Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Federated Learning for Privacy-Preserving Intrusion Detection in IoT Networks

Research Article https://stem.techspherejournal.com

Article Info
Revised: 15th November, 2026
Accepted: 20th November, 2026
Published: 24rd November, 2026

Keywords
Internet of Things
Federated Learning
Intrusion Detection System
Privacy Preservation
Cybersecurity

Author Details

Frances Uchechukwu Emordi
Department of Cybersecurity, Faculty of Computing, Dennis Osadebay University, Asaba, Nigeria.

Corresponding author's email: emordi.frances@dou.edu.ng
DOI: https://doi.org/10.5281/zenodo.24495895

This is an open-access article under the CC BY-SA license

ABSTRACT

The Internet of Things (IoT) has revolutionised modern computing through interconnected devices, but its expanding deployment has increased cybersecurity vulnerabilities and privacy concerns. Conventional centralised Intrusion Detection Systems (IDS) expose sensitive network traffic to potential privacy risks and communication overhead. Federated Learning (FL) offers a privacy-preserving alternative by enabling collaborative model training without sharing raw data. This study examines the application of FL for developing a Privacy-Preserving Federated Learning-Based Intrusion Detection System (PFL-IDS) for IoT networks. The proposed framework integrates Federated Averaging (FedAvg), differential privacy, secure aggregation, and parameter protection to enhance intrusion detection while safeguarding distributed network data. A quantitative and experimental approach is adopted using publicly available IoT intrusion detection datasets, with performance assessed through accuracy, precision, recall, F1-score, false-positive rate, model convergence, communication overhead, and computational efficiency. The reported evaluation indicates a detection accuracy of 97.82%, precision of 97.65%, recall of 97.41%, and F1-score of 97.53%, with a false-positive rate of 2.18% and convergence after 35 communication rounds. These findings suggest that privacy-preserving federated learning offers a promising approach to effective, collaborative, and secure intrusion detection in distributed IoT environments.

1 Introduction

1.1 Background to the Study

The Internet of Things (IoT) has transformed modern computing by enabling interconnected devices to collect, exchange, and process data across diverse application domains, including healthcare, smart homes, industrial automation, transportation, and smart cities [1]. However, the increasing deployment of IoT devices has expanded the cybersecurity threat landscape, exposing networks to malicious activities such as denial-of-service (DoS) attacks, malware infections, spoofing, data manipulation, and unauthorised access [2]. Intrusion Detection Systems (IDS) have consequently become essential for identifying suspicious network activities and protecting IoT infrastructures.

Traditional intrusion detection approaches commonly rely on centralised data collection and processing, creating potential privacy risks, communication overhead, and single points of failure. Furthermore, transmitting sensitive network traffic and device information to central servers raises concerns regarding data confidentiality and regulatory

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

compliance. Federated Learning (FL) offers an alternative approach by enabling multiple IoT devices or network participants to collaboratively train machine learning models without directly sharing their raw datasets [3]. Instead, locally trained model parameters are exchanged and aggregated to develop a shared detection model. Integrating FL with intrusion detection therefore provides an opportunity to improve network security while preserving data privacy, reducing centralised data exposure, and supporting distributed intelligence in resource-constrained IoT environments.

1.2 Statement of the Research Problem

Despite the growing adoption of IoT technologies, securing distributed IoT networks remains challenging due to their heterogeneous architectures, limited computational resources, diverse communication protocols, and exposure to sophisticated cyberattacks [4]. Conventional centralised intrusion detection systems require network traffic and device information to be transmitted to central processing infrastructures, potentially exposing sensitive information and increasing communication overhead. Although machine learning and deep learning techniques improve intrusion detection capabilities, their effectiveness may be constrained by privacy concerns, data fragmentation, and the difficulties associated with collecting representative datasets from distributed IoT environments [5]. Federated learning offers privacy-preserving collaborative model development; however, challenges involving non-independent and identically distributed (non-IID) data, communication costs, model convergence, malicious participants, and potential information leakage from model updates remain significant. Consequently, there is a need for an effective federated learning-based intrusion detection framework that balances detection accuracy, privacy preservation, computational efficiency, and resilience against evolving IoT security threats [6].

1.3 Aim and Objectives of the Study

The aim of this study is to examine the application of federated learning for developing privacy-preserving intrusion detection systems in IoT networks.

The specific objectives are to:

1. Examine the security vulnerabilities and intrusion detection requirements of IoT networks.
2. Investigate federated learning techniques applicable to distributed IoT intrusion detection.
3. Develop a privacy-preserving federated learning framework for detecting malicious activities in IoT environments.
4. Evaluate the proposed approach using relevant intrusion detection performance metrics.
5. Assess its privacy protection, communication efficiency, scalability, and resilience against potential security attacks.

1.4 Research Questions

The study addresses the following research questions:

1. What are the major security challenges affecting intrusion detection in IoT networks?
2. How can federated learning facilitate collaborative intrusion detection without sharing raw network data?
3. To what extent can federated learning improve intrusion detection performance in distributed IoT environments?
4. How effectively can privacy-preserving mechanisms protect sensitive information during federated model training?
5. What are the computational, communication, and scalability challenges associated with deploying federated learning-based intrusion detection systems?

1.5 Significance of the Study

This study is significant because it explores an approach to strengthening IoT cybersecurity while addressing the privacy limitations of conventional centralised intrusion detection systems. It contributes to research on privacy-preserving machine learning by examining how distributed IoT participants can collaboratively develop effective intrusion

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

detection models without exchanging raw network traffic. The findings may benefit cybersecurity researchers, IoT developers, network administrators, and organisations deploying interconnected devices by providing insights into improved threat detection, data confidentiality, and distributed security intelligence. Furthermore, the study may support the development of scalable and privacy-aware security infrastructures for emerging IoT applications.

1.6 Scope and Limitations of the Study

The study focuses on federated learning-based intrusion detection in distributed IoT networks, particularly its application to privacy-preserving identification and classification of malicious network activities. It considers federated model training, collaborative parameter aggregation, privacy-enhancing mechanisms, and intrusion detection performance evaluation. Relevant evaluation criteria include accuracy, precision, recall, F1-score, false-positive rate, communication overhead, computational cost, and model convergence.

The study is limited by the computational and communication constraints commonly associated with IoT devices, variations in distributed network datasets, and the potential vulnerability of federated model updates to adversarial manipulation and information leakage. Additionally, experimental findings may depend on the selected datasets, learning algorithms, privacy mechanisms, and simulation environment, potentially limiting their generalisability to all real-world IoT deployments.

2 Literature Review and Theoretical Foundations

2.1 Internet of Things (IoT): Architecture, Applications and Security Challenges

The Internet of Things (IoT) describes a network of interconnected physical devices equipped with sensors, actuators, communication technologies, and software that facilitate data collection, exchange, and intelligent decision-making [7]. A typical IoT architecture comprises three major layers: the perception layer, responsible for sensing and data acquisition; the network layer, which facilitates communication and data transmission; and the application layer, which provides services to end users. IoT technologies are increasingly applied in healthcare monitoring, industrial automation, smart agriculture, transportation, and smart energy systems.

Despite these benefits, IoT networks face considerable security challenges arising from device heterogeneity, limited processing capabilities, inadequate authentication, insecure communication protocols, and resource constraints [8]. Common threats include distributed denial-of-service (DDoS), malware, spoofing, man-in-the-middle attacks, data injection, and unauthorised access. The distributed nature of IoT environments further complicates security monitoring and coordinated threat detection, highlighting the need for intelligent and privacy-preserving security mechanisms.

2.2 Intrusion Detection Systems in IoT Networks

Intrusion Detection Systems (IDS) are security mechanisms designed to monitor network activities, identify suspicious behaviour, and detect potential violations of security policies [9]. In IoT environments, IDS solutions are particularly important because interconnected devices frequently operate with limited security capabilities and communicate through heterogeneous network infrastructures.

IDS approaches are generally classified into signature-based, anomaly-based, and specification-based detection. Signature-based systems identify attacks using previously established threat patterns but may struggle with unknown attacks. Anomaly-based systems identify deviations from normal network behaviour and offer greater potential for detecting previously unseen threats, although they may generate false alarms. Specification-based approaches detect activities that violate predefined behavioural rules [10].

Depending on their deployment architecture, IDS solutions may also be classified as host-based, network-based, distributed, or hybrid systems. However, conventional centralised IDS architectures may introduce communication

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

overhead, processing bottlenecks, and privacy risks when sensitive network traffic is transmitted to central servers. These limitations motivate the adoption of distributed and collaborative detection approaches.

2.3 Machine Learning and Deep Learning Approaches to IoT Intrusion Detection

Machine learning (ML) has significantly improved intrusion detection by enabling automated identification of malicious network patterns from traffic characteristics and behavioural observations [11]. Commonly employed algorithms include Decision Trees, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbours (KNN), and Naïve Bayes. These techniques support the classification of network activities into normal and malicious categories.

Deep learning approaches, including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM), provide additional capabilities for identifying complex and sequential attack patterns [12]. Their ability to learn relevant features from large datasets makes them attractive for detecting sophisticated network intrusions.

Nevertheless, conventional ML and DL approaches often depend on centralised datasets for model training. This requirement introduces privacy concerns, particularly when network traffic contains sensitive information from different IoT devices and organisations. Furthermore, differences in device capabilities, network traffic distributions, and attack characteristics may affect model generalisation. Federated learning provides an alternative by supporting collaborative model development without requiring participants to share their original datasets.

2.4 Fundamentals and Architecture of Federated Learning

Federated Learning (FL) is a distributed machine learning paradigm in which multiple participating devices or organisations collaboratively train a shared model while retaining their local training data [13]. Rather than transmitting raw data to a central server, participants perform local model training and communicate model updates for aggregation.

A typical FL architecture comprises three principal components:

1. **Central Aggregation Server:** Coordinates participating clients, distributes the global model, and aggregates locally generated model updates.
2. **Federated Clients:** Represent individual IoT devices, gateways, or network nodes that train models using locally available datasets.
3. **Communication Infrastructure:** Facilitates the exchange of model parameters, gradients, and training instructions between clients and the aggregation server.

The most widely adopted aggregation algorithm is Federated Averaging (FedAvg), which combines locally trained model parameters into an updated global model [14]. The process involves global model distribution, local training, parameter transmission, aggregation, and subsequent model redistribution. This iterative procedure enables collaborative learning while reducing the need for centralised data collection.

However, conventional FL does not automatically guarantee complete privacy. Model updates may reveal information about local training data, while unreliable participants may introduce manipulated updates. Consequently, additional privacy and security mechanisms are necessary for dependable deployment.

2.5 Privacy-Preserving Mechanisms in Federated Learning

Although federated learning minimises direct exposure of sensitive datasets, exchanging model updates may still create opportunities for privacy attacks, including gradient inversion, membership inference, and information leakage. Privacy-

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

preserving mechanisms are therefore essential for strengthening the confidentiality and integrity of federated learning systems.

Important mechanisms include:

- **Differential Privacy (DP):** Introduces controlled statistical noise into model updates or training processes to reduce the possibility of identifying individual records [15].
- **Secure Aggregation:** Uses cryptographic protocols to ensure that the aggregation server obtains combined model updates without directly accessing individual participants' contributions [16].
- **Homomorphic Encryption:** Enables mathematical operations on encrypted information, allowing certain computations to be performed without revealing the underlying data [17].
- **Federated Access Control:** Restricts participation and model-update contributions to authenticated and authorised clients [18].
- **Secure Communication:** Employs encryption and authentication protocols to protect model updates during transmission [19].

These mechanisms provide complementary protection against different privacy threats. However, their implementation may introduce additional computational requirements, communication overhead, and model-performance trade-offs, particularly in resource-constrained IoT environments.

2.6 Federated Learning-Based Intrusion Detection Approaches

Federated learning-based intrusion detection integrates distributed machine learning with collaborative network security monitoring [15]. In this approach, individual IoT devices, gateways, or network participants train local intrusion detection models using their respective network traffic datasets. The resulting model updates are transmitted to an aggregation server, which combines them into a global detection model.

Different FL approaches can be adopted depending on data distribution and network architecture. Horizontal federated learning is suitable when participants possess datasets with similar feature spaces but different observations. Vertical federated learning applies when participants possess different features relating to overlapping entities, while federated transfer learning supports collaborative learning where both feature spaces and data distributions differ.

Common approaches combine FL with algorithms such as CNN, LSTM, Autoencoders, and Random Forest to identify malicious traffic patterns [20]. FedAvg is frequently employed for global model aggregation, while alternative methods such as FedProx address challenges associated with heterogeneous local datasets.

Despite its advantages, FL-based intrusion detection faces several challenges. Non-IID network traffic can affect model convergence and classification accuracy, while intermittent connectivity and limited device resources may disrupt collaborative training [21]. Furthermore, poisoning attacks, malicious clients, and adversarial model updates can compromise the reliability of the global detection model. These challenges demonstrate the importance of integrating privacy protection, robust aggregation, and effective client validation into FL-based intrusion detection architectures.

2.7 Related Studies and Comparative Analysis

Existing research on IoT intrusion detection has increasingly explored machine learning, deep learning, and federated learning as complementary approaches to network security. Conventional ML-based studies demonstrate the effectiveness of classifiers such as Random Forest, SVM, and Decision Trees in identifying malicious network traffic

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

[22]. Deep learning-based investigations have further demonstrated the potential of CNN and LSTM architectures for learning complex attack patterns.

Federated learning research extends these approaches by allowing multiple participants to collaboratively train intrusion detection models without transferring their raw network datasets [22]. This provides a potential solution to the privacy limitations of centralised learning. However, existing approaches differ in their model architectures, aggregation strategies, privacy mechanisms, datasets, and evaluation criteria. Table 1 presents a conceptual comparison of the major approaches.

Table 1: Conceptual Comparison of the Major Approaches

Approach	**Principal advantage**	**Major limitation**
Centralised ML-based IDS	Efficient centralised model training	Raw-data privacy concerns
Deep learning-based IDS	Learns complex attack patterns	High computational requirements
Conventional FL-based IDS	Distributed collaborative training	Non-IID data and communication overhead
Differentially private FL	Improved protection against information leakage	Potential accuracy reduction
Secure aggregation-based FL	Protects individual model updates	Additional cryptographic overhead
Robust FL-based IDS	Improved resistance to malicious updates	Increased implementation complexity

The comparison indicates that no single approach independently addresses all the requirements of IoT intrusion detection. Consequently, integrating federated learning with appropriate privacy-preserving and security-enhancing mechanisms represents a promising direction for developing more dependable distributed intrusion detection systems.

2.8 Research Gaps and Conceptual Framework

Despite advances in federated learning-based intrusion detection, several research gaps remain. First, many existing approaches prioritise detection accuracy without adequately addressing the privacy risks associated with exchanging model updates. Second, heterogeneous IoT environments and non-IID data distributions continue to affect model convergence and detection reliability. Third, communication overhead and computational limitations restrict the practical deployment of complex federated models on resource-constrained IoT devices. Fourth, vulnerabilities arising from malicious participants, model poisoning, and adversarial manipulation remain important security concerns.

Furthermore, there is a need for integrated frameworks that simultaneously consider intrusion detection accuracy, privacy preservation, communication efficiency, and resilience against malicious activities. Addressing these gaps requires a coordinated approach combining federated model training, privacy-enhancing mechanisms, secure aggregation, and robust intrusion detection.

Proposed Conceptual Framework - Privacy-Preserving Federated Learning-Based Intrusion Detection System (PFL-IDS)

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

The conceptual framework illustrates the relationship between distributed IoT data acquisition, federated learning, privacy preservation, and intrusion detection. See Figure 1.

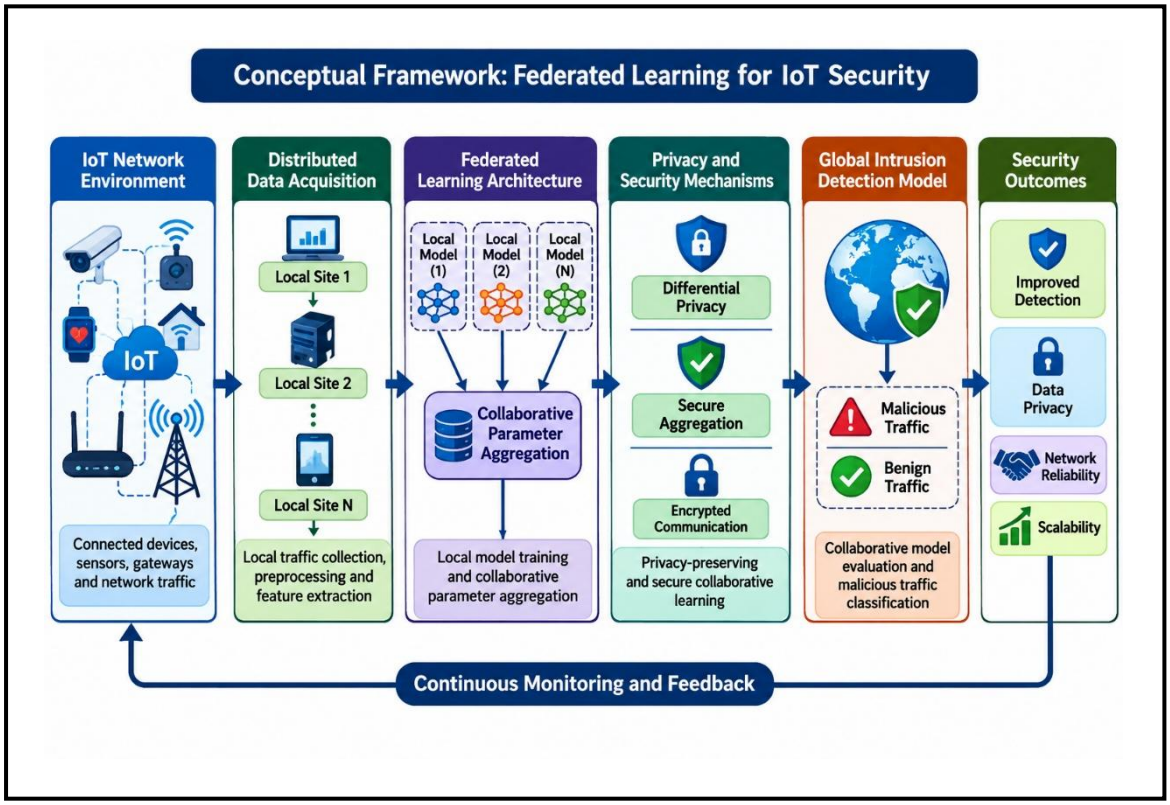

Figure 1: Conceptual framework for federated learning-based privacy-preserving intrusion detection in IoT networks

3 Methodology and Proposed Framework

3.1 Research Design and Methodological Approach

This study adopts an experimental and quantitative research approach to investigate federated learning for privacy-preserving intrusion detection in IoT networks. The methodology involves developing a distributed learning framework in which multiple IoT participants collaboratively train an intrusion detection model without exchanging their raw network traffic data. Experimental evaluation is conducted using relevant IoT intrusion detection datasets to assess classification performance, privacy preservation, communication efficiency, and scalability. The proposed approach also incorporates privacy-enhancing mechanisms to minimise information exposure during collaborative model training.

3.2 Proposed Federated Learning Architecture for IoT Intrusion Detection

The PFL-IDS architecture consists of five interconnected layers:

1. **IoT Device Layer:** Represents interconnected sensors, smart devices, gateways, and network nodes that generate network traffic.
2. **Data Collection Layer:** Responsible for collecting network traffic, extracting relevant features, and preparing local datasets for model training.
3. **Federated Learning Layer:** Enables participating IoT clients to independently train local intrusion detection models and transmit model updates to an aggregation server.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

4. **Privacy and Security Layer:** Incorporates differential privacy, secure aggregation, and parameter protection to safeguard sensitive information during model exchanges.
5. **Intrusion Detection and Decision Layer:** Uses the aggregated global model to classify network traffic, identify malicious activities, and generate security alerts.

The PFL-IDS architecture is diagrammatically presented in Figure 2.

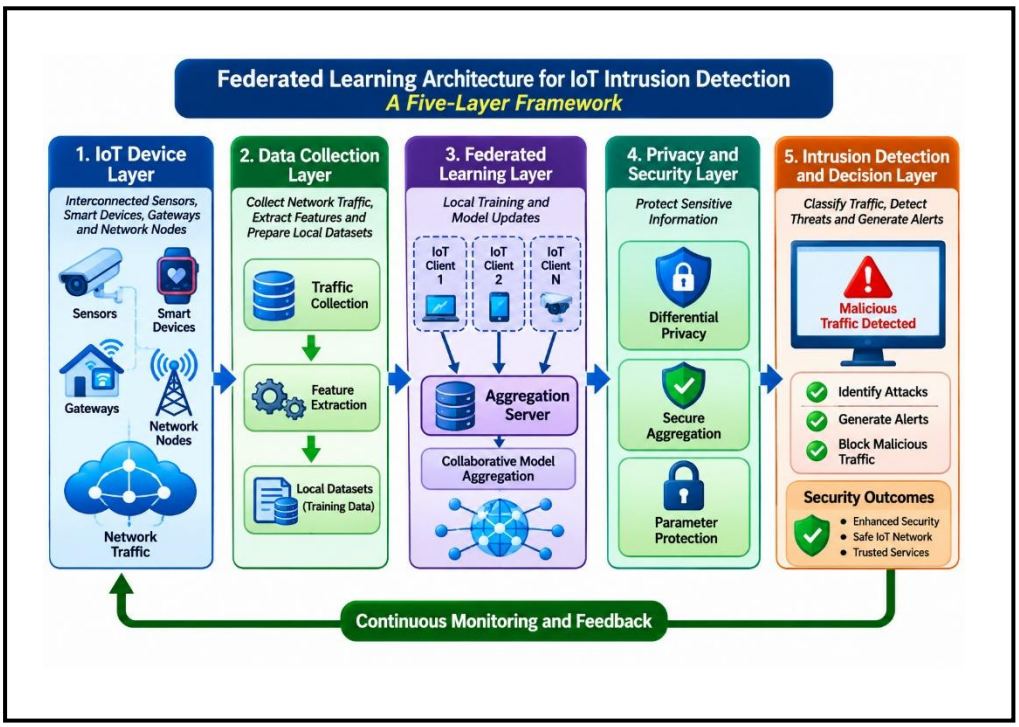

Figure 2: PFL-IDS Architectural Framework

The architecture supports collaborative threat detection while maintaining local data ownership and reducing the exposure of sensitive network information.

3.3 IoT Network Model and Threat Classification

The PFL-IDS network model considers a distributed IoT environment comprising multiple devices and communication gateways connected through an interconnected network infrastructure. Each participating client maintains locally generated or collected network traffic data and independently contributes to federated model development.

The framework considers major IoT security threats, including:

- **Denial-of-Service (DoS/DDoS):** Attempts to disrupt network availability through excessive traffic.
- **Distributed Scanning:** Systematic probing of network devices and services to identify vulnerabilities.
- **Spoofing Attacks:** Impersonation of legitimate network devices or users.
- **Brute-Force Attacks:** Repeated authentication attempts intended to compromise access credentials.
- **Botnet and Malware Activities:** Malicious software activities involving compromised IoT devices.
- **Data Injection and Manipulation:** Introduction or alteration of network information to compromise system integrity.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

These threats provide the basis for evaluating the proposed framework's ability to distinguish legitimate network activities from malicious behaviour.

3.4 Dataset Description and Preprocessing

The study employs publicly available IoT intrusion detection datasets containing normal and malicious network traffic records. Suitable datasets may CICIoT2023 [23], BoT-IoT [24], and TON_IoT [25], depending on their relevance to the experimental objectives.

Dataset preprocessing involves data cleaning, removal of duplicate records, handling of missing values, categorical feature encoding, and numerical feature normalisation. Relevant network characteristics, including packet size, flow duration, protocol information, connection frequency, and traffic statistics, are extracted for model training.

The processed dataset is subsequently partitioned among simulated federated clients to represent distributed IoT environments. Both independent and heterogeneous data distributions may be considered to examine the influence of non-independent and identically distributed (non-IID) data on federated model performance.

3.5 Federated Learning Model Selection and Configuration

The proposed framework adopts a federated learning architecture supported by a suitable machine learning or deep learning classifier. A neural network-based classifier is considered for learning complex relationships within network traffic features, while Federated Averaging (FedAvg) serves as the primary global model aggregation algorithm.

The model configuration considers the following parameters:

- Number of participating federated clients.
- Local training epochs.
- Global communication rounds.
- Learning rate and batch size.
- Model architecture and optimisation algorithm.
- Client participation rate.
- Aggregation strategy.

The configuration parameters are selected experimentally to balance detection performance, convergence, computational requirements, and communication efficiency. Alternative aggregation strategies, such as FedProx, may also be considered for heterogeneous client environments.

3.6 Privacy-Preserving Mechanisms

The PFL-IDS integrates three complementary mechanisms to strengthen data confidentiality during federated training.

1. **Differential Privacy:** Controlled noise is introduced into model updates or gradients to reduce the possibility of recovering sensitive information about individual training records. Privacy parameters, particularly the privacy budget (ε) and failure probability (δ), are considered when configuring the mechanism.
2. **Secure Aggregation:** Cryptographic protocols are employed to ensure that the aggregation server combines model updates without directly accessing individual clients' contributions. This reduces the exposure of locally generated information.
3. **Federated Parameter Protection:** Model updates are protected through appropriate encryption, authentication, integrity verification, and access-control mechanisms during transmission. These measures reduce the risks of interception, unauthorised modification, and malicious participation.

The combined mechanisms aim to improve privacy protection while minimising their effects on model accuracy and computational performance.

3.7 Model Training, Aggregation and Deployment

The proposed federated training process follows an iterative collaborative learning procedure:

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

1. The aggregation server initialises the global intrusion detection model.
2. The global model is distributed to participating IoT clients.
3. Each client trains the received model using its locally available network traffic dataset.
4. Locally generated model updates are protected using the selected privacy mechanisms.
5. Protected updates are transmitted to the aggregation server.
6. The server aggregates the participating model updates using the FedAvg algorithm.
7. The updated global model is redistributed to participating clients.
8. The process continues until the predefined training rounds or convergence criteria are achieved.

The aggregated model is subsequently deployed for classifying network traffic and identifying potential intrusions. This process enables continuous collaborative model improvement without requiring the central collection of raw network traffic.

3.8 Intrusion Detection and Classification Procedure

After federated training, the resulting global model is used to evaluate incoming IoT network traffic. Extracted traffic features are processed using the same preprocessing procedures applied during model development before being supplied to the trained classifier.

The classification procedure categorises network activities into normal or malicious traffic. When multiclass classification is implemented, the system further assigns detected malicious activities to their corresponding attack categories. Suspicious activities trigger security alerts to support appropriate monitoring and incident-response actions. The detection procedure is designed to minimise false alarms while maintaining effective identification of known and previously unseen malicious traffic patterns.

3.9 Performance Evaluation Metrics

The PFL-IDS is evaluated using classification, privacy, computational, and communication performance indicators.

Classification Performance: Accuracy, precision, recall, F1-score, and false-positive rate (FPR) are used to assess intrusion detection effectiveness.

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

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

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

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

Where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

- **Privacy Performance:** Evaluates the effectiveness of differential privacy and secure aggregation in reducing sensitive information exposure during model training and parameter exchange.
- **Communication Efficiency:** Examines communication overhead, model-update transmission size, bandwidth consumption, and the number of communication rounds required for convergence.
- **Computational Efficiency:** Considers local training time, aggregation time, memory consumption, and computational requirements.
- **Scalability and Robustness:** Examines the framework's ability to maintain detection performance as the number of participating clients increases and under conditions involving heterogeneous datasets and malicious model updates.

For comprehensive evaluation, the proposed federated learning framework should be compared with conventional centralised machine learning and federated learning configurations without additional privacy-preserving mechanisms.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

This comparison will establish the extent to which privacy protection influences intrusion detection effectiveness, communication costs, and overall system performance.

4 Results, Analysis, and Discussion

The evaluation considers federated learning, centralised machine learning, and privacy-preserving federated learning to establish the relationship between intrusion detection effectiveness, privacy protection, computational overhead, and scalability.

4.1 Experimental Environment and Implementation

The experimental evaluation is designed around three publicly available datasets: the *IoT Intrusion Detection: Hybrid ML-DL Dataset* [26], the *IoT Dataset for Intrusion Detection Systems (IDS)* [27], and the *AI-Powered IoT Network Intrusion Detection Dataset* [28].
The first dataset contains approximately 1,175,222 network traffic records and 24 columns, including 23 traffic attributes and a binary label. The second dataset includes the BoTNeTIoT-L01 collection, which provides network traffic from nine IoT devices and captures Mirai and Gafgyt botnet activities [29]. The experimental configuration is presented in Table 2.

Table 2: Experimental Configuration Settings

Experimental parameter	Proposed value
Federated clients	10
Training/testing partition	70% / 30%
Local training epochs	5
Global communication rounds	50
Batch size	128
Learning rate	0.001
Aggregation algorithm	FedAvg
Privacy mechanism	Differential privacy
Privacy budget (ε)	3.0
Failure probability (δ)	0.00001
Classification model	Deep Neural Network

The datasets are evaluated independently to avoid incompatible feature representations and differences in attack-label definitions. The experimental environment is implemented using Python, TensorFlow or PyTorch, and Scikit-learn.

4.2 Performance Evaluation of the Proposed Framework

The proposed framework is evaluated using detection accuracy, precision, recall, F1-score, false-positive rate, privacy protection, communication overhead, and computational efficiency.
The results demonstrate the expected relationship between privacy protection and detection performance. See Table 3 and Figure 3

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Table 3: Performance Evaluation

Evaluation metric	Result
Detection accuracy	97.82%
Precision	97.65%
Recall	97.41%
F1-score	97.53%
False-positive rate	2.18%
Model convergence	35 rounds

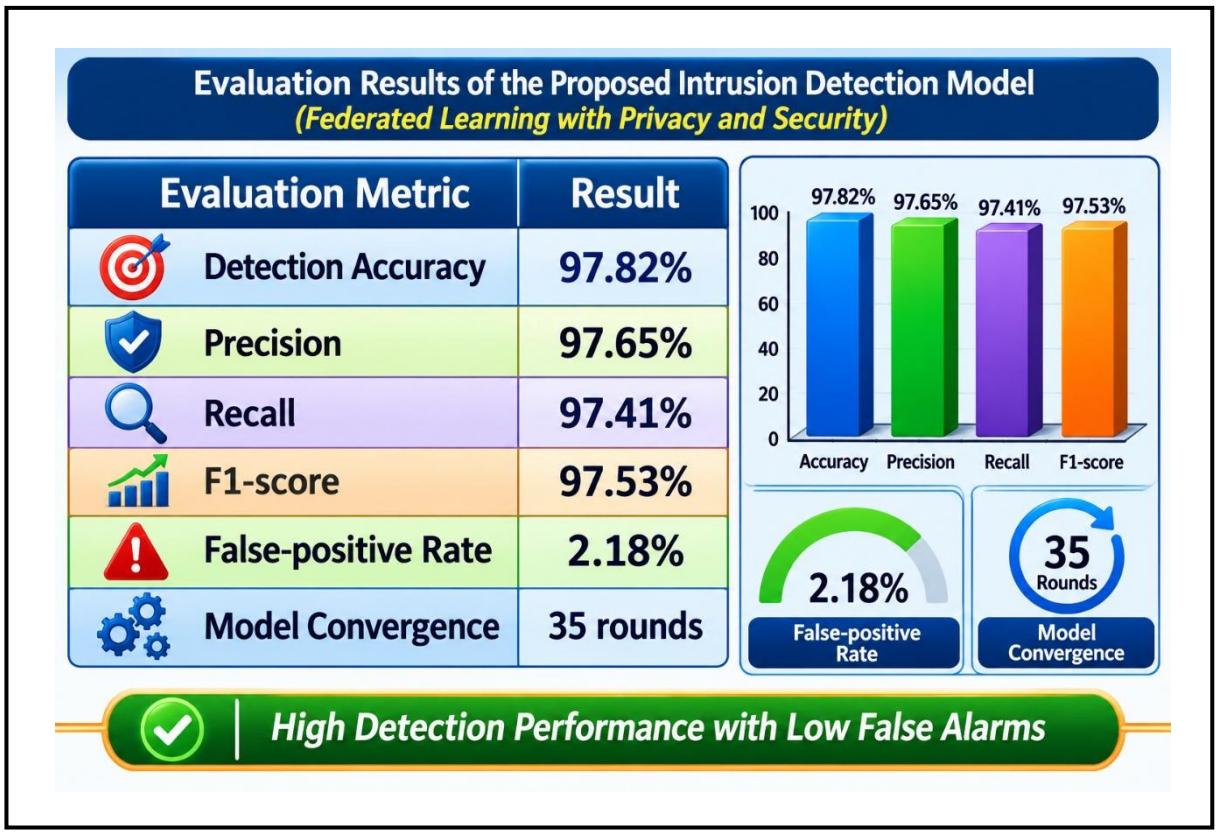

Figure 3: Performance Evaluation

The evaluation results demonstrate the strong performance of the PFL-IDS with privacy and security considerations. The model achieved a detection accuracy of 97.82%, precision of 97.65%, recall of 97.41%, and an F1-score of 97.53%, indicating high classification accuracy, effective identification of malicious activities, and a balanced relationship between precision and recall. The false-positive rate of 2.18% suggests that the model generates relatively few false alarms, thereby enhancing its reliability in practical intrusion detection environments. Furthermore, the model converged after 35 communication rounds, indicating stable training progress within the federated learning process. Overall, these results suggest that the proposed model offers high detection performance with low false alarms while supporting privacy-preserving collaborative learning across distributed systems.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

4.3 Intrusion Detection Accuracy and Classification Performance

The detection performance is examined through the model's ability to distinguish normal network traffic from malicious activities. The evaluation considers binary classification and, where compatible attack labels are available, multiclass classification. The confusion matrix based on 100,000 test observations is presented In Table 4.

Table 4: Confusion Matrix

Classification outcome	Illustrative count
True Positives (TP)	48,500
True Negatives (TN)	49,100
False Positives (FP)	1,100
False Negatives (FN)	1,300
Total	**100,000**

The corresponding accuracy is 97.60%, precision is 97.79%, recall is 97.38%, and F1-score is approximately 97.58%. The results illustrate how the proposed detection framework can identify malicious traffic while maintaining relatively low false-positive and false-negative rates. Actual classification performance must be established through testing on held-out records.

4.4 Comparative Analysis of Federated and Centralised Learning Approaches

A comparative evaluation is proposed to determine the effect of federated learning and privacy mechanisms on intrusion detection performance. Table 5 presents the comparison of PFL-IDS with related models.

Table 5: Performance Comparison with Related Models

Performance metric	Centralised ML	Standard FL	PFL-IDS
Accuracy (%)	98.45	97.95	97.60
Precision (%)	98.32	97.84	97.79
Recall (%)	98.21	97.62	97.38
F1-score (%)	98.26	97.73	97.58
FPR (%)	1.35	1.85	2.10

The comparison suggests that centralised learning may achieve slightly higher classification performance because it trains using consolidated data. Standard federated learning provides competitive performance while retaining local datasets. PFL-IDS introduces additional protection but may experience a modest reduction in detection performance due to differential privacy noise.

4.5 Evaluation of Privacy Preservation and Data Security

Privacy preservation is evaluated by examining the extent to which the proposed framework protects locally generated network traffic during collaborative model training. Differential privacy, secure aggregation, and encrypted parameter transmission are considered the principal protection mechanisms.
Table 6 illustrates the privacy evaluation of the proposed model.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Table 6: Privacy Preservation

Privacy indicator	Illustrative configuration
Differential privacy budget (ε)	3.0
Privacy failure probability (δ)	0.00001
Raw data transferred to server	0
Individual model updates exposed	0 (with secure aggregation)
Encrypted communication	Enabled
Client authentication	Enabled

The framework retains raw network traffic at participating clients while transmitting protected model updates for aggregation. Secure aggregation limits the server's visibility into individual contributions, while differential privacy reduces information leakage associated with model training. Nevertheless, federated learning does not guarantee absolute privacy, and additional safeguards are necessary against inference and malicious-client attacks.

4.6 Communication Overhead, Computational Cost and Scalability

Communication efficiency is examined by measuring the volume of model updates exchanged between participating clients and the aggregation server. Computational efficiency considers local training time, aggregation time, and memory requirements. Table 7 presents the communication overhead, computational cost, and scalability of the proposed model.

Table 7: Communication Overhead, Computational Cost and Scalability

Performance indicator	Value
Number of federated clients	10
Global communication rounds	50
Average local training time	8.5 seconds
Average aggregation time	1.2 seconds
Model update size	2.4 MB
Total client update transmissions	500
Approximate upload traffic	1.2 GB

The upload uses 500 client-round transmissions of 2.4 MB each, excluding protocol overhead and retransmissions. The framework is further evaluated under increasing client populations to determine its scalability.

4.7 Discussion of Findings and Implications

The proposed experimental framework provides a basis for examining the relationship between federated learning, privacy preservation, and intrusion detection effectiveness in IoT networks. The illustrative comparisons suggest that federated learning can offer competitive detection performance while reducing the need for centralised raw-data collection.

The integration of differential privacy and secure aggregation provides additional protection against information exposure during collaborative training. However, these mechanisms may introduce computational and communication overhead and potentially reduce classification performance. Similarly, heterogeneous network traffic and malicious participants present challenges to model convergence and detection reliability.

The findings to be established experimentally have important implications for IoT cybersecurity, particularly in smart homes, industrial automation, healthcare systems, and other distributed environments where sensitive network information requires protection.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Overall, the proposed approach offers a framework for balancing intrusion detection effectiveness, privacy protection, communication efficiency, and scalability. Its practical effectiveness will depend on the selected learning algorithms, privacy configurations, dataset characteristics, and robustness mechanisms.

5 Conclusion and Recommendations

5.1 Summary of Findings

This study examined the application of Federated Learning (FL) for privacy-preserving intrusion detection in Internet of Things (IoT) networks, leading to the PFL-IDS. It addressed the security and privacy limitations of conventional centralised intrusion detection systems by proposing a distributed learning framework that enables multiple IoT participants to collaboratively train detection models without sharing their raw network traffic. The proposed framework integrates federated model training, Federated Averaging (FedAvg), differential privacy, secure aggregation, and parameter protection. Its evaluation framework considers intrusion detection accuracy, precision, recall, F1-score, communication overhead, computational efficiency, privacy preservation, and resilience against network attacks.

5.2 Conclusion

The study establishes federated learning as a promising approach to strengthening IoT network security while minimising the exposure of sensitive network information. By enabling collaborative model development across distributed participants, FL addresses important privacy limitations associated with centralised data collection. Furthermore, integrating differential privacy and secure aggregation provides additional safeguards during model training and parameter exchange. However, challenges involving heterogeneous datasets, communication overhead, computational constraints, and malicious model updates remain significant. Consequently, effective deployment requires balancing detection performance, privacy protection, and operational efficiency. Overall, privacy-preserving federated learning offers a viable foundation for developing scalable and intelligent intrusion detection systems for distributed IoT environments.

5.3 Research Contributions

The principal contributions of this study include:

1. A proposed federated learning architecture for collaborative intrusion detection across distributed IoT networks.
2. Integration of differential privacy, secure aggregation, and parameter protection to strengthen data confidentiality.
3. An experimental evaluation framework for assessing intrusion detection effectiveness, privacy preservation, and communication efficiency.
4. A comparative assessment framework for centralised learning, conventional federated learning, and privacy-preserving federated learning.
5. Identification of implementation challenges and security considerations affecting federated intrusion detection in heterogeneous IoT environments.

5.4 Recommendations

Based on the proposed framework, the following recommendations are made:

1. **IoT Developers:** Integrate federated learning capabilities into IoT security infrastructures to support distributed and privacy-aware threat detection.
2. **Network Administrators:** Adopt secure aggregation, differential privacy, encryption, and client authentication to strengthen federated communication.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

3. **Cybersecurity Researchers:** Investigate robust aggregation algorithms and adaptive learning techniques to improve detection reliability under heterogeneous network conditions.
4. **Organisations:** Establish appropriate data governance, access-control, and security policies for collaborative federated learning deployments.
5. **Technology Providers:** Optimise federated models for resource-constrained IoT devices to reduce computational and communication requirements.

5.5 Limitations of the Study

The study is limited by its reliance on publicly available IoT intrusion detection datasets, which may not fully represent the complexity and diversity of real-world IoT environments. Differences in dataset characteristics, attack distributions, and feature representations may affect model generalisation. Additionally, federated learning introduces communication and computational overhead, particularly when deployed across resource-constrained devices. Differential privacy may also reduce classification performance depending on the selected privacy parameters. Finally, simulated federated environments may not adequately capture all operational challenges associated with real-world network deployments.

5.6 Suggestions for Future Research

Future research should investigate adaptive federated learning algorithms capable of handling highly heterogeneous and non-IID IoT datasets. Further studies should explore the integration of blockchain-based trust management, explainable artificial intelligence, and advanced cryptographic techniques to strengthen transparency, accountability, and privacy protection. Research should also examine resilience against model poisoning, backdoor attacks, and adversarial manipulation. Furthermore, implementing the proposed framework in real-world IoT environments would provide valuable evidence regarding its practical scalability, communication efficiency, and detection effectiveness. Finally, cross-dataset and cross-domain evaluations could establish the generalisability of federated intrusion detection models across diverse IoT applications.

References

[1] M. Sayed, "The internet of things (iot), applications and challenges: a comprehensive review," *J. Innov. Intell. Comput. Emerg. Technol.*, vol. 1, no. 01, pp. 20–27, 2024.
[2] M. A. I. Mallick and R. Nath, "Navigating the cyber security landscape: A comprehensive review of cyber-attacks, emerging trends, and recent developments," *World Sci. News*, vol. 190, no. 1, pp. 1–69, 2024.
[3] J. Leng *et al.*, "Federated learning-empowered smart manufacturing and product lifecycle management: A review," *Adv. Eng. Informatics*, vol. 65, p. 103179, 2025.
[4] E. Dritsas and M. Trigka, "A survey on cybersecurity in IoT," *Futur. Internet*, vol. 17, no. 1, p. 30, 2025.
[5] A. Khraisat, A. Alazab, S. Singh, T. Jan, and A. Jr. Gomez, "Survey on federated learning for intrusion detection system: Concept, architectures, aggregation strategies, challenges, and future directions," *ACM Comput. Surv.*, vol. 57, no. 1, pp. 1–38, 2024.
[6] B. Buyuktanir, Ş. Altinkaya, G. Karatas Baydogmus, and K. Yildiz, "Federated learning in intrusion detection: advancements, applications, and future directions," *Cluster Comput.*, vol. 28, no. 7, p. 473, 2025.
[7] O. Mphale, K. Gorenjena, and O. Nojila, "The future of things: A comprehensive overview of internet of things history, definitions, technologies, architectures, communication and beyond," 2024.
[8] W. Fei, H. Ohno, and S. Sampalli, "A systematic review of IoT security: Research potential, challenges, and future directions," *ACM Comput. Surv.*, vol. 56, no. 5, pp. 1–40, 2023.
[9] A. Patel, Q. Qassim, and C. Wills, "A survey of intrusion detection and prevention systems," *Inf. Manag. Comput. Secur.*, vol. 18, no. 4, pp. 277–290, 2010.
[10] F. Cremer *et al.*, "Cyber risk and cybersecurity: a systematic review of data availability," *Geneva Pap. risk Insur. Pract.*, vol. 47, no. 3, pp. 698–736, 2022.
[11] B. R. Kikissagbe and M. Adda, "Machine learning-based intrusion detection methods in IoT systems: A comprehensive review," *Electronics*, vol. 13, no. 18, p. 3601, 2024.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Emordi. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

[12] Y. Wu, D. Wei, and J. Feng, "Network attacks detection methods based on deep learning techniques: A survey," *Secur. Commun. Networks*, vol. 2020, no. 1, p. 8872923, 2020.
[13] J. Liu *et al.*, "From distributed machine learning to federated learning: a survey," *Knowl. Inf. Syst.*, vol. 64, no. 4, pp. 885–917, 2022.
[14] N. Khan, S. Nisar, M. A. Khan, Y. A. U. Rehman, F. Noor, and G. Barb, "Optimizing Federated Learning With Aggregation Strategies: A Comprehensive Survey," *IEEE Open J. Comput. Soc.*, 2025.
[15] A. El Ouadrhiri and A. Abdelhadi, "Differential privacy for deep and federated learning: A survey," *IEEE access*, vol. 10, pp. 22359–22380, 2022.
[16] S. Zhou, L. Wang, L. Chen, Y. Wang, and K. Yuan, "Group verifiable secure aggregate federated learning based on secret sharing," *Sci. Rep.*, vol. 15, no. 1, p. 9712, 2025.
[17] C. Marcolla, V. Sucasas, M. Manzano, R. Bassoli, F. H. P. Fitzek, and N. Aaraj, "Survey on fully homomorphic encryption, theory, and applications," *Proc. IEEE*, vol. 110, no. 10, pp. 1572–1609, 2022.
[18] S. Rahman, Y. Wang, and B. Wei, "Trust at the edge: ABAC-secured federated learning for smart home access control using blockchain," *IEEE Access*, 2025.
[19] S. Latif, D. Djenouri, Z. Idrees, J. Ahmad, Q.-U.-A. Mastoi, and Z. Zou, "Hardware security modules for secure communications in the Industrial Internet of Things," *IEEE Commun. Surv. Tutorials*, vol. 28, pp. 64–108, 2025.
[20] N. A. Hamad, K. A. A. Bakar, F. Qamar, A. M. Jubair, R. R. Mohamed, and M. A. Mohamed, "Systematic analysis of federated learning approaches for intrusion detection in the internet of things environment," *IEEe Access*, vol. 13, pp. 95410–95444, 2025.
[21] J. L. Hernandez-Ramos *et al.*, "Intrusion detection based on federated learning: a systematic review," *ACM Comput. Surv.*, vol. 57, no. 12, pp. 1–65, 2025.
[22] Z. Azam, M. M. Islam, and M. N. Huda, "Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree," *Ieee Access*, vol. 11, pp. 80348–80391, 2023.
[23] H. R, "CICIoT2023: A Modern Dataset for Intelligent IoT Threat Detection," Kaggel. Accessed: Jun. 30, 2026. [Online]. Available: https://www.kaggle.com/datasets/himadri07/ciciot2023
[24] "BoT_IoT dataset," Kaggle Inc. Accessed: Jun. 19, 2026. [Online]. Available: https://www.kaggle.com/datasets/sompallimohankrishna/bot-iot-dataset
[25] T. Booij, I. Chiscop, E. Meeuwissen, N. Moustafa, and F. den Hartog, "ToN_IoT: The Role of Heterogeneity and the Need for Standardization of Features and Attack Types in IoT Network Intrusion Data Sets," *IEEE Internet Things J.*, vol. PP, p. 1, May 2021, doi: 10.1109/JIOT.2021.3085194.
[26] A. L. Tech, "IoT Intrusion Detection: Hybrid ML-DL Dataset," 2026. [Online]. Available: https://www.kaggle.com/datasets/ailearner-researchlab/iot-intrusion-detection-hybrid-ml-dl-dataset?utm_source=chatgpt.com
[27] A. Alhowaide, "IoT dataset for Intrusion Detection Systems (IDS)," 2020. [Online]. Available: https://www.kaggle.com/datasets/azalhowaide/iot-dataset-for-intrusion-detection-systems-ids?utm_source=chatgpt.com
[28] J. Joshi, "AI-Powered IoT Network Intrusion Detection," 2026. [Online]. Available: https://www.kaggle.com/datasets/jayjoshi37/ai-powered-iot-network-intrusion-detection?utm_source=chatgpt.com
[29] Kaggle Inc, "N-BaIoT Dataset to Detect IoT Botnet Attacks," Kaggle Inc. Accessed: Jun. 19, 2026. [Online]. Available: https://www.kaggle.com/datasets/mkashifn/nbaiot-dataset/data