MLP	N/A	MQTTset
8	[23]	Packet grouping, session merging	DT, RF, KNN, XGBoost	N/A	BoT-IoT, MedBIoT, MQTT-IoT-IDS2020
9	[24]	N/A	RF, DT, GB, AdaBoost, KNN	N/A	BoTNeTIoT-L01
10	[25]	SCC + SSA	LightGBM (Bayesian optimized)	No standard SCC thresholds	ToN-IoT, UNSW-NB15
11	[15]	IPCA, Quantization	DNN-BiLSTMQ	Requires effective dimensionality reduction	CIC-IDS2017, CIC-IoT2023
12	[6]	mFCBF	LightGBM, XGBoost	Class imbalance leads biased results	CICIDS-2017, NSL-KDD, BoT-IoT
13	[14]	Chi-square	BiLSTM	DL models remain computationally heavy	UNSW-NB15
14	[26]	PCA	LSTM, BiLSTM, RF, SVM, XGBoost, FFNN	High training overhead	Custom Dataset

Worms [27]. The lack of contemporary attack scenarios in older benchmarks like KDD99 is addressed with this dataset, which is highly recommended in the literature.

2) BoT-IoT: This dataset was specifically designed to simulate botnet traffic within a realistic IoT network environment [28], incorporating traffic from devices such as weather stations, smart fridges, and motion-activated lights. With over 99% attack traffic and fewer than 1% benign traffic, it is distinguished by a significant class imbalance that frequently calls for the use of data balancing strategies like SMOTE during preprocessing [29]. The dataset focuses on botnet-centric attacks that fall into Keylogging, Data Exfiltration, Service Scan, OS Fingerprint, and UDP categories.

3) CICIDS2017: Published by the Canadian Institute for Cybersecurity [30], it includes the most recent attacks with both malicious and benign network traffic traces. There are 84 features in this labeled dataset. There are 14 attack types in this dataset. Some of the attack types are significant in the aspect IoT detection [11]. The dataset is particularly noted for providing raw PCAP files, allowing researchers to extract features and assign labels at the packet level for granular analysis [17].

4) ToN-IoT: It is a heterogeneous dataset that uniquely collects data from multiple layers of an IoT ecosystem, including the edge, fog, and cloud layers, making it highly capable for evaluating distributed IDS architectures. ToN-IoT was created by the Cyber Range and the UNSW Canberra IoT Labs. The ToN-IoT dataset is available in the ToN-IoT repository [31]. Nine attack types, including scanning, DoS, DDoS, ransomware, backdoors, data injection, cross-site scripting (XSS), password cracking, and Man-in-the-Middle (MITM) attacks, are included in the dataset along with 43 feature attributes [25].

C. RQ3: What metrics are used to evaluate the performance of lightweight IDS in IoT systems?

These metrics are categorized into two main groups: network performance metrics and data mining metrics. The majority of research used well-known mining metrics, such as accuracy, F1-Score, precision, and recall, to evaluate lightweight models. Several researchers have employed network performance metrics to evaluate their models using network traffic and resource utilization data like CPU usage, Energy usage and Model Running Time [1].

Data mining metrics are derived from the Confusion Matrix, which tracks True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) [25]. TP denote the number of positive samples that are correctly classified, while TN represent negative samples that are correctly identified. FP correspond to negative samples that are incorrectly classified as positive, and FN refer to positive samples that are incorrectly classified as negative. Based on these definitions, several standard performance evaluation metrics are defined as follows.

Accuracy measures the overall correctness of the model by computing the proportion of correctly classified samples among all samples:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision evaluates the reliability of positive predictions and is defined as the ratio of correctly predicted positive samples

to all samples predicted as positive:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall, also known as sensitivity, quantifies the model's ability to identify all actual positive samples:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

The F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's classification performance:

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

False Positive Rate (FPR) indicates the percentage of normal traffic incorrectly classified as attacks. A low FPR is critical to prevent the IDS from disrupting legitimate IoT services:

$$\text{FPR} = \frac{FP}{FP + TN} \tag{5}$$

D. RQ4: What key challenges, gaps and future research directions were identified in lightweight IDS models in IoT devices?

This study aims to critically evaluate existing lightweight IDS techniques to assess their effectiveness and limitations. The analysis supports researchers in selecting appropriate methods during the design of IDS solutions for IoT devices. The subsequent section discusses the open issues, challenges, and future research directions for lightweight ML and DL based IDS models in IoT environments.

V. Open Issues, Challenges, and Future Research Directions

This section discusses the open challenges identified in the literature and outlines future research directions for lightweight machine learning- and deep learning-based intrusion detection systems for IoT, aiming to guide further improvements to existing approaches

- **Challenge 1- Heterogeneous data in IoT networks:** The IoT ecosystem includes a wide variety of devices with distinct processing power, hardware features, and functionality [1]. That becomes handling inconsistent feature dimensions and discrete features across such a fragmented environment poses substantial challenges for IDS models, which make it difficult to detect threats that may change significantly across different device types [2].
- **Challenge 2- Feature selection techniques:** High-dimensional data with many redundant or unnecessary features is frequently present in IoT network traffic, which can reduce detection accuracy and raise computational overhead [32]. Additionally, harmful traffic may be misclassified as a result of poor feature selection [29]. Promising results are obtained when feature selection approaches are applied in the research to reduce the complexity of the suggested models. Additionally, hybrid feature selection techniques produce excellent results for lightweight IDS model creation and feature selection.
- **Challenge 3- Real-time intrusion detection:** The dynamic nature of IoT attacks, such as Distributed Denial of Service (DDoS), requires immediate identification and mitigation to prevent system breakdown, demanding IDS solutions capable of operating in real-time [11].
- **Challenge 4- Dataset Limitation:** The use of outdated datasets, like KDD99 or NSL-KDD, which do not accurately reflect modern IoT protocols or threat signatures, usually restricts research in IoT IDS [7]. Due to privacy concerns, it is also difficult to properly train supervised models because there is a lack of high-quality, labeled, real-world traffic data [33].
- **Challenge 5- Model computational complexity:** The limited processing power, memory, and battery life of IoT devices significantly constrain the implementation of Deep Learning models on these devices [17]. Lightweight designs that use techniques like model quantization, structural compression, and reduced neural networks must balance detection efficacy and efficiency in order to fit under the rigid resource constraints of IoT devices [15].
- **Challenge 6- Multiple attack detection:** In real-world IoT traffic and benchmark datasets, normal traffic and high-volume attacks (like DoS/DDoS) vastly outnumber low-volume attacks (like Worms, Backdoors, or Infiltration). IDS often exhibit poor performance on minority classes because they are biased toward the majority classes [34]. Effective zero-day attack detection requires IDS to use more than just conventional techniques. AEs and other advanced approaches should be used [35].
- **Challenge 7- Model interpretability and explainability:** The majority of lightweight IDS models based on machine learning and deep learning operate as black-box systems, where the internal decision-making process is not easily understandable. Although these methods have high detection accuracy, they don't provide much insight into how predictions are made. It's challenging to comprehend, validate, and trust the detection decisions due to this lack of interpretability [24].

The following points are recommended as future directions in the field of lightweight IDS solutions based on ML and DL to detect various threats in IoT, depending on the concerns and limitations revealed by the analyzed studies.

- The use of cloud computing, and blockchain to address the challenge of massive data generated by IoT devices. Collaborative edge-cloud architecture: local gateways perform specific tasks to handle diverse traffic patterns before it reaches the cloud **Challenge 1**.
- To develop lightweight IDS techniques for IoT devices, further research combining feature selection with ML and DL classification algorithms will be worthwhile. To choose the best set of features with minimal computing complexity and resource usage **Challenge 2**.
- Research is moving towards implementing incremental learning, which allows systems to update with new data without the latency of complete retraining, thereby facil-

itating faster response times **Challenge 3**.

- Prioritize the creation of high-quality, real-time datasets that accurately reflect modern IoT attack vectors and protocols to replace outdated benchmarks. In addition to employing sophisticated synthetic data generation techniques, researchers plan to create unsupervised and semi-supervised learning models to solve the lack of labelled data **Challenge 4**.
- The focus of future research is on advanced model compression techniques, such as pruning redundant neural connections and utilizing quantization-aware training to drastically reduce model size for deployment on tiny devices **Challenge 5**.
- Additional studies are needed to explore and analyze new forms of attacks aimed at various layers of IoT. Furthermore, it is advisable to examine multiple attack scenarios and create effective detection systems. This will address the issues in **Challenge 6**.
- Incorporating Explainable Artificial Intelligence (XAI) methods like Local interpretable model-agnostic explanations (LIME), Shapley additive explanations (SHAP), or feature importance analysis can enhance model behavior analysis and deliver insights into how input features affect prediction results. As a result, trust and accountability can be achieved. This will address the issues in **Challenge 7**.

VI. Conclusion

This study presents a comprehensive systematic review of lightweight intrusion detection techniques for IoT systems. Relevant literature was systematically retrieved from five major academic databases: Google Scholar, ScienceDirect, IEEE Xplore, Scopus, and Springer, covering publications from 2022 to 2026. The identified studies underwent multiple screening and refinement stages according to predefined criteria, resulting in a final selection of 34 primary studies. These selected works were then critically analyzed to address the formulated research questions. The analysis indicates significant, steady growth in research publications on lightweight IDS solutions over the examined period.

The literature indicates that integrating ML and DL is currently the most effective approach for securing IoT networks against evolving cyber threats, surpassing traditional signature-based methods in detecting complex attacks while there is an increasing interest in incorporating deep learning approaches. In addition, several studies adopt hybrid models that integrate ML and DL methods to improve detection performance while maintaining computational efficiency. The analysis further demonstrates that feature engineering is a key component in developing lightweight IDS models, with correlation-based feature selection methods being more widely applied than alternative techniques. Furthermore, advanced optimization techniques such as quantization have proven capable of shrinking DL models by over 99%. Most implementations also rely on Python as the primary programming environment for IDS model development.

In summary, this review identifies several open challenges that hinder the deployment of effective lightweight intrusion detection systems in IoT environments. These issues include device and traffic heterogeneity, high-dimensional feature spaces, the need for real-time detection, constraints imposed by current benchmark datasets, and the computational complexity of deep learning models. Future research should focus on cooperative edge-cloud architectures, effective feature selection in conjunction with lightweight learning models, and incremental learning mechanisms to enable prompt adaptation to novel attacks.

References

[1] P. R. Agbedanu, R. Musabe, J. Rwigema, I. Gatare, and Y. Pavlidis, "IPCA-SAMKNN: A novel network IDS for resource constrained devices," in *2022 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*. IEEE, pp. 540–545. [Online]. Available: https://ieeexplore.ieee.org/document/10180926

[2] N. Dat-Thinh, H. Xuan-Ninh, and L. Kim-Hung, "MidSiot: A multistage intrusion detection system for internet of things," vol. 2022, no. 1, p. 9173291. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2022/9173291

[3] Z. Sun, G. An, Y. Yang, and Y. Liu, "Optimized machine learning enabled intrusion detection 2 system for internet of medical things," vol. 6, p. 100056. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2773186323000506

[4] Y. Otoum, D. Liu, and A. Nayak, "DL-IDS: a deep learning–based intrusion detection framework for securing IoT," vol. 33, no. 3, p. e3803, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ett.3803.

[5] N. Sarwar, I. S. Bajwa, M. Z. Hussain, M. Ibrahim, and K. Saleem, "IoT network anomaly detection in smart homes using machine learning," vol. 11, pp. 119 462–119 480. [Online]. Available: https://ieeexplore.ieee.org/document/10287977/

[6] J. P. Kushwaha, S. Bhadauria, and S. Tapaswi, "mFCBF based lightweight intrusion detection system for IoT networks," vol. 28, no. 6, p. 416. [Online]. Available: https://link.springer.com/10.1007/s10586-024-05022-0

[7] A. R. Gad, M. Haggag, A. A. Nashat, and T. M. Barakat, "A distributed intrusion detection system using machine learning for IoT based on ToN-IoT dataset," vol. 13, no. 6.

[8] S. Sadhwani, A. Sriwastawa, R. Muthalagu, and P. M. Pawar, "Feature engineering based intelligent intrusion detection system for IoT network security," vol. 16, no. 9, pp. 6487–6511. [Online]. Available: https://link.springer.com/10.1007/s13042-025-02631-y

[9] G. A. Mukhaini, M. Anbar, S. Manickam, T. A. Al-Amiedy, and A. A. Momani, "A systematic literature review of recent lightweight detection approaches leveraging machine and deep learning mechanisms in internet of things networks," vol. 36, no. 1, p. 101866. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1319157823004202

[10] M. Bhavsar, K. Roy, J. Kelly, and O. Olusola, "Anomaly-based intrusion detection system for IoT application," vol. 3, no. 1, p. 5. [Online]. Available: https://link.springer.com/10.1007/s43926-023-00034-5

[11] M. Ahsan, N. Rifat, M. Chowdhury, and R. Gomes, "Intrusion detection for IoT network security with deep neural network," in *2022 IEEE International Conference on Electro Information Technology (eIT)*. IEEE, pp. 467–472, ISSN: 2154-0373. [Online]. Available: https://ieeexplore.ieee.org/document/9814006

[12] K. Saurabh, T. Kumar, U. Singh, O. P. Vyas, and R. Khondoker, "NFDLM: A lightweight network flow based deep learning model for DDoS attack detection in IoT domains," in *2022 IEEE World AI IoT Congress (AIIoT)*, pp. 736–742. [Online]. Available: http://arxiv.org/abs/2207.10803

[13] M. Samantaray, R. C. Barik, and A. Kumar Biswal, "Ensemble-based lightweight machine learning optimization for IoT network intrusion detection," in *2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*. IEEE, pp. 1101–1106. [Online]. Available: https://ieeexplore.ieee.org/document/10969343

[14] H. Benaddi, M. Jouhari, and O. Elharrouss, "A lightweight hybrid approach for intrusion detection systems using a chi-square feature selection approach in IoT," vol. 32, p. 101624. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2542660525001386

[15] S. F. Misrak and H. M. Melaku, "Lightweight intrusion detection system for IoT with improved feature engineering and advanced dynamic quantization," vol. 5, no. 1, p. 97. [Online]. Available: https://link.springer.com/10.1007/s43926-025-00203-8

[16] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," p. n71. [Online]. Available: https://www.bmj.com/lookup/doi/10.1136/bmj.n71

[17] X.-H. Nguyen, X.-D. Nguyen, H.-H. Huynh, and K.-H. Le, "Realguard: A lightweight network intrusion detection system for IoT gateways," vol. 22, no. 2, p. 432. [Online]. Available: https://www.mdpi.com/1424-8220/22/2/432

[18] M. I. Alghamdi, "A hybrid model for intrusion detection in IoT applications," vol. 2022, no. 1, p. 4553502. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2022/4553502

[19] S. Kaushik, A. Bhardwaj, A. Alomari, S. Bharany, A. Alsirhani, and M. Mujib Alshahrani, "Efficient, lightweight cyber intrusion detection system for IoT ecosystems using MI2g algorithm," vol. 11, no. 10, p. 142. [Online]. Available: https://www.mdpi.com/2073-431X/11/10/142

[20] D. Rani, N. S. Gill, P. Gulia, and J. M. Chatterjee, "An ensemble-based multiclass classifier for intrusion detection using internet of things," vol. 2022, pp. 1–16. [Online]. Available: https://www.hindawi.com/journals/cin/2022/1668676/

[21] R. Zhao, G. Gui, Z. Xue, J. Yin, T. Ohtsuki, B. Adebisi, and H. Gacanin, "A novel intrusion detection method based on lightweight neural network for internet of things," vol. 9, no. 12, pp. 9960–9972. [Online]. Available: https://ieeexplore.ieee.org/document/9566308/

[22] M. S. Mahmood and N. B. Al Dabagh, "Improving IoT security using lightweight based deep learning protection model," vol. 30, no. 1, pp. 119–129. [Online]. Available: https://tj-es.com/ojs/index.php/tjes/article/view/875

[23] M. He, Y. Huang, X. Wang, P. Wei, and X. Wang, "A lightweight and efficient IoT intrusion detection method based on feature grouping," vol. 11, no. 2, pp. 2935–2949. [Online]. Available: https://ieeexplore.ieee.org/document/10177934

[24] M. Z. Mahmud, S. Islam, S. R. Alve, and A. J. Pial, "Optimized IoT intrusion detection using machine learning technique," in *2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON)*, pp. 167–172. [Online]. Available: http://arxiv.org/abs/2412.02845

[25] D. Zhang, D. Huang, Y. Chen, S. Lin, and C. Li, "A lightweight IoT intrusion detection method based on two-stage feature selection and bayesian optimization," vol. 9, no. 3, pp. 359–389. [Online]. Available: http://www.aimspress.com/article/doi/10.3934/electreng.2025017

[26] B. Boswell, S. Barrett, S. Rajaganapathy, and G. Dorai, "FLARE: Feature-based lightweight aggregation for robust evaluation of IoT intrusion detection." [Online]. Available: http://arxiv.org/abs/2504.15375

[27] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*. IEEE, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7348942/

[28] 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," vol. 100, pp. 779–796. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167739X18327687

[29] S. A. Khanday, H. Fatima, and N. Rakesh, "Implementation of intrusion detection model for DDoS attacks in lightweight IoT networks," vol. 215, p. 119330. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S095741742202348X

[30] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization:," in *Proceedings of the 4th International Conference on Information Systems Security and Privacy*. SCITEPRESS - Science and Technology Publications, pp. 108–116.

[31] N. Moustafa, "A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_iot datasets," vol. 72, p. 102994. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2210670721002808

[32] K. Albulayhi, Q. Abu Al-Haija, S. A. Alsuhibany, A. A. Jillepalli, M. Ashrafuzzaman, and F. T. Sheldon, "IoT intrusion detection using machine learning with a novel high performing feature selection method," vol. 12, no. 10, p. 5015. [Online]. Available: https://www.mdpi.com/2076-3417/12/10/5015

[33] M. Vishwakarma and N. Kesswani, "DIDS: A deep neural network based real-time intrusion detection system for IoT," vol. 5, p. 100142. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S277266222200073X

[34] A. Basati and M. M. Faghih, "DFE: efficient IoT network intrusion detection using deep feature extraction," vol. 34, no. 18, pp. 15 175–15 195. [Online]. Available: https://doi.org/10.1007/s00521-021-06826-6

[35] S. K. R. Mallidi and R. R. Ramisetty, "Optimizing intrusion detection for IoT : A systematic review of machine learning and deep learning approaches with feature selection and data balancing," vol. 15, no. 2, p. e70008. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/10.1002/widm.70008