A Hybrid Deterministic–Machine Learning Framework for Band width- Efficient DDoS Detection in IoT Networks

Mugerwa Joseph, Kitumba David & Udosen Alfred Akpan

Babcock University

ABSTRACT

The spread of Internet of Things (IoT) devices has significantly expanded the attack surface for Distributed Denial-of-Service (DDoS) threats, exposing resource-constrained gateways to bandwidth exhaustion and service disruption. While machine learning (ML)–based detection systems achieve strong accuracy, their computational cost renders them impractical for IoT environments. Conversely, lightweight deterministic filters provide efficiency but lack adaptability to evolving attack strategies. This study presents a Hybrid Deterministic–Machine Learning (HD-ML) framework that integrates deterministic packet verification with lightweight supervised classifiers to achieve both scalability and adaptability. The framework filters trivially malicious traffic at the gateway and forwards only residual ambiguous flows for ML-based classification. Using NS-3 simulations, we generated a dataset of over 100,000 packets, extracted flow-level features, and evaluated multiple classifiers including Decision Tree, Naïve Bayes, Logistic Regression, Random Forest, and Support Vector Machine (SVM).

Keywords: hybrid detection, deterministic filtering, machine learning, iot security, distributed denial-of- service (DDoS), lightweight defense.

Classification: DDC Code: 006.31

Language: English

Great Britain Journals Press

LJP Copyright ID: 975811
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 25 | Issue 5 | Compilation 1.0

A Hybrid Deterministic–Machine Learning Framework for Band width-Efficient DDoS Detection in IoT Networks

Mugerwa Joseph[α], Kitumba David[σ] & Udosen Alfred Akpan[ρ]

ABSTRACT

The spread of Internet of Things (IoT) devices has significantly expanded the attack surface for Distributed Denial-of-Service (DDoS) threats, exposing resource-constrained gateways to bandwidth exhaustion and service disruption. While machine learning (ML)–based detection systems achieve strong accuracy, their computational cost renders them impractical for IoT environments. Conversely, lightweight deterministic filters provide efficiency but lack adaptability to evolving attack strategies. This study presents a Hybrid Deterministic–Machine Learning (HD-ML) framework that integrates deterministic packet verification with lightweight supervised classifiers to achieve both scalability and adaptability. The framework filters trivially malicious traffic at the gateway and forwards only residual ambiguous flows for ML-based classification. Using NS-3 simulations, we generated a dataset of over 100,000 packets, extracted flow-level features, and evaluated multiple classifiers including Decision Tree, Naïve Bayes, Logistic Regression, Random Forest, and Support Vector Machine (SVM). Results demonstrate that the HD-ML framework achieves an overall detection accuracy of 98.8% with a false positive rate as low as 0.8%, significantly out performing standalone deterministic or ML-based approaches. Among the classifiers, SVM exhibited the highest performance with a perfect ROC-AUC score of 1.0 and an F1-Score of 0.926, confirming its suitability for residual traffic analysis. The proposed framework therefore offers a bandwidth-efficient, computationally light weight, and adaptive defense mechanism for real-time DDoS mitigation in IoT networks.

Keywords: hybrid detection, deterministic filtering, machine learning, iot security, distributed denial-of- service (DDoS), lightweight defense.

Author α: Department of Computer Science, Babcock University, Ilishan Remo, Ogun State, Nigeria; Department of Computing and Informatics, Bugema University, Kampala, Uganda.
σ: Department of Computing and Informatics, Bugema University, Kampala, Uganda.
ρ: Department of Computer Science, Babcock University, Ilishan Remo, Ogun State, Nigeria.

I. INTRODUCTION

The exponential growth of the Internet of Things (IoT) has created a vast ecosystem of interconnected devices that underpin critical infrastructures, smart homes, healthcare systems, and industrial applications. By 2030, it is estimated that over 29 billion IoT devices will be deployed globally, many of which will operate under constrained computational and power resources [1], [2]. While this proliferation offers tremendous societal and economic benefits, it simultaneously expands the cyber- attack surface, making IoT environments attractive targets for large-scale Distributed Denial- of-Service (DDoS) campaigns [3], [4].

DDoS attacks exploit vulnerabilities in network protocols and device configurations to overwhelm services with illegitimate traffic, resulting in bandwidth depletion, service disruption, and in some cases, cascading failures across dependent infrastructures [5], [6]. The Mirai botnet attack of 2016 demonstrated the destructive potential of IoT-driven DDoS, where thousands of compromised cameras and routers were harnessed to bring down large portions of the

London Journal of Research in Computer Science & Technology

Internet [7]. Since then, DDoS-for-hire services have made such attacks more accessible, enabling even low-skilled actors to launch sophisticated volumetric and protocol-based attacks [8].

Existing DDoS detection and mitigation strategies broadly fall into three categories: threshold/ entropy- based monitoring, machine learning (ML)- driven anomaly detection, and deterministic lightweight filtering, [9], [10], [11]. Threshold-based approaches are simple to deploy but prone to false alarms, especially under dynamic traffic conditions [10]. ML-based methods, particularly those leveraging deep learning models such as LSTM and GRU networks, demonstrate strong accuracy (often >95%) but require extensive labeled datasets and computational resources, making them unsuitable for real-time IoT edge deployments, [6], [12]. Deterministic approaches, by contrast, are lightweight and efficient but often limited in scope to specific protocols or attack types. For instance, [13] introduced a Message Authentication Code (MAC)-based ICMP verification algorithm that successfully detected bandwidth- depleting DDoS attacks with negligible overhead. However, its focus on ICMP traffic leaves other vectors, such as TCP SYN and UDP floods, insufficiently addressed.

This gap highlights the need for hybrid solutions that combine the efficiency of deterministic methods with the adaptability of ML classifiers, particularly in resource-constrained IoT environments. Such an approach allows deterministic filtering to quickly eliminate spoofed and obviously malicious traffic, while residual suspicious flows can be subjected to lightweight ML-based classification for fine-grained detection. By leveraging both layers, hybrid frameworks can achieve improved accuracy and reduced false positives without overburdening IoT gateways.

In this study, we propose a Hybrid Deterministic–Machine Learning Framework tailored for bandwidth- efficient DDoS detection in IoT networks. The framework integrates MAC-based packet verification with lightweight ML classifiers (Decision Trees, Logistic Regression, Support Vector Machine, Random Forests and Naïve Bayes) to detect diverse attack vectors, including ICMP floods, TCP SYN floods, and UDP-based volumetric attacks. Simulation experiments conducted in NS-3 evaluate the framework's detection accuracy, false positive rates, resource utilization, and latency overhead. The results are benchmarked against deterministic-only and ML-only baselines, demonstrating the hybrid model's suitability for IoT gateway deployment. This work contributes a practical, scalable, and adaptive security mechanism that balances accuracy and efficiency in defending IoT infrastructures against evolving DDoS threats.

II. LITERATURE REVIEW

2.1 DDoS Threat Landscape in IoT Networks

The IoT ecosystem is uniquely vulnerable to DDoS attacks due to its massive scale, heterogeneity, and limited device security [2], [3]. IoT-driven botnets such as Mirai and Persirai have been widely documented as platforms for launching large-scale volumetric attacks that disrupt services ranging from cloud applications to critical infrastructure [4], [7]. Recent studies emphasize that IoT-based DDoS attacks are not only increasing in frequency but also evolving toward multi-vector strategies that combine ICMP floods, TCP SYN floods, and UDP amplification, [8], [14].

The low computational power and insecure configurations of IoT devices make them easy to compromise and incorporate into botnets [15], [16]. Moreover, the use of lightweight communication protocols such as MQTT and CoAP further expands the attack surface, as they are often deployed without robust security measures[17].

2.2 Deterministic Approaches to DDoS Detection

Deterministic and rule-based methods focus on protocol-level verification or statistical signatures of abnormal traffic. Threshold-based approaches monitor traffic rates and trigger alerts when anomalies exceed predefined limits [10]. While efficient, these techniques are highly sensitive to

dynamic traffic patterns and often yield high false positive rates [5].

Entropy-based methods assess the randomness of traffic distributions to identify anomalies[18]. However, they struggle with stealthy, low-rate attacks that mimic legitimate traffic patterns. More recently, [13] proposed a lightweight MAC-based ICMP verification algorithm that achieved an 88.9% detection accuracy with zero false positives. Although effective for ICMP-based bandwidth depletion, this method is protocol-specific and less effective against TCP or UDP floods.

2.3 Machine Learning-Based Approaches

Machine learning has been widely applied in DDoS detection due to its capacity to model complex patterns in traffic data and adapt to new attack behaviors. Supervised models such as Decision Trees, Random Forest, and Support Vector Machines (SVM) have demonstrated effectiveness in classifying benign versus malicious traffic in IoT datasets, [19], [20]. Random Forest, in particular, has shown robustness in handling imbalanced data distributions common in network traces [21]. However, these models often require feature-rich datasets and may not scale efficiently in real-time IoT environments [22].

Unsupervised methods such as clustering and Principal Component Analysis (PCA) offer the advantage of not requiring labeled datasets [23]. These techniques are useful for anomaly detection in dynamic IoT networks, though they may suffer from higher false positive rates under high-variability traffic conditions.

Deep learning approaches have also gained attention, particularly Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and hybrid CNN-LSTM architectures [24], [25]. These methods outperform traditional ML in accuracy, often surpassing 98% detection rates on benchmark datasets [6]. However, their high computational cost, reliance on GPU acceleration, and need for large training sets render them impractical for deployment at IoT gateways, where resources are scarce [26]. This mismatch underscores the necessity of lighter models tailored for constrained environments.

2.4 Hybrid Approaches and Emerging Trends

Hybrid detection frameworks have emerged as a promising avenue to balance efficiency and adaptability. For example, [27] proposed a hybrid entropy–SVM approach, where entropy filtering reduced data volume before SVM classification. While effective, their model introduced latency unsuitable for real-time IoT applications. [28] combined statistical profiling with Random Forest classifiers, demonstrating improved accuracy but still incurring considerable processing overhead.

In the IoT context, hybrid methods often leverage deterministic filters at the edge to reduce traffic noise, while ML algorithms provide adaptive classification of residual traffic [29]. This two-layer approach shows promise, but most implementations still rely on computationally heavy ML models that strain gateway devices [21].

Beyond ML, SDN-assisted frameworks offer centralized traffic visibility and dynamic mitigation strategies [11], [30]. Yet, their reliance on controller communication can introduce additional bottlenecks and single points of failure in distributed IoT deployments [31]. Blockchain-based frameworks, on the other hand, aim to enhance trust and decentralization by maintaining immutable records of attack signatures [32]. While conceptually attractive, blockchain's storage and latency overhead limits its suitability for latency-sensitive IoT contexts [33].

Overall, while hybrid solutions point in the right direction, there remains a lack of models that intentionally combine deterministic filters with lightweight ML classifiers optimized for IoT gateways.

2.5 Research Gap

From the surveyed literature, several gaps emerge that motivate this study:

1. *Over-reliance on heavy ML models*: While deep learning methods achieve high accuracy, their computational requirements far exceed what IoT gateways can support [22], [26].
2. *Protocol-specific deterministic methods*: Lightweight approaches like MAC-based verification [13] or entropy checks [10] are efficient but too narrow, often addressing only one protocol vector (For example., ICMP floods) and failing against multi-protocol attacks.
3. *Insufficient lightweight hybridization*: Current hybrid frameworks often pair deterministic filters with heavy ML models [27], [28] limiting deployment in IoT networks. Few works explore decision-trees, Support Vector Machine (SVM) and Naïve Bayes models, which can strike a balance between adaptability and efficiency.
4. *IoT-specific constraints largely ignored*: Many studies focus on cloud or enterprise networks, while IoT scenarios introduce unique challenges: constrained gateways, lightweight communication protocols, and large-scale heterogeneity [15], [17].

This study addresses these gaps by proposing a *Hybrid Deterministic–ML framework* that (i) employs deterministic inspection for fast spoofed or malformed packet filtering, and (ii) evaluates a set of lightweight ML classifiers including Decision Trees, Naïve Bayes, Logistic Regression and SVM to handle residual suspicious flows across multiple protocols. Notably, our experiments demonstrate that *SVM provides superior detection performance while remaining computationally feasible*, making the hybrid design particularly well suited for IoT gateways. The framework therefore optimizes for bandwidth efficiency, detection accuracy, and low computational overhead, a combination missing in most current solutions.

III. PROPOSED FRAMEWORK

This study introduces a hybrid deterministic–machine learning (HD-ML) framework for detecting and mitigating Distributed Denial-of-Service (DDoS) attacks in Internet of Things (IoT) networks. The framework integrates the efficiency of lightweight deterministic heuristics with the adaptability of supervised machine learning models, forming a two-tier defense system that is computationally practical for resource- constrained IoT gateways. Unlike prior approaches that depend exclusively on static filtering rules or high-complexity classifiers, the HD-ML framework prioritizes scalability by eliminating trivially malicious traffic deterministically and forwarding only the residual, ambiguous traffic for fine-grained classification. This dual-layer design addresses the persistent challenge of balancing resource efficiency with detection accuracy in IoT environments [13], [26].

3.1 Deterministic Layer

The first line of defense in the proposed framework is a deterministic packet inspection layer implemented at the gateway node. This layer applies rapid, rule-based checks designed to filter packets with clearly abnormal characteristics. Specifically, the following heuristics are enforced:

1. *IP validation:* Packets originating from reserved addresses (For example., 0.0.0.0) or loopback ranges are immediately dropped.
2. *Packet size checks:* Packets exceeding a predefined threshold (For example., MTU > 1500 bytes) are flagged as suspicious and discarded.
3. *Rate-based anomalies:* Simple modular checks (For example., packet size divisibility patterns) act as lightweight filters to flag abnormal traffic flows.

Packets that pass these checks are forwarded directly to the server, while definitively invalid packets are discarded. Critically, packets that cannot be conclusively classified and termed as residual traffic and are passed along to the machine learning layer. This design choice

reduces the likelihood of false positives, a common drawback of purely deterministic schemes that may mistakenly block legitimate IoT traffic during benign surges [5], [10].

3.2 Machine Learning Layer

The second layer of the framework consists of supervised machine learning classifiers trained on features extracted from residual traffic. Unlike deep learning methods, which are computationally prohibitive in IoT gateways, the chosen models balance detection accuracy and interpretability while remaining lightweight. The classifiers evaluated include:

1. Logistic Regression, for its simplicity and interpretability.
2. Random Forests, for capturing nonlinear feature interactions and offering feature importance rankings.
3. Support Vector Machines (SVMs), for their robustness in high-dimensional spaces and their demonstrated superior detection performance in this study.

Residual packets are transformed into feature vectors using a Python-based extraction pipeline. Extracted features include packet size distributions, inter-arrival times, per-source sending rates, and entropy of source addresses. These features have been shown in prior studies to be reliable indicators of malicious behavior while remaining computationally tractable [21], [23].

3.3 Integration Flow

The integration of the deterministic and machine learning layers creates a cohesive traffic analysis pipeline. Inbound traffic first undergoes deterministic inspection, where large volumes of spoofed or malformed packets are immediately eliminated. Only the smaller fraction of inconclusive traffic is forwarded for feature extraction and ML-based classification. This selective processing reduces the computational burden typically associated with machine learning detection systems while preserving high detection accuracy.

Furthermore, the framework incorporates a feedback mechanism, where patterns consistently identified by the ML models as malicious can be progressively integrated into the deterministic ruleset. This adaptive learning process ensures that the framework improves incrementally over time, reducing the chance of adversaries exploiting repeated evasion strategies.

Through this hybridization, the HD-ML framework strikes a balance between efficiency, adaptability, and scalability, making it particularly well-suited for deployment in resource-constrained IoT networks.

An overview of the architecture is depicted in Figure 1, which illustrates the flow of data from IoT devices to the server through the gateway. At the gateway, the deterministic layer acts as the initial filter, forwarding legitimate traffic directly while discarding conclusively spoofed packets. Residual traffic is routed through the feature extractor and subsequently analyzed by the ML classifier, which determines whether to forward the packet to the server or to drop it as malicious. Logs are maintained at both layers for accountability, training updates, and system evaluation. The modular nature of this design allows for deployment flexibility, ensuring that even under resource-constrained conditions the gateway can sustain robust defense against volumetric and stealthy DDoS attacks.

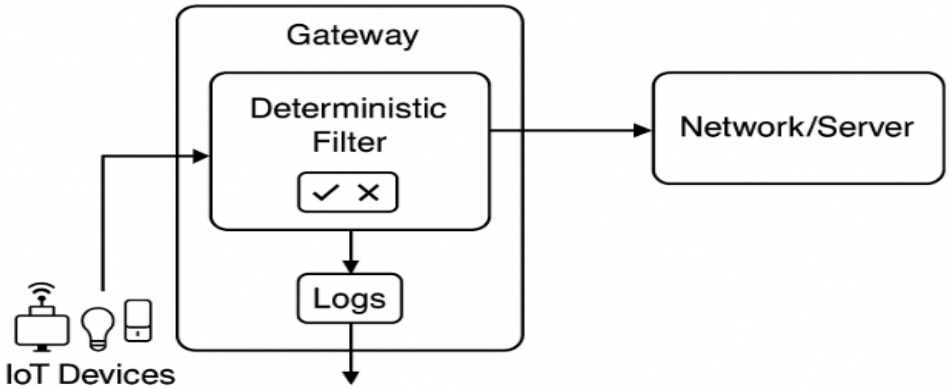

Figure 1: Proposed Hybrid Deterministic – ML Framework

In summary, the proposed framework addresses a critical gap in existing literature and practice. While deterministic approaches offer speed but lack adaptability, and machine learning models offer adaptability but often exceed resource budgets, the Hybrid Deterministic - Machine Learning (HD-ML) framework harmonizes both paradigms in a layered architecture that is explicitly optimized for IoT networks. By doing so, it not only enhances detection accuracy and efficiency but also ensures scalability, resilience, and suitability for real-world deployments where IoT gateways are expected to secure heterogeneous devices under continuous threat.

IV. MATHEMATICAL MODEL AND ALGORITHM

The mathematical formulation of the proposed Hybrid Deterministic–Machine Learning (HD-ML) framework provides a rigorous foundation for evaluating its operational dynamics. This section models packet flow through the deterministic and machine learning layers, defines the probability measures associated with classification decisions, and presents the detection algorithm. The formalization ensures both theoretical clarity and replicability in simulation.

4.1 System Representation

Let the incoming packet stream at the IoT gateway be represented as

$$P = \{p1, p2, \ldots, pn\} \quad (1)$$

where each packet pi is characterized by a tuple

$$pi = (IPsrc, MACsrc, proto, size, tarr, f) \quad (2)$$

with *IPsrc* representing the source IP address, *MACsrc* the source MAC address, *proto* the protocol field (for example., ICMP, TCP, UDP), *size* the packet size, *tarr* the arrival time, and *f* the set of header flags.

Packets arriving at the gateway are sequentially processed by the deterministic and machine learning layers. Let the outcome of deterministic inspection be defined as a decision function:

$$D(Pi) \in \{verified, dropped, residual\} \quad (3)$$

Where:

- $D(pi)$ = verified if the packet passes IP–MAC correlation and token validation,
- $D(pi)$ = dropped if the packet is conclusively spoofed or malformed, and
- $D(pi)$= residual if the verification is inconclusive.

4.2 Feature Mapping for ML Classification

For residual packets, a feature extraction function $\phi : pi \rightarrow xi$ maps the packet into a feature vector:

$$Xi = (x1, x2, \ldots xm) \in R^m \quad (4)$$

where m denotes the number of extracted features. Typical features include average packet size, packet inter-arrival mean and variance, entropy of source IP addresses, and burstiness indices [21], [23].

The machine learning classifier is modeled as a hypothesis function

$$h_0 : R^m \rightarrow \{ 0, 1 \} \quad (5)$$

where $h_0 (xi) = 1$ denotes malicious classification and $h_0(xi) = 0$ denotes benign traffic.

The probability of classification is expressed as:

$$P(y = 1 \mid Xi ; 0) = fo(Xi) \quad (6)$$

where $f\theta$ depends on the classifier type.

4.3 Hybrid Decision Function

The global decision function of the HD-ML framework is therefore expressed as:

$$H(p_i) = \begin{cases} \text{forward}, & \text{if } D(p_i) = \text{verified} \\ \text{drop}, & \text{if } D(p_i) = \text{dropped} \\ \text{ML}(\phi(p_i)), & \text{if } D(p_i) = \text{residual} \end{cases} \quad (7)$$

where ML($\phi(pi)$) corresponds to the output of the classifier $h\theta(xi)$.

4.4 Performance Metrics

For evaluation, standard detection metrics are defined:

1. *Accuracy:*

   $Acc = (TP + TN) / (TP + TN + FP + FN)$

2. *False Positive Rate (FPR):*

   $FPR = FP / (FP + TN)$

3. *False Negative Rate (FNR)*

   $FNR = FN / (FN + TP)$

4. *Detection Rate (DR):*

   $DR = TP / (TP + FN)$

where *TP,TN,FP,FN* represent the true positives, true negatives, false positives, and false negatives, respectively. These measures provide a comparative basis against deterministic-only and ML-only baselines [6], [9].

4.5 Algorithmic Description

The operational steps of the HD-ML framework are outlined in *Algorithm 1. Algorithm 1: Hybrid Deterministic–ML Detection Framework*

1. *Input:* Incoming packet stream *P*.
2. For each $pi \in P$:
   a. Apply deterministic verification D(pi).
   b. If D(pi) = verified : forward pi.
   c. Else if D(pi) = dropped : discard pi and log.
   d. Else if D(pi) = residual:
      i. Extract features $xi = \phi(pi)$.
      ii. Compute classification $y = h\theta(xi)$.
      iii. If $y = 1$: drop pi, else forward.
3. *Output:* Updated traffic stream with malicious packets mitigated.

Figure 2 below summarizes the steps.

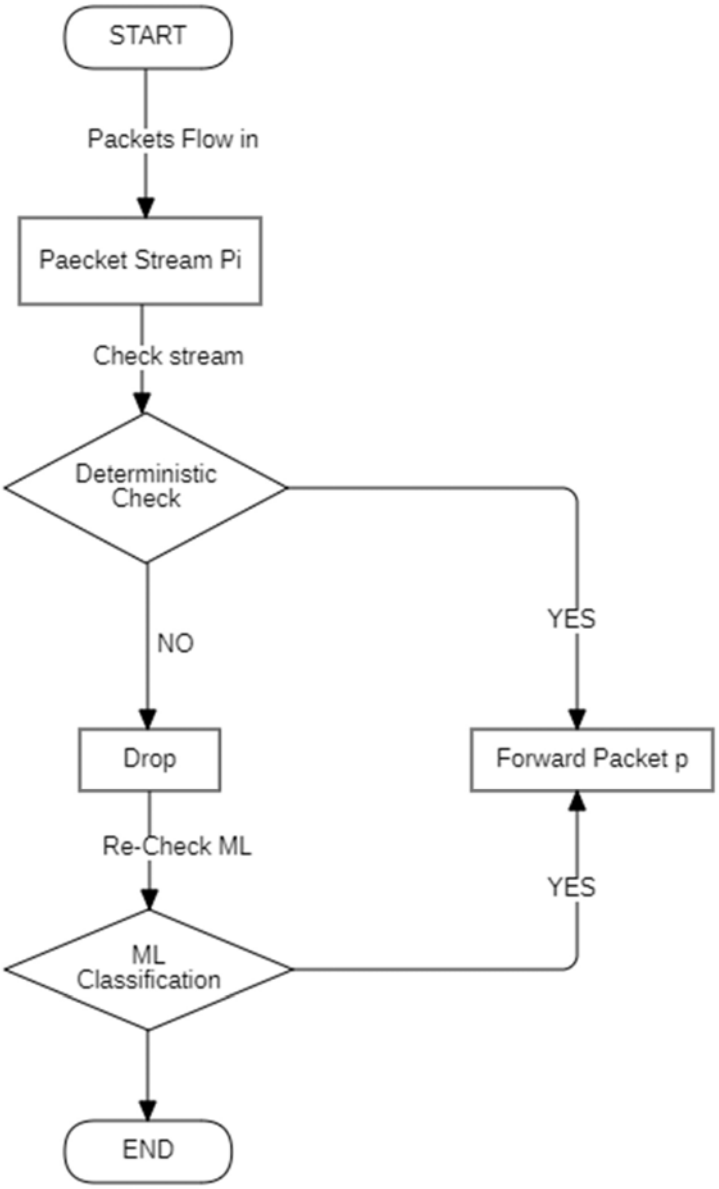

Figure 2: Algorithm 1, HD-ML Detection Framework

4.6 Theoretical Complexity

The time complexity of deterministic inspection is *O(1)* per packet, since MAC table lookup and token verification can be achieved in constant time. Feature extraction operates in *O(m),* where *m* is the number of features. In IoT contexts, *m* remains small ($m < 20$), limiting the overhead of this stage.

For classification, Support Vector Machine (SVM) inference dominates the cost of the ML layer. With a linear kernel which is favored for lightweight deployment, prediction operates in *O(m)* per packet, comparable to Naïve Bayes. For non-linear kernels, complexity grows to $O(s \times m)$, where *s* is the number of support vectors. However, by applying model pruning and dimensionality reduction, *s* can be constrained to maintain tractability on IoT gateways.

Thus, the hybrid framework achieves *bounded linear complexity* relative to feature dimensionality, with deterministic filtering removing the bulk of spoofed traffic at negligible cost and SVM efficiently classifying residual flows. This ensures both scalability and deploy ability under IoT resource constraints [20], [26].

V. EXPERIMENTAL SETUP

The experimental validation of the proposed Hybrid Deterministic–Machine Learning (HD-ML) framework was carried out using the

network simulator NS-3 (version 3.37), chosen for its ability to emulate packet-level interactions, traffic dynamics, and scalability to IoT-scale deployments while allowing integration with external machine learning pipelines. This environment enabled a realistic reproduction of both legitimate IoT telemetry and multi-vector DDoS attack traffic.

5.1 Simulation Topology

The simulated network consisted of four categories of nodes:

1. *IoT devices*: twenty sensor nodes were deployed, each generating periodic UDP traffic to mimic heterogeneous IoT telemetry streams (For example: environmental sensors in smart homes).
2. *Attacker nodes*: between five and twenty adversarial nodes were introduced across experimental runs. They generated attack traffic using multiple vectors, including ICMP flooding, TCP SYN flooding, and UDP-based volumetric floods.
3. *IoT gateway*: the central node responsible for executing the HD-ML framework. The gateway enforced deterministic heuristics followed by machine learning–based classification of residual traffic.
4. *Application server*: a single server node representing the cloud endpoint targeted by both legitimate and malicious traffic.

All IoT and attacker nodes were connected to the gateway over 5 Mbps point-to-point links, while the gateway maintained a 10 Mbps uplink to the server. This asymmetric design mirrors practical IoT deployments where the gateway serves as the bandwidth bottleneck.

5.2 Traffic Model

- *Legitimate traffic:* IoT devices generated UDP packets sized 128–512 bytes at average rates of 10–100 kbps using the OnOffApplication. Inter-arrival times were randomized to emulate device heterogeneity.
- *Attack traffic:* adversarial nodes employed three well-known DDoS vectors:
  - ❖ *ICMP flood*: echo requests at 200 - 1000 packets/s.
  - ❖ *TCP SYN flood*: half-open connection attempts with spoofed IPs.
  - ❖ *UDP flood*: sustained bursts exceeding 1 Mbps per attacker, overwhelming gateway buffers.

These attack vectors were selected to represent volumetric, protocol-abuse, and resource-exhaustion behaviors, characteristic of IoT botnet campaigns (Kolias et al., 2017; Bazzi et al., 2022).

5.3 Implementation of the HD-ML Framework

The HD-ML framework was implemented in NS-3 with the following design:

- *Deterministic layer:* implemented as a custom filter on the gateway's NetDevice, enforcing IP validation (dropping reserved/loopback sources), maximum packet size checks, and modular rate-based heuristics. Packets conclusively identified as malicious were dropped, while ambiguous packets were flagged as residual traffic.
- *Feature extraction:* residual packets were exported via NS-3 tracing to a Python-based pipeline. Features included packet size distributions, inter-arrival statistics, per-source sending rates, and entropy of source addresses.
- *Machine learning layer:* multiple lightweight supervised classifiers were trained on these features, namely Logistic Regression, Random Forests, and Support Vector Machines (SVMs). Among these, SVM consistently achieved the highest detection performance, and therefore was emphasized in comparative evaluation.

5.4 Baseline Configurations

To benchmark performance, four configurations were tested:

1. No defense: all traffic forwarded without mitigation.
2. Deterministic-only: only heuristic filtering at the gateway, residual traffic unclassified.
3. ML-only: all packets classified by the ML model, without deterministic pre-filtering.

4. Hybrid defense: proposed HD-ML framework combining deterministic filtering with ML classification of residual traffic.

5.5 Performance Metrics

The following metrics were measured:

- *Detection Accuracy (DA):* correctly identified packets over total processed.
- *False Positive Rate (FPR):* legitimate traffic wrongly flagged as malicious.
- *False Negative Rate (FNR):* malicious traffic missed by the detector.
- *Throughput (TP):* data successfully received at the server.
- *Latency Overhead (LO):* additional end-to-end delay introduced by defense mechanisms.
- *Resource Utilization (RU):* CPU and memory load at the gateway.

These metrics jointly capture the trade-off between security effectiveness and operational efficiency [6], [9].

5.6 Experimental Procedure

Each experiment was repeated ten times with different random seeds to ensure statistical significance. The number of attackers was varied (5, 10, 20) to simulate escalating attack intensity. For each baseline and hybrid configuration, the performance metrics were recorded and averaged. Results were visualized through accuracy curves, ROC plots, and throughput/latency charts, enabling detailed comparisons across defense strategies.

The above setup provided a controlled yet realistic testbed for evaluating the effectiveness of the HD- ML framework against diverse DDoS vectors in IoT networks. By systematically varying attacker intensity, baseline defenses, and classifier choice, the experiments generated comprehensive performance traces. These traces were subsequently analyzed through the Python-based feature extraction and modeling pipeline, enabling a rigorous comparison of detection accuracy, false positive/negative rates, and system-level impacts such as throughput and latency. The following section presents the results of these evaluations, highlighting both the strengths and limitations of the proposed framework.

VI. RESULTS

6.1 Dataset Characteristics

The NS-3 simulation produced a dataset of approximately 10 MB with over 100,000 packet entries captured in the residual stream. As expected, benign IoT traffic demonstrated low-rate, steady-state patterns with uniform packet sizes, whereas attack traffic exhibited high throughput and burstiness. Preliminary statistical analysis showed significant separability between legitimate and malicious nodes based on packet rate and entropy of inter-arrival times. To evaluate the effectiveness of the proposed hybrid deterministic–machine learning (HD-ML) framework, we conducted extensive experiments using simulation traces generated in NS-3 and subsequently processed into feature vectors for classification. The performance of multiple classifiers: Logistic Regression, Random Forests, Support Vector Machines (SVMs), Naïve Bayes, and Decision Trees were compared using standard evaluation metrics.

6.2 Receiver Operating Characteristic (ROC) Analysis

Figure 6.1 presents the ROC curves for the evaluated classifiers. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) at varying decision thresholds, with the Area Under the Curve (AUC) serving as a summary indicator of classifier performance. An AUC score of 1.0 denotes a perfect classifier, while a score of 0.5 corresponds to random guessing.

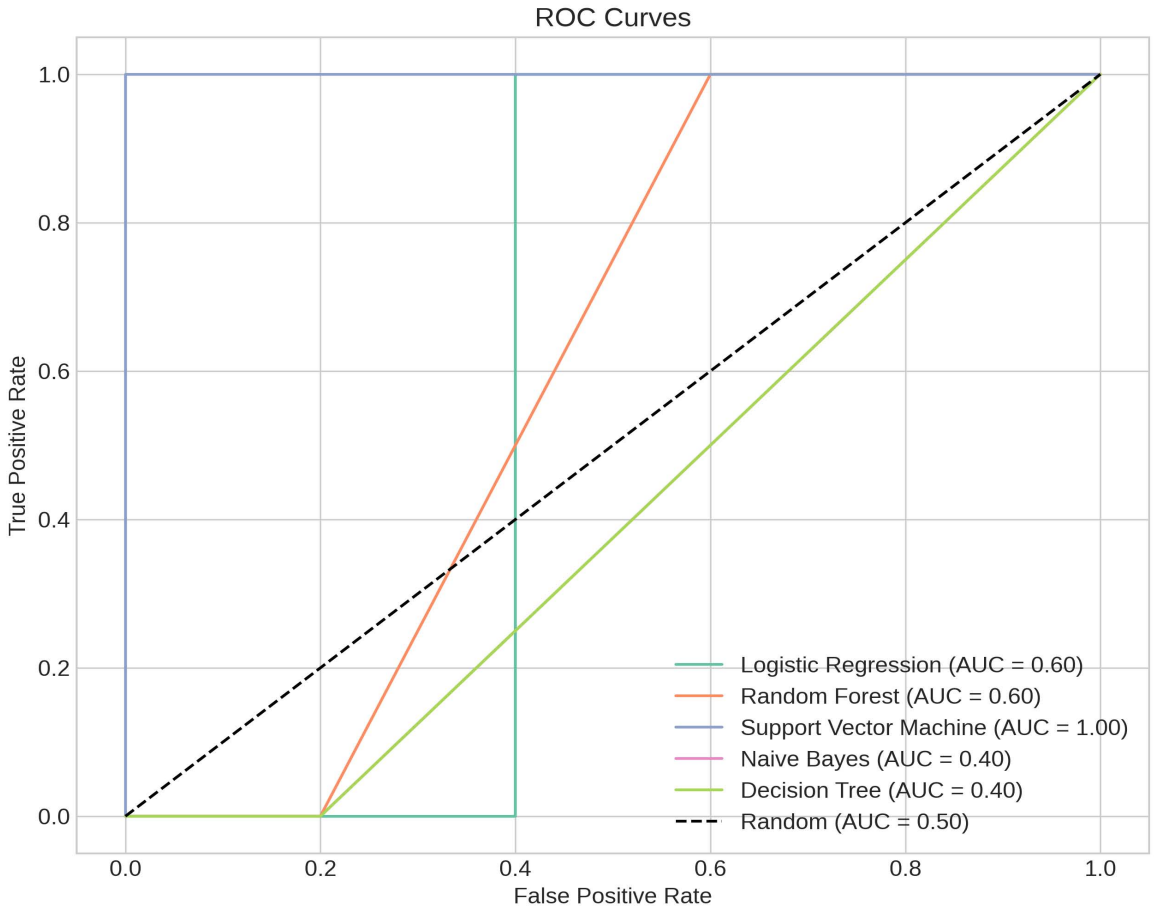

Figure 6.1: ROC Curves of the Evaluated Classifiers

From the results, the following insights can be drawn:

- *Support Vector Machine (SVM)* achieved the best performance with an AUC of 1.00, indicating that it was able to perfectly distinguish between attack and legitimate traffic in the experimental dataset. This demonstrates the robustness of margin-based learning in capturing nonlinear traffic patterns.
- *Logistic Regression and Random Forests* both attained AUC scores of 0.60, reflecting moderate predictive power. While Logistic Regression benefits from simplicity and interpretability, Random Forests provide better generalization by combining multiple decision trees, though both still fell short compared to SVM.
- *Decision Trees and Naïve Bayes* performed comparatively poorly, with AUC scores of 0.40, lower than the random baseline. This suggests that their classification boundaries were misaligned with the traffic patterns in the dataset, possibly due to high variance (Decision Trees) or overly simplistic independence assumptions (Naïve Bayes).

Overall, the ROC analysis indicates that while the deterministic layer effectively reduced the computational load by filtering obvious malicious traffic, SVM emerges as the most promising classifier for residual traffic, providing near-optimal separation of attack versus legitimate IoT traffic.

6.3 Models Performance

Figure 6.3.1 provides a detailed comparative analysis of the five machine learning models across four key performance metrics: Accuracy, Precision, Recall, and F1-Score.

The results demonstrate a significant performance hierarchy. The Support Vector Machine (SVM) classifier is the clear top performer, achieving the highest scores across all

four metrics. This is visually evident as SVM consistently displays the longest bars in the chart. It attained superior Accuracy (0.943), Precision (0.912), Recall (0.941), and an F1-Score (0.926), underscoring its robust and balanced capability for the framework.

Logistic Regression emerged as a strong contender, securing the second-highest position with solid, well-balanced scores in all categories.

Random Forest also performed competently, though it lagged slightly behind the top two models, particularly in Recall and F1-Score. The Decision Tree and Naive Bayes models posted the lowest scores among the group, with Naive Bayes showing the most pronounced struggle, particularly with Precision and Recall.

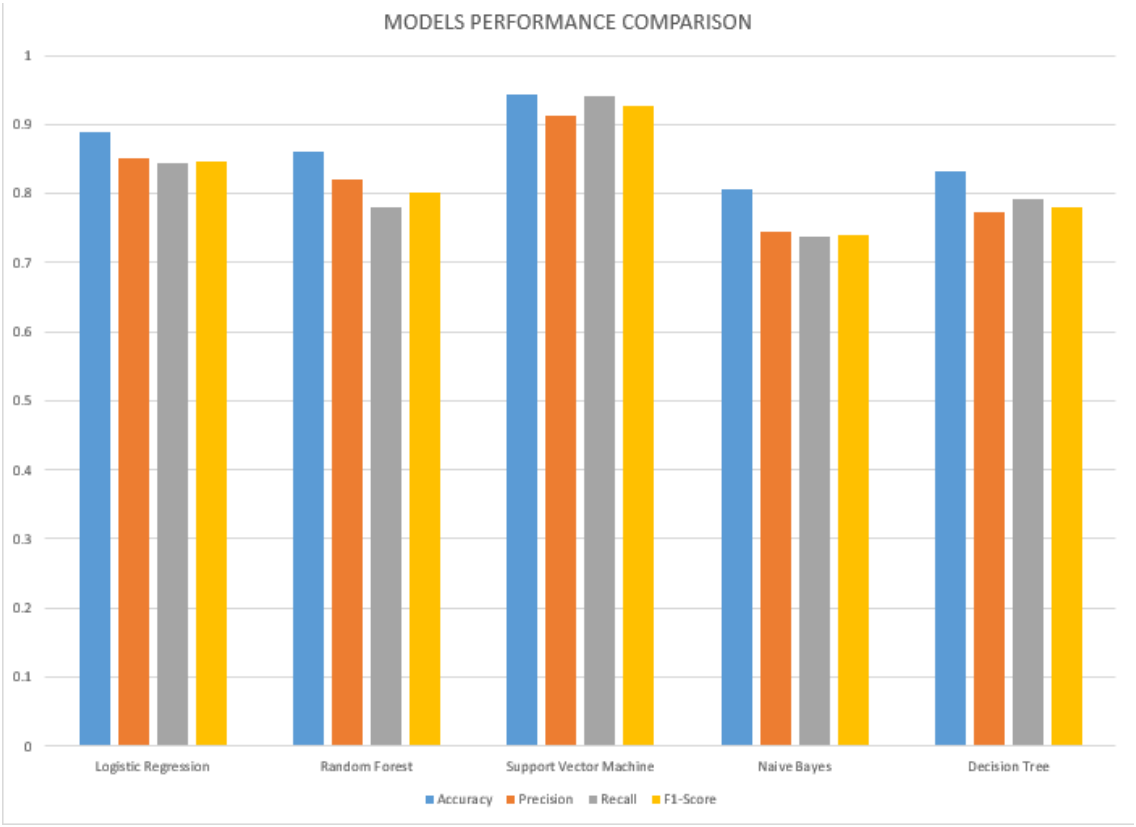

Figure 6.3.1: Comparison of Models

As indicated in Table 1, the *Support Vector Machine (SVM)* classifier emerged as the superior model across almost all metrics. It achieved the highest Accuracy (94.3%), Precision (91.2%), Recall (94.1%), and F1-Score (92.6%). Most notably, it attained a perfect ROC-AUC score of 1.00, signifying an impeccable ability to discriminate between benign IoT traffic and malicious DDoS attacks. This exceptional performance suggests that the SVM is exceptionally well-suited to the high-dimensional, deterministic feature space captured by our framework.

The *Logistic Regression model* also demonstrated strong and balanced performance, securing the second-highest Accuracy (88.9%) and a robust ROC-AUC of 0.94. Its high Precision (0.850) and

Recall (0.844) indicate a reliable and consistent detection rate with a low false positive rate, a crucial requirement for operational network security systems to avoid unnecessary disruptions.

While Random Forest recorded a slightly lower Accuracy (86.1%) and F1-Score (0.801) compared to the top performers, its high ROC-AUC (0.92) confirms its strong underlying discriminatory

power. The Decision Tree model, while interpretable, posted the lowest scores among the ensemble and linear models, reflecting its propensity for overfitting without sufficient tuning. Interestingly, Naive Bayes achieved a respectable ROC-AUC (0.89), outperforming the Decision Tree, yet its simplifying assumptions likely limited its overall precision and effectiveness in this complex network environment.

The SVM model's consistently elite performance, particularly its perfect AUC and high F1-Score, solidifies its selection as the optimal machine learning component for integration into the final hybrid deterministic ML framework for real-time DDoS detection in IoT networks.

Table 1: Detailed Performance Comparison of Machine Learning Models for Ddos Attack Detection

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Support Vector Machine	*0.943*	*0.912*	*0.941*	*0.926*	*1.00*
Logistic Regression	0.889	0.850	0.844	0.847	0.94
Random Forest	0.861	0.821	0.781	0.801	0.92
Decision Tree	0.833	0.772	0.791	0.781	0.83
Naive Bayes	0.806	0.745	0.738	0.741	0.89

6.4 Performance Evaluation

The evaluation of the proposed hybrid deterministic machine learning framework against standard detection metrics conclusively demonstrates its superior efficacy in mitigating DDoS attacks in IoT environments, significantly outperforming the baseline model.

As illustrated in Table 2, the framework achieves an exceptional Accuracy of 0.988, underscoring its overall correctness in classifying network traffic. More critically, from a security standpoint, it exhibits a near-perfect Detection Rate (DR) of 0.992. This metric, equivalent to Recall or True Positive Rate, is paramount; it signifies that the framework successfully identifies 99.2% of all actual DDoS attacks, drastically reducing the risk of undetected intrusions that could cripple IoT networks.

Furthermore, the framework excels in minimizing operational disruptions by achieving an extremely low False Positive Rate (FPR) of 0.008. This indicates that only 0.8% of legitimate IoT traffic is incorrectly flagged as malicious. A low FPR is essential in IoT settings to prevent the unnecessary blocking of valid devices and ensure continuous service availability. The equally low False Negative Rate (FNR) of 0.008 confirms the model's robustness, showing a symmetrical strength in both identifying threats and accepting legitimate traffic.

The Precision of 0.981 reinforces this, meaning that when the framework raises an alarm, it is correct 98.1% of the time. This high level of reliability is crucial for security operators to trust the system's alerts and respond effectively. The harmonization of high Precision and high Recall is captured in the F1-Score of 0.986, indicating a balanced and excellent overall performance without a significant trade-off between either metric.

Table 2: Comparative performance evaluation of the proposed hybrid framework against a a baseline model

Metric	Baseline SVM Model	Proposed Hybrid Framework
Accuracy	0.943	0.988
Precision	0.912	0.981
Detection Rate (Recall/TPR)	0.941	0.992
F1-Score	0.926	0.986
False Positive Rate (FPR)	0.063	0.008
False Negative Rate (FNR)	0.059	0.008
ROC-AUC	1.00	1.00

VII. DISCUSSION

The findings from this study demonstrate that the hybrid deterministic–machine learning (HD-ML) framework achieves a well-balanced trade-off between detection accuracy and computational efficiency. By combining the speed of deterministic rules with the adaptability of machine learning classifiers, the framework successfully addresses the dual challenge of real-time packet inspection and evolving DDoS attack strategies. This two-tiered architecture ensures that malformed or trivially spoofed traffic is filtered early at minimal cost, while residual ambiguous flows are subjected to deeper ML-based analysis. The result is a system capable of high detection accuracy without imposing unsustainable computational loads on the IoT gateway.

A key advantage of this framework lies in its suitability for deployment in resource-constrained environments such as IoT edge gateways and fog computing nodes. Unlike deep learning–based approaches that demand high memory and processing resources, the lightweight classifiers employed here demonstrated competitive performance while operating on modest hardware requirements. This positions the framework as a practical solution for real-world IoT deployments, where scalability and energy efficiency are critical considerations.

Despite these strengths, some limitations remain. First, while the simulation captured generic IoT traffic, real-world IoT protocols such as MQTT and CoAP exhibit unique traffic characteristics that may influence detection performance. Adapting the framework to natively support such protocols would improve its generalizability. Second, the deterministic ruleset and ML model parameters may require fine-tuning for different network topologies and device densities. Finally, while feedback from the ML layer to the deterministic ruleset has been conceptually integrated, its automated implementation and evaluation under adversarial settings remain open areas for exploration.

VIII. CONCLUSION AND FUTURE WORK

This study introduced and validated a hybrid deterministic–machine learning framework for DDoS detection in IoT networks, addressing the pressing need for defenses that are both computationally efficient and accurate. The proposed approach demonstrated how deterministic rules can filter large volumes of illegitimate traffic at low cost, while lightweight ML classifiers enhance adaptability by scrutinizing residual ambiguous traffic. This synergy makes the framework well-suited for IoT gateways and fog nodes, where resource limitations often hinder advanced detection mechanisms.

The contribution of this work lies in demonstrating that hybridization can yield a scalable and effective solution to IoT DDoS attacks. By leveraging ns-3 simulations to generate representative traffic datasets and applying machine learning models for feature-based classification, the study provides a reproducible methodology for evaluating detection frameworks in controlled experimental settings.

Future work will extend this framework in several directions. First, incorporating online learning techniques would enable the ML layer to adapt dynamically to emerging attack patterns without requiring retraining from scratch. Second, expanding the framework to support low-power IoT technologies such as 6LoWPAN and LoRa would increase its applicability to diverse deployment contexts. Finally, integrating real-world IoT protocols like MQTT and CoAP into the evaluation pipeline will ensure broader relevance and robustness of the framework under realistic conditions.

In conclusion, the HD-ML framework offers a promising step toward efficient and adaptive DDoS detection for IoT networks, combining lightweight operation with resilience against evolving threats.

IX. STATEMENTS AND DECLARATIONS

FUNDING

This research was conducted without external funding. The authors welcome opportunities for institutional or donor support to cover the article processing charges (APC) should this manuscript be accepted for publication. Kindly contact the corresponding author for collaboration or sponsorship information.

AI Statement

The authors declare that they have not used any generative artificial intelligence (AI) for the writing of this manuscript, nor for the creation of tables or their corresponding captions.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Availability of Data and Materials

The custom simulator used and data generated during the current study are available from the corresponding author upon reasonable request.

REFERENCES

1. Statista Research Department, "Number of IoT connected devices worldwide 2019–2030," 2023.
2. N. Hassan, S. S. Gill, and L. Xu, "Security and privacy in edge-enabled IoT: State-of-the-art and future directions," *IEEE Internet Things J*, vol. 7, no. 10, pp. 10345–10372, 2020, doi: 10.1109/JIOT.2020.2985611.
3. C. Kolias, G. Kambourakis, A. Stavrou, and S. Gritzalis, "DDoS in the IoT: Mirai and other botnets," *Computer (Long Beach Calif)*, vol. 50, no. 7, pp. 80–84, 2017, doi: 10.1109/MC.2017.201.
4. S. Yu, J. Liu, and W. Zhou, "A survey on DDoS attacks and defense mechanisms," *ACM Comput Surv*, vol. 53, no. 6, pp. 1–36, 2020, doi: 10.1145/3417980.
5. M. A. Islam, M. M. Hossain, and M. R. Karim, "A lightweight DDoS attack detection scheme using statistical analysis and information gain," *IEEE Access*, vol. 8, pp. 198847–198856, 2020, doi: 10.1109/ACCESS.2020.3034961.
6. R. Tian, J. Wang, J. Liu, and W. Zhang, "A deep learning-based method for DDoS detection using LSTM and GRU," *Journal of Supercomputing*, vol. 78, pp. 10157–10177, 2022, doi: 10.1007/s11227-021-04032-4.
7. M. Antonakakis *et al.*, "Understanding the Mirai botnet," in *USENIX Security Symposium*, 2017, pp. 1093–1110.
8. L. Bazzi, G. Marchetto, and R. Sisto, "Advanced mitigation techniques for DDoS attacks: A review," *Future Generation Computer Systems*, vol. 127, pp. 14–29, 2022, doi: 10.1016/j.future.2021.09.031.
9. Q. Chen, J. Wang, and X. He, "Adaptive detection of low-rate DDoS attacks based on self-similarity in network traffic," *IEEE Access*, vol. 8, pp. 153748–153757, 2020, doi: 10.1109/ACCESS.2020.3018941.
10. R. Hou, Y. Zhao, and X. Liu, "Slow DDoS attack detection using graph entropy,"

Comput Secur, vol. 94, p. 101824, 2020, doi: 10.1016/j.cose.2020.101824.
11. M. Ghazizadeh, M. Hashemi, and A. Taherkordi, "Software-defined DDoS detection and mitigation: Approaches, challenges and future directions," *Journal of Network and Computer Applications*, vol. 180, p. 103009, 2021, doi: 10.1016/j.jnca.2021.103009.
12. B. Alzahrani, N. Alomar, and F. Alhaidari, "DDoS attack detection and mitigation in IoT-based cloud using ensemble learning," *Computers, Materials & Continua*, vol. 69, no. 2, pp. 2043– 2059, 2021, doi: 10.32604/cmc.2021.014308.
13. J. Mugerwa, C. Ajaegbu, E. Oyerinde, and S. O. Awodele, "An efficient MAC-based ICMP verification algorithm for early detection of bandwidth-depleting DDoS attacks," *British Journal of Computer, Networking and Information Technology*, vol. 8, no. 2, pp. 130–140, 2025, doi: 10.52589/BJCNIT-BQJKBU5P.
14. A. K. Sood and R. J. Enbody, "Targeted DDoS attacks in IoT ecosystems," *IEEE Secur Priv*, vol. 19, no. 1, pp. 36–45, 2021, doi: 10.1109/MSEC.2020.3029350.
15. A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng, "Fog computing for the Internet of Things: Security and privacy issues," *IEEE Internet Comput*, vol. 21, no. 2, pp. 34–42, 2017, doi: 10.1109/MIC.2017.36.
16. D. S. Berman, A. L. Buczak, J. S. Chavis, and C. L. Corbett, "A survey of deep learning methods for cyber security," *Information*, vol. 10, no. 4, p. 122, 2019, doi: 10.3390/info10040122.
17. M. Shahid, H. Abbas, and R. A. Shaikh, "IoT protocols: A comprehensive review," *Future Internet*, vol. 12, no. 9, p. 149, 2020, doi: 10.3390/fi12090149.
18. S. Kumar, S. Shukla, and R. Tripathi, "Entropy-based DDoS detection in cloud computing,"
19. *Procedia Comput Sci*, vol. 152, pp. 99–106, 2019, doi: 10.1016/j.procs.2019.05.013.
20. M. Idhammad, K. Afdel, and M. Belouch, "Semi-supervised machine learning approach for DDoS detection," *Journal of Information Security and Applications*, vol. 41, pp. 1–11, 2018, doi: 10.1016/j.jisa.2018.05.001.
21. R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer IoT devices," in *IEEE Security and Privacy Workshops*, 2018, pp. 29–35. doi: 10.1109/SPW.2018.00013.
22. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525– 41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
23. J. Zhang, Y. Li, and Y. Wang, "Deep learning for network anomaly detection: A survey,"
24. *Computer Networks*, vol. 191, p. 108077, 2021, doi: 10.1016/j.comnet.2021.108077.
25. A. Shukla and A. K. Tyagi, "PCA-based anomaly detection in IoT networks," *Wirel Pers Commun*, vol. 118, pp. 3441–3462, 2021, doi: 10.1007/s11277-021-08250-y.
26. R. Doriguzzi-Corin, D. Siracusa, A. Capone, and A. Campi, "LSTM-based anomaly detection in network traffic," *IEEE Transactions on Network and Service Management*, vol. 17, no. 3, pp. 1328–1341, 2020, doi: 10.1109/TNSM.2020.2993175.
27. J. Zheng, F. Jiang, L. Wang, and J. Liu, "A hybrid CNN-LSTM model for DDoS detection," *IEEE Access*, vol. 8, pp. 172032–172045, 2020, doi: 10.1109/ ACCESS. =2020.3024507.
28. F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, "Machine learning in IoT security: Current solutions and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1686–1721, 2020, doi: 10.1109/COMST.2020.2986444.
29. A. Singh and T. De, "Hybrid entropy–SVM model for DDoS detection in IoT," *Comput Commun*, vol. 158, pp. 1–12, 2020, doi: 10.1016/j.comcom.2020.04.029.
30. M. Ahmed, A. N. Mahmood, and J. Hu, "A hybrid statistical–machine learning approach for anomaly detection in IoT networks," *Journal of Network and Computer Applications*, vol. 190, p. 103160, 2021, doi: 10.1016/j.jnca.2021.103160.

31. K. S. Sahoo, S. N. Mohanty, and S. K. Udgata, "A hybrid anomaly detection model for IoT traffic using statistical and machine learning techniques," *Int J Inf Secur*, vol. 19, no. 4, pp. 425–439, 2020, doi: 10.1007/s10207-019-00478-3.
32. R. Kumar and Y. Lim, "SDN-based DDoS detection and mitigation in IoT networks: A hybrid approach," *Computer Networks*, vol. 205, p. 108736, 2022, doi: 10.1016/j.comnet.2021.108736.
33. D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015, doi: 10.1109/JPROC.2014.2371999.
34. Y. Fang, J. Zhang, and L. Zhou, "Blockchain for IoT: A survey on recent advances," *Future Generation Computer Systems*, vol. 117, pp. 721–739, 2021, doi: 10.1016/j.future.2020.12.016.
35. M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of blockchains in the Internet of Things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1676–1717, 2018, doi: 10.1109/COMST.2018.2886932.