



AI-Driven Intrusion Detection in Internet of Things Networks: A Lightweight and Adaptive Framework for Next-Generation Cybersecurity

Dr. Arjun Deshmukh¹

Research Scholar¹

Department of Computer Science and Engineering

Western Institute of Technology, Ahmedabad, Gujarat, India

arjun.deshmukh@witindia.in

Abstract

The Internet of Things (IoT) has emerged as a fundamental component of modern digital infrastructure, connecting sensors, smart appliances, industrial controllers, healthcare equipment, vehicles, surveillance systems, and numerous other physical objects to communication networks and intelligent computing platforms. Although IoT connectivity provides significant advantages in automation, monitoring, resource optimization, and real-time decision-making, it also introduces a complex and continuously expanding cybersecurity environment. The heterogeneity of IoT devices, limited computational resources, diverse communication protocols, inadequate security configurations, and large-scale deployment make conventional security mechanisms increasingly difficult to apply effectively.

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), provides an opportunity to improve IoT security by learning behavioral patterns from network traffic and identifying malicious or anomalous activities. Existing research has investigated decision trees, random forests, support vector machines, neural networks, convolutional neural networks, recurrent neural networks, autoencoders, federated learning, and hybrid architectures for intrusion detection. However, high classification performance in laboratory environments does not necessarily guarantee successful deployment on resource-constrained IoT and edge devices. Recent research continues to identify computational complexity, energy consumption, latency, dataset limitations, adaptability, and inconsistent deployment evaluation as important challenges. ([MDPI](#))

This paper presents a comprehensive examination of AI-driven intrusion detection for IoT environments and proposes a conceptual lightweight and adaptive framework. The proposed architecture distributes security intelligence across IoT devices, network monitoring components, edge gateways, and centralized security infrastructure. Lightweight anomaly screening is performed near the data source, while computationally intensive analysis is reserved for suspicious or uncertain events. The framework further incorporates adaptive learning, explainable detection, privacy-aware data processing, and a feedback mechanism for continuous improvement. The paper analyzes major IoT threats, conventional IDS techniques, machine learning approaches, deep learning architectures, edge intelligence, federated learning, evaluation metrics, deployment limitations, and future research opportunities. Rather than optimizing only classification accuracy, the proposed framework considers detection effectiveness, latency, resource consumption, scalability, privacy, adaptability, and interpretability. The study provides a foundation for future experimental research into practical AI-enabled intrusion detection systems capable of protecting heterogeneous and evolving IoT environments.

Keywords: Internet of Things, Artificial Intelligence, Machine Learning, Intrusion Detection, Cybersecurity, Deep Learning, Edge Computing, Anomaly Detection, Federated Learning, IoT Security

1. Introduction

The rapid growth of connected devices has fundamentally changed the way digital systems interact with the physical world. The Internet of Things enables physical objects to sense their surroundings, communicate with other devices, exchange information, and interact with software platforms. IoT technology is now associated with smart homes, industrial automation, healthcare, transportation, agriculture, energy management, surveillance, environmental monitoring, and smart-city infrastructure.

A typical IoT ecosystem may contain thousands of heterogeneous devices operating simultaneously. These devices differ considerably in hardware capability, operating system, firmware, communication protocol, power availability, and security configuration. Some devices possess substantial processing capability, whereas others operate with extremely limited memory, storage, and energy resources.

This heterogeneity creates a significant cybersecurity challenge.

Traditional enterprise cybersecurity architectures were largely designed around relatively stable computing devices and clearly defined network boundaries. IoT environments are considerably more dynamic. Devices may be physically accessible, intermittently connected, remotely managed, automatically updated, or operated for long periods without direct human interaction.

Furthermore, IoT devices can become attractive targets for attackers because compromising a large number of relatively weak devices may provide substantial computing, network, or operational capabilities.

The security problem becomes particularly serious when IoT devices participate in cyber-physical processes. An attack against a conventional workstation may result in data loss or service disruption. In contrast, compromising an industrial controller, medical device, smart-grid component, or transportation system may produce physical consequences.

Consequently, IoT cybersecurity requires mechanisms capable of identifying malicious behavior before it develops into a major security incident.

1.1 Role of Artificial Intelligence

Artificial intelligence provides a promising approach because it allows security systems to identify patterns within large volumes of network data.

A conventional rule-based IDS may identify an attack only when a previously defined rule or signature matches the observed behavior. An AI-based IDS can instead learn statistical or behavioral relationships from historical observations.

The fundamental concept can be represented as:

$$f: X \rightarrow Y \text{ if } X \rightarrow Y$$

where XX represents observed network characteristics and YY represents the predicted security state.

The security state may be represented as:

$$Y \in \{\text{Normal}, \text{Suspicious}, \text{Malicious}\} \text{ if } Y \in \{\text{Normal}, \text{Suspicious}, \text{Malicious}\}$$

This approach enables the IDS to examine behavioral characteristics rather than relying exclusively on predefined signatures.

Recent reviews of IoT cybersecurity research show continued interest in ML, DL, federated learning, explainable AI, and edge-based detection. ([MDPI](#))

1.2 Research Motivation

The motivation for this study is based on the gap between **AI model performance** and **real-world IoT deployment requirements**.

A model may achieve excellent classification performance using a benchmark dataset but still be unsuitable for a constrained edge environment if:

- It requires excessive memory.
- Inference takes too long.
- Model updates require substantial bandwidth.
- Energy consumption is excessive.
- It generates too many false alarms.
- It cannot adapt to new attack patterns.
- Its decisions cannot be interpreted.

Therefore, a practical IoT IDS should be evaluated as a complete system rather than simply as a classification algorithm.

1.3 Problem Statement

The central research problem addressed by this paper is:

How can an AI-based intrusion detection system provide accurate, adaptive, explainable, and low-latency threat detection in heterogeneous IoT networks while maintaining acceptable computational, memory, energy, and communication requirements?

1.4 Research Objectives

The objectives of this study are:

1. To examine major cybersecurity threats affecting IoT environments.
2. To review conventional intrusion detection mechanisms.

3. To analyze machine-learning and deep-learning approaches for IoT intrusion detection.
4. To investigate the importance of edge computing for low-latency security.
5. To identify limitations in existing AI-based IoT IDS research.
6. To propose a lightweight and adaptive IDS architecture.
7. To examine privacy-preserving approaches such as federated learning.
8. To identify suitable evaluation metrics beyond classification accuracy.
9. To discuss future directions for practical AI-based IoT cybersecurity.

1.5 Contributions

The major contributions of this paper are:

- A structured analysis of the IoT cybersecurity environment.
- A review of AI-based intrusion detection techniques.
- A conceptual multi-layer architecture for lightweight IoT intrusion detection.
- Integration of edge intelligence into the proposed architecture.
- An adaptive learning mechanism for evolving attack behavior.
- Consideration of privacy and explainability.
- A practical evaluation framework covering performance and deployment requirements.

2. Internet of Things Security Architecture

An IoT ecosystem can generally be divided into several functional layers.

Layer 1: Perception Layer

This layer contains physical sensors and actuators.

Examples include:

- Temperature sensors
- Cameras
- Motion sensors
- Smart meters

- Industrial sensors
- Wearable devices

Layer 2: Network Layer

The network layer transports information between devices, gateways, and servers.

Technologies may include:

- Wi-Fi
- Ethernet
- Bluetooth
- Zigbee
- Cellular networks
- LPWAN technologies

Layer 3: Edge/Fog Layer

The edge layer processes information closer to the data source.

This layer is particularly important for intrusion detection because security analysis can occur without sending every event to a distant cloud server.

Layer 4: Cloud Layer

Cloud platforms provide large-scale:

- Storage
- Analytics
- Machine learning
- Security intelligence
- Historical data processing

Layer 5: Application Layer

The application layer provides services to users and organizations.

Examples include:

- Smart-home applications
- Industrial dashboards
- Healthcare applications
- Traffic management systems
- Smart-grid applications

Security must therefore be considered across the entire architecture rather than at only one layer.

3. IoT Cybersecurity Threats

3.1 Denial-of-Service Attacks

Denial-of-Service (DoS) attacks attempt to make a service unavailable.

Attackers may generate excessive traffic or exploit weaknesses that consume device resources.

In an IoT environment, the consequences can be significant because devices may have limited processing and communication capacity.

3.2 Distributed Denial-of-Service

A Distributed Denial-of-Service (DDoS) attack involves multiple compromised devices.

The IoT environment can be particularly attractive for such attacks because large numbers of poorly secured devices can potentially be compromised and coordinated.

A simplified representation is:

$$DDoS = \sum_{i=1}^n Traffic_i$$

where $Traffic_i$ represents malicious traffic generated by compromised device i .

3.3 Network Scanning

Attackers frequently scan networks before launching additional attacks.

Indicators include:

- Numerous connection attempts
- Rapid changes in destination ports
- Sequential address access
- High failure rates

These behavioral characteristics can be learned by ML models.

3.4 Brute-Force Attacks

Repeated authentication attempts may indicate credential attacks.

Relevant features include:

- Number of authentication attempts
- Failed authentication ratio
- Time interval between attempts
- Number of targeted accounts

3.5 Malware and Botnets

IoT malware can convert devices into remotely controlled components of a botnet.

The resulting device may:

- Scan other systems
- Participate in DDoS attacks
- Communicate with command-and-control servers
- Attempt lateral movement
- Generate malicious traffic

3.6 Man-in-the-Middle Attacks

An attacker may attempt to intercept communication between IoT devices and legitimate services.

This may result in:

- Data modification
- Credential theft
- Session manipulation
- Unauthorized commands

3.7 Data Exfiltration

Sensitive information may be transferred to unauthorized external destinations.

Anomaly detection can identify unusual outbound communication patterns.

4. Conventional Intrusion Detection

4.1 Signature-Based IDS

Signature-based systems compare observed activity against known attack signatures.

The primary advantage is efficient identification of known attacks.

However:

Unknown Attack $\notin$ Signature Database $\rightarrow$ Unknown \ Attack \ not in Signature Database

may result in the attack being missed.

Therefore, signature-based detection alone is insufficient for continuously changing IoT environments.

4.2 Anomaly-Based IDS

Anomaly-based systems model normal behavior.

A simple anomaly score may be represented as:

$$A(x) = D(x, \mu)$$

where:

- x is the observed event,
- μ is the learned normal profile,
- D is a distance function.

If:

$$A(x) > \tau$$

the event is considered anomalous.

4.3 Hybrid IDS

Hybrid systems combine signature and anomaly detection.

The advantage is that known threats can be detected rapidly while unusual behavior can be subjected to behavioral analysis.

5. Machine Learning for IoT Security

Machine learning can be divided broadly into supervised, unsupervised, and semi-supervised approaches.

5.1 Supervised Learning

Supervised models require labeled training data.

Examples include:

- Decision Trees
- Random Forest
- SVM
- Logistic Regression
- Neural Networks

The model learns:

$$P(Y|X)P(Y|X)$$

where XX represents network features and YY represents the attack class.

5.2 Unsupervised Learning

Unsupervised learning does not require predefined attack labels.

Examples include:

- k-Means
- DBSCAN
- Autoencoders
- Clustering-based anomaly detection

These techniques can be useful when labeled attack data are limited.

5.3 Semi-Supervised Learning

Semi-supervised learning combines labeled and unlabeled observations.

This can be valuable because obtaining large quantities of accurately labeled IoT attack traffic is difficult.

6. Deep Learning Approaches

6.1 Artificial Neural Networks

A basic neuron can be represented as:

$$z = \sum_{i=1}^n w_i x_i + b$$

followed by:

$$y = f(z)$$

Neural networks can model nonlinear relationships between network features.

6.2 Convolutional Neural Networks

CNNs can identify local patterns in structured representations of network traffic.

6.3 Recurrent Neural Networks

RNNs are suitable for sequential data.

This is important because network attacks may exhibit temporal behavior.

6.4 LSTM

Long Short-Term Memory networks are designed to preserve useful information across longer sequences.

They may therefore be useful for attacks whose behavioral patterns develop over time.

6.5 Autoencoders

Autoencoders learn to reconstruct input data.

The reconstruction error is:

$$E(x) = \|x - \hat{x}\|^2$$

A large reconstruction error can indicate anomalous behavior.

7. Edge Computing for IoT Intrusion Detection

Centralized cloud-based security can introduce communication delays.

The total response time can be approximated as:

$$T = T_c + T_n + T_p + T_r$$

where:

- T_c = data capture time,
- T_n = network transmission time,
- T_p = processing time,
- T_r = response time.

Edge processing primarily reduces the network component T_n .

Recent research emphasizes the importance of edge-aware deployment because IoT-edge systems combine heterogeneous devices, limited resources, and security requirements that may demand rapid response. ([MDPI](#))

8. Proposed Lightweight Adaptive IDS

The proposed architecture consists of five principal components:

IoT Devices



Traffic Monitoring



Edge Detection



Adaptive Security Intelligence



Central Security Management

8.1 IoT Monitoring

Devices generate basic telemetry such as:

- Packet frequency
- Session duration
- Destination information
- Connection failures
- Protocol type
- Authentication events

8.2 Feature Processing

Features are cleaned and normalized before being supplied to the ML model.

A normalized feature can be calculated as:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

8.3 Lightweight Detection

A computationally inexpensive model performs the first screening.

Normal traffic is immediately accepted.

Suspicious traffic is passed to a deeper analysis stage.

8.4 Adaptive Analysis

The system maintains a confidence score:

$$C(x) = \max_j P(Y_j | x)$$

If:

$$C(x) < \theta$$

the observation is considered uncertain and forwarded for advanced analysis.

8.5 Response

Depending on the risk level, the system can:

- Generate an alert.
- Restrict communication.
- Isolate a device.
- Increase monitoring.
- Request administrator confirmation.

9. Federated Learning

Centralized training requires data to be transferred to a common server.

Federated learning provides an alternative.

Each participating edge node trains a local model:

$$w_{k,t+1} = \text{Train}(D_k, w_t) \quad w_k^{t+1} = \text{Train}(D_k, w^t)$$

The central system can then aggregate updates.

A simplified aggregation process is:

$$w_{t+1} = \sum_{k=1}^K \frac{n_k}{N} w_{k,t+1} \quad w_{t+1} = \sum_{k=1}^K \frac{n_k}{N} w_k^{t+1}$$

This approach can reduce the need to transfer raw traffic data.

However, federated learning introduces its own problems, including:

- Model poisoning
- Communication overhead
- Non-IID data
- Client reliability
- Privacy leakage

Consequently, federated learning should be combined with appropriate security controls.

10. Explainable AI

Security personnel may be reluctant to trust a system that provides an unexplained prediction.

An explainable IDS should ideally report:

Prediction: Malicious

Confidence: High

Main indicators:

- Abnormal connection frequency
- Unexpected destination
- Multiple failed authentication attempts
- Unusual outbound traffic

Explainability can improve:

- Analyst trust
- Incident investigation
- False-positive analysis
- Regulatory accountability

Recent IoT security research specifically identifies explainable AI as an important direction for trusted intelligent IDS systems. ([ScienceDirect](#))

11. Evaluation Framework

A future implementation of the proposed architecture should evaluate more than accuracy.

11.1 Accuracy

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

11.2 Precision

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

11.3 Recall

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

11.4 F1-Score

$$F1 = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

11.5 False Positive Rate

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

11.6 Detection Latency

The system should measure the elapsed time between attack occurrence and detection.

11.7 Resource Utilization

The following should also be measured:

- CPU usage
- RAM consumption
- Model size
- Inference time
- Energy consumption
- Communication overhead

This broader evaluation is important because recent reviews note that deployment reporting for IoT-edge AI security remains inconsistent, particularly regarding latency, memory, energy, and communication overhead. ([MDPI](#))

12. Dataset Considerations

An effective experimental study should use representative IoT security datasets.

Potential datasets include:

- IoT-23
- Bot-IoT
- TON_IoT
- UNSW-NB15
- CIC-based intrusion datasets

Dataset selection should consider:

1. Number of devices.
2. Attack categories.
3. Traffic realism.
4. Class imbalance.
5. Feature availability.
6. Dataset age.
7. Reproducibility.

A model should ideally be evaluated using more than one dataset because excellent performance on a single dataset does not necessarily demonstrate generalization.

13. Research Challenges

13.1 Data Imbalance

Some attacks may occur much less frequently than normal traffic.

This can cause a classifier to favor the majority class.

13.2 Concept Drift

Network behavior changes over time.

A model trained during one period may become less effective later.

13.3 Resource Constraints

Deep-learning models can be computationally expensive.

13.4 Privacy

Network traffic may contain sensitive information.

13.5 Adversarial Machine Learning

Attackers may deliberately manipulate inputs to deceive AI models.

13.6 Model Poisoning

Adaptive systems may be attacked through malicious training data.

13.7 Interoperability

Different IoT manufacturers may use different protocols and data representations.

13.8 False Positives

An overly sensitive system can generate excessive alerts.

14. Future Directions

14.1 TinyML-Based IDS

TinyML may enable selected security intelligence to operate directly on constrained IoT devices.

14.2 Self-Adaptive Security

Future IDS systems could automatically adjust detection thresholds based on changing environmental conditions.

14.3 Federated Security Intelligence

Multiple organizations could potentially collaborate on attack detection without directly sharing raw network traffic.

14.4 AI Agents for Security Operations

Agent-based AI could potentially assist with:

- Alert investigation
- Threat correlation
- Incident prioritization
- Automated response

14.5 Neuromorphic Computing

Neuromorphic architectures may provide future opportunities for highly efficient edge intelligence. Recent IoT security reviews identify neuromorphic computing among possible future directions for ultra-fast IDS. ([ScienceDirect](#))

14.6 Quantum-Resistant Security

As connected infrastructure becomes more dependent on cryptographic mechanisms, future research must also consider long-term cryptographic resilience.

15. Discussion

The analysis presented in this paper indicates that the development of an effective AI-based IDS for IoT requires a multidisciplinary approach.

Machine learning provides powerful classification capabilities, but classification performance alone is not enough.

A model may have high accuracy while:

- consuming excessive memory,
- producing unacceptable latency,
- generating excessive false alarms,
- failing on unseen attacks,
- or requiring unrealistic computational infrastructure.

The proposed architecture addresses these limitations through distributed processing.

The first stage performs inexpensive screening at the edge. Only suspicious events are forwarded for deeper analysis. This principle reduces the amount of computationally intensive processing required.

Another important characteristic is adaptability. IoT networks are dynamic environments. Devices may be added or removed, applications may change, and legitimate traffic patterns may evolve.

Consequently, a fixed model is unlikely to remain optimal indefinitely.

Adaptive learning can help address this problem, but the learning process itself must be protected. An attacker capable of injecting malicious training examples could manipulate the IDS.

Privacy also remains a significant concern. Centralized collection of network traffic can create additional risks. Federated learning offers one possible solution by moving training closer to data sources, although federated systems introduce new attack surfaces.

Explainability is equally important. Security administrators need sufficient information to understand and validate automated decisions.

The future IoT IDS should therefore be viewed as an **intelligent security ecosystem**, rather than simply an ML classifier.

16. Limitations

This paper presents a conceptual framework and literature-based analysis rather than a completed experimental implementation.

Therefore, no fabricated experimental accuracy, latency, or resource-consumption results are reported.

The proposed architecture requires future validation through:

- Real IoT devices.
- Multiple datasets.
- Edge hardware.
- Different network configurations.
- Unknown attack scenarios.
- Adversarial testing.
- Long-term deployment.
- Federated-learning experiments.

These experiments would be necessary before making quantitative claims regarding the superiority of the proposed framework.

17. Conclusion

The rapid expansion of the Internet of Things has produced significant opportunities for intelligent automation while simultaneously creating an increasingly complex cybersecurity environment. The diversity, scale, resource limitations, and connectivity of IoT devices make conventional security mechanisms insufficient when used in isolation.

Artificial intelligence offers important capabilities for identifying malicious behavior, detecting anomalies, classifying attacks, and supporting automated security responses. Machine learning and deep learning techniques have therefore become important areas of IoT cybersecurity research. However, their practical deployment requires careful consideration of computational cost, latency, energy consumption, privacy, explainability, model drift, and adversarial robustness.

This paper proposed a lightweight and adaptive AI-driven intrusion detection framework based on distributed monitoring, edge intelligence, multi-stage detection, adaptive learning, and centralized security management.

The key principle is to avoid treating every network event with the same computational intensity. Normal traffic can be processed using lightweight mechanisms, while suspicious events can receive deeper analysis.

Future research should therefore focus not merely on developing increasingly complex AI models, but on creating **efficient, explainable, adaptive, privacy-aware, and deployment-ready security systems**.

Recent research in IoT-edge cybersecurity similarly indicates that the field is increasingly constrained by practical deployment issues rather than by a lack of new classification algorithms. ([MDPI](#))

The integration of edge computing, federated learning, explainable AI, lightweight models, and adaptive security mechanisms provides a promising foundation for the next generation of IoT intrusion detection.

References

1. Xue, Q., Xue, P., Wang, Z., & Ma, H. (2026). *Artificial Intelligence for Cybersecurity in IoT-Edge Systems: A Structured Review of Methods, Datasets, Evaluation, and Deployment Challenges*. Electronics, 15(11), 2409. ([MDPI](#))
 2. *Unveiling IoT ecosystem security: A review of intelligent IDS, trends, challenges, and future directions*. Computers and Electrical Engineering, 128, 110626, 2025. ([ScienceDirect](#))
 3. *A Survey of Machine Learning Approaches to IoT Security*. Algorithms, 19(5), 384, 2026. ([MDPI](#))
 4. *A secure IoT-edge architecture with data-driven AI techniques for early detection of cyber threats in healthcare*. Discover Internet of Things, 5, 54, 2025. ([Springer](#))
 5. *Internet of cybersecurity things in the third decade of the 21st century: A forward vision*. Computers & Security, 161, 104747, 2026. ([ScienceDirect](#))
 6. *Integrating AI, Blockchain, and Edge Computing for Zero-Trust IoT Security: A Comprehensive Review of Advanced Cybersecurity Framework*. Computers, Materials & Continua, 85(3), 4307–4344, 2025. ([ScienceDirect](#))
 7. Fagan, M., Megas, K., Marron, J., Brady, K., Cuthill, B., Herold, R., Lemire, D., & Hoehn, N. *IoT Device Cybersecurity Guidance for the Federal Government*. National Institute of Standards and Technology.
 8. NIST. *IoT Device Cybersecurity Capability Core Baseline*. NISTIR 8259A.
 9. *Security and privacy challenges in healthcare IoT systems with cloud, edge, and fog computing: A systematic literature review*. Computer Science Review, 62, 101005, 2026. ([Keele Repository](#))
 10. *Trustworthy and Privacy-Preserving AI for IoT-Enabled Healthcare Systems*. Frontiers in Big Data, 2026. ([Frontiers](#))
- ...