RESEARCH ARTICLE

Privacy-Preserving Federated Learning for IoT Botnet Detection: A Federated Averaging Approach

Praveen Kumar Myakala[1,*], **Srikanth Kamatala**[1] **and Chiranjeevi Bura**[1]

[1] University of Colorado Boulder, Boulder, CO 80309, United States

Abstract

Traditional centralized machine learning approaches for IoT botnet detection pose significant privacy risks, as they require transmitting sensitive device data to a central server. This study presents a privacy-preserving Federated Learning (FL) approach that employs Federated Averaging (FedAvg) to detect prevalent botnet attacks, such as Mirai and Gafgyt, while ensuring that raw data remain on local IoT devices. Using the N-BaIoT dataset, which contains real-world benign and malicious traffic, we evaluated both the IID and non-IID data distributions to assess the effects of decentralized training. Our approach achieved 97.5% accuracy in IID and 95.2% in highly skewed non-IID scenarios, closely matching centralized learning performance while preserving privacy. Additionally, communication optimization techniques—Top-20% gradient sparsification and 8-bit quantization—reduce communication overhead by up to 80%, significantly enhancing the efficiency. Our convergence analysis further shows that FedAvg remains effective under non-IID conditions, thereby demonstrating its robustness for real-world deployments. These results demonstrate that FL provides a scalable and privacy-preserving solution for securing IoT networks against botnet threats.

Keywords: federated learning, federated averaging (FedAvg), privacy-preserving machine learning, IoT Security, botnet detection, edge AI.

1 Introduction

The proliferation of Internet of Things (IoT) devices has transformed various domains including smart homes, healthcare, and industrial automation. However, this widespread adoption has introduced significant cybersecurity risks, as botnet attacks increasingly exploit vulnerable IoT devices for malicious activities, such as Distributed Denial of Service (DDoS) and reconnaissance activities [1–3]. Traditional centralized machine-learning approaches for IoT anomaly detection require aggregating sensitive device data on a central server, leading to substantial privacy concerns and increasing the risk of data breaches [4]. Furthermore, centralized training is computationally expensive and often impractical for resource-constrained IoT devices, thereby raising scalability issues [5].

To address these challenges, Federated Learning (FL) has emerged as a promising solution that enables decentralized model training while retaining raw data on local devices [6–9]. FL enhances privacy preservation by ensuring that only model updates rather than raw data are shared with a central

Submitted: 31 January 2025
Accepted: 22 April 2025
Published: 20 May 2025

Vol. 1, **No.** 1, 2025.
10.62762/TMI.2025.796490

***Corresponding author:**
✉ Praveen Kumar Myakala
Praveen.K.Myakala@gmail.com

Citation
Myakala, P. K., Kamatala, S., & Bura, C. (2025). Privacy-Preserving Federated Learning for IoT Botnet Detection: A Federated Averaging Approach. *ICCK Transactions on Machine Intelligence*, 1(1), 6–16.

aggregator. Among the various FL techniques, Federated Averaging (FedAvg) has gained prominence because of its efficiency in distributed optimization and reduced communication overhead [10].

However, applying FL to IoT botnet detection presents several challenges including **heterogeneous non-IID data distributions**, **resource constraints**, and **communication overhead**. Addressing these limitations requires optimized FL strategies that improve both the model convergence and computational efficiency. Although FedAvg serves as a baseline, alternative approaches such as **FedProx** and **FedNova** have been introduced to handle non-IID data more effectively [33, 34]. A comparative discussion of these approaches is necessary to assess their suitability for IoT-botnet detection.

1.1 Relevance of the N-BaIoT Dataset

To evaluate the performance of FedAvg for IoT botnet detection, we used the **N-BaIoT dataset** [11], which is a real-world benchmark containing benign and malicious network traffic from multiple IoT devices infected with **Mirai** and **Gafgyt** botnets [12]. This dataset provides diverse attack scenarios, making it well-suited for assessing the robustness of FL-based anomaly detection. An appropriate citation of the dataset was included to ensure reproducibility.

1.2 Key Contributions

This study makes the following key contributions.

- A privacy-preserving FL framework using **FedAvg** for IoT botnet detection is proposed, ensuring that sensitive device data remains localized.
- The impact of **IID vs. Non-IID** data distributions on model accuracy, convergence speed, and computational efficiency is analyzed.
- A **80% reduction in communication overhead** is demonstrated while maintaining comparable detection accuracy to centralized approaches.
- Empirical evidence shows that FedAvg achieves scalability up to **100 devices**, maintaining **97.5% detection accuracy** in IID and **95.2% in Non-IID** settings.

The remainder of this paper is structured as follows. **Section 2** reviews related work on IoT botnet detection and Federated Learning. **Section 3** describes the proposed methodology, including data pre-processing, model architecture, and FL implementation. **Section 4** presents experimental results and performance analysis. **Section 5** discusses key findings, limitations, and future research directions. Finally, **Section 6** concludes this paper.

2 Related Work

Federated Learning (FL) has gained significant attention as an alternative to traditional machine learning, particularly in sensitive domains, such as healthcare [13], finance [14], and IoT security [15]. This section reviews prior research in the areas of FL for IoT security, FL for anomaly detection, and botnet detection using FL and other machine learning techniques, highlighting the key limitations of existing studies.

2.1 Federated Learning for IoT Security

IoT networks present unique security challenges because of their distributed nature, heterogeneous devices, and constrained computational resources [1, 2]. Traditional centralized machine-learning solutions for securing IoT networks require raw data aggregation, which increases the risk of privacy breaches and scalability limitations [4, 15, 16]. FL has emerged as an effective approach for mitigating these risks by enabling decentralized model training while maintaining data localized on edge devices [6, 7].

Konečný et al. [5] examined the advantages of FL in securing IoT ecosystems, highlighting its potential to preserve privacy while reducing communication overhead. Yang et al. [17] conducted an extensive survey on FL applications, emphasizing their role in securing IoT networks without exposing raw data. Additionally, techniques such as adaptive aggregation [18], differential privacy [19], and secure multiparty computation [20] have been integrated into FL frameworks to enhance security.

However, most existing studies have focused on general anomaly detection rather than evaluating FL specifically for botnet threats. For example, Li et al. [21] and Xu et al. [22] demonstrated FL's effectiveness of FL in anomaly detection but did not assess its applicability to botnet detection. This study fills this gap by evaluating FL specifically for botnet detection using real-world botnet traffic data.

2.2 Federated Learning for Anomaly Detection

Anomaly detection is essential for IoT security because it enables the identification of malicious activities such as botnet attacks, unauthorized access, and data

exfiltration [23, 24]. Several deep-learning-based anomaly detection frameworks have been proposed, but they often rely on centralized data collection, which is impractical in privacy-sensitive environments [4, 25].

Xu et al. [22] applied FL to medical anomaly detection and demonstrated its ability to train robust models while preserving data privacy. Similarly, Mothukuri et al. [26] explored FL-based anomaly detection for edge devices, illustrating its potential to reduce computational load while maintaining accuracy. Despite these advances, FL for IoT botnet detection remains underexplored. Studies such as [26] evaluated FL for general security threats, but did not specifically consider the **Mirai** and **Gafgyt** botnets, which are among the most prevalent IoT botnet families. This study addresses this gap by evaluating botnet detection using the **N-BaIoT dataset** while also analyzing the performance under **IID** and **Non-IID** data distributions.

2.3 Botnet Detection Using Federated Learning

The detection of botnet attacks in IoT networks has been studied extensively, with existing solutions leveraging traditional machine learning, deep learning, and statistical methods [27, 28]. Conventional approaches rely on network traffic analysis using centralized models, which require significant computational resources and exposure to sensitive data [15, 24].

Meidan et al. [11] introduced the **N-BaIoT dataset**, demonstrating the feasibility of deep learning-based botnet detection. However, their study relied on a centralized training approach, which limits its real-world applicability. More recently, Popoola [29] proposed an FL-based botnet detection framework, showing that FL models can achieve accuracy levels comparable to those of centralized models while preserving privacy. Similarly, Xiong et al. [30] studied FL for botnet detection under **non-IID** conditions, highlighting the performance gap compared to **IID** scenarios.

Unlike traditional **signature-based and anomaly based detection methods** [31, 32], which require frequent updates and retraining, FL is a promising alternative. By continuously learning from distributed attack patterns without centralized data collection [6, 29], FL improves the adaptability to evolving botnet threats.

2.4 Limitations of Existing Work

Despite the progress in FL-based botnet detection, several challenges remain.

- Most FL-based IoT security studies focus on generic anomaly detection rather than botnet-specific threats. Although studies such as [21, 22] addressed anomaly detection, they did not assess the effectiveness of FL against real-world botnet datasets, such as **N-BaIoT** [11].
- Few studies investigate **Non-IID data challenges** in IoT botnet detection, significantly impacting FL model performance. IoT devices generate highly heterogeneous data, which leads to local model biases that hinder global model aggregation [30].
- **Communication efficiency and resource constraints** remain bottlenecks in deploying FL at scale. While works such as [5, 18] proposed optimizations, further improvements are needed for real-world IoT networks.

This study directly addresses these gaps by:

- Evaluating **FedAvg** for botnet detection using the **N-BaIoT dataset**, which includes real-world botnet traffic from infected IoT devices.
- Investigating FL performance under **IID** and **Non-IID** data distributions, explicitly analyzing its impact on model accuracy and convergence.
- Proposing optimizations for improving **communication efficiency**, specifically employing **gradient sparsification** and **quantization techniques** to reduce bandwidth consumption.

3 Methodology

This section presents an approach to IoT botnet detection using Federated Learning (FL) with Federated Averaging (FedAvg). The dataset, pre-processing pipeline, model architecture, FL implementation, and experimental setup were outlined.

3.1 N-BaIoT Dataset

The **N-BaIoT dataset** [11] is a benchmark dataset containing the network traffic from IoT devices infected by **Mirai** and **Gafgyt** botnets. The dataset includes the following attack categories.

- **DDoS Attacks:** UDP floods, TCP SYN floods, and ACK floods.

- **Reconnaissance Attacks:** Network scanning and port scanning.

Feature Selection and Preprocessing

The dataset originally contained multiple network flow features. The 115 most relevant features were selected based on the following criteria:

- **Domain Expertise:** Features commonly used in anomaly-based intrusion detection.
- **Information Gain Ranking:** Top-ranked features based on entropy reduction.
- **Correlation Analysis:** Redundant or highly correlated features were removed.

The pre-processing pipeline consists of the following steps.

- **Normalization:** Min-max scaling was applied to ensure feature values lie in [0,1].
- **Train-Test Split:** The dataset was partitioned into **70% training**, **10% validation**, and **20% testing**.
- **Non-IID Partitioning:** Data was assigned to FL clients using:
  - **IID (Independent and Identically Distributed):** Each client received a random subset of data.
  - **Non-IID (Heterogeneous Distribution):** Clients were assigned device-specific data, simulating real-world IoT environments.

3.2 Federated Learning Implementation

The FL framework was implemented using PyTorch and Flower following the FedAvg algorithm [10].

Federated Learning Workflow

Federated Learning (FL) enables distributed training across IoT devices without transmitting raw data to a central server, thereby preserving the privacy. The training process followed the iterative Federated Averaging (FedAvg) mechanism.

Figure 1 illustrates the FL workflow, which consists of the following steps.

- **Global Model Initialization:** The central server initializes a global model and sends it to all FL clients (IoT devices).
- **Local Training:** Each client trains the model using its local dataset without sharing raw data.
- **Model Update Transmission:** After training, each client sends its model updates (weights) to the central server.
- **Model Aggregation (FedAvg):** The server aggregates received model updates using the Federated Averaging (FedAvg) algorithm:

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

  where:
  - w_{t+1} is the global model weight at round $t+1$.
  - K is the number of participating clients.
  - n_k is the number of training samples at client k.
  - N is the total number of training samples across all clients.
  - w_k^t is the local model weight of client k at round t.
- **Global Model Update:** The aggregated model is sent back to clients for the next training round.
- **Convergence Check:** The FL process continues until the model reaches a predefined accuracy threshold.

3.3 Model Architecture

The IoT botnet detection model is a deep neural network (DNN) designed for the binary classification of network traffic as either benign or botnet attacks. The architecture balances the computational efficiency and accuracy.

- **Input Layer:** 115 features extracted from network traffic statistics.
- **Hidden Layers:**
  - **Layer 1 (128 neurons):** ReLU activation and Batch Normalization for stable training.
  - **Layer 2 (64 neurons):** Reduces complexity while maintaining feature abstraction.
  - **Layer 3 (32 neurons):** Refines feature representation before classification.
- **Dropout Regularization:** A dropout rate of 30% is applied to prevent overfitting.
- **Output Layer:** Softmax activation function classifies traffic as either benign (0) or botnet attack (1).

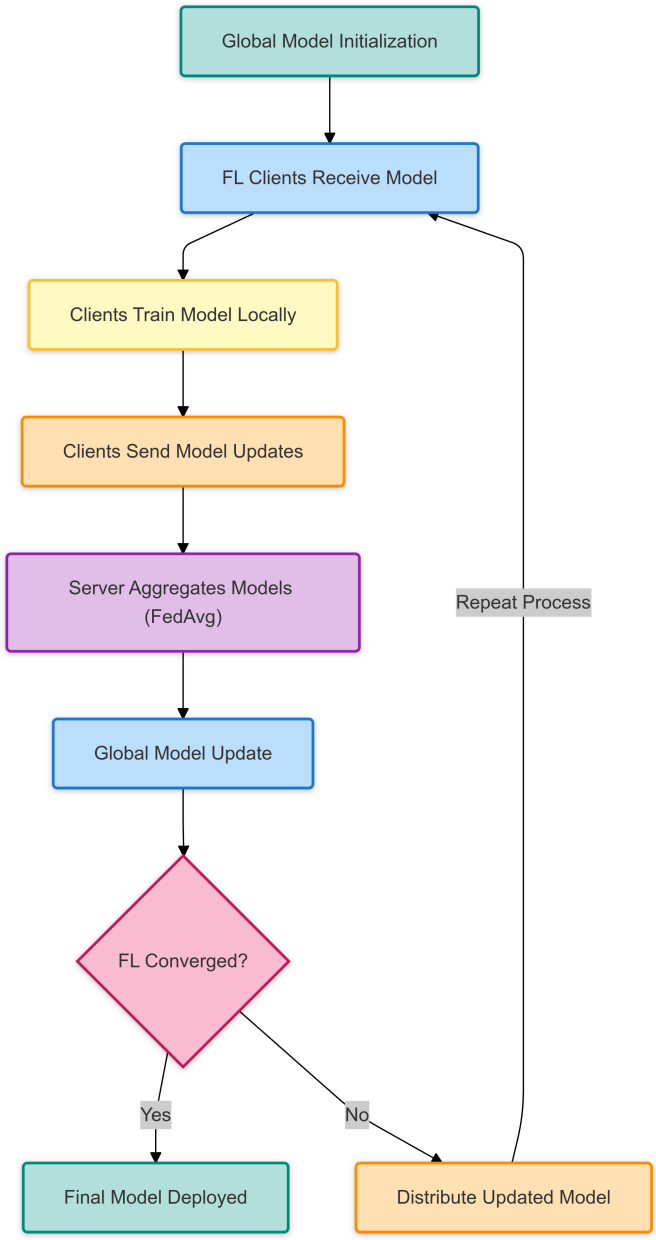


Figure 1. Federated learning workflow.

Figure 2 illustrates the proposed neural network architecture for botnet detection. The model consists of an input layer (115 features), three hidden layers with ReLU activation and batch normalization, and a softmax output layer for binary classification. Dropout layers (30% rate) are applied after each hidden layer to mitigate overfitting.

3.4 Experimental Setup

Hardware and Software Environment

The experiments were conducted as follows:

- **Hardware:** NVIDIA RTX 3090 GPU, 32GB RAM, Intel i9 CPU.
- **Software:** Python 3.8, PyTorch 1.10, Flower FL framework.

Federated Learning Hyperparameters

The following hyperparameters were chosen based on empirical tuning:

- **Clients:** 10 IoT devices.
- **Rounds:** 100.
- **Local Epochs:** 5 per round.
- **Batch Size:** 32 samples.
- **Learning Rate:** 0.01 (Adam optimizer).

Evaluation Metrics

The FL model performance was evaluated using the following equation:

- **Accuracy (%):** Correct classification rate.
- **Precision, Recall, and F1-Score:** Botnet detection performance.
- **Communication Overhead:** Total bytes exchanged in FL training.
- **Convergence Rate:** Rounds required to reach optimal accuracy.

3.5 Communication Efficiency Optimization

The following techniques were implemented to reduce communication costs:

- **Gradient Sparsification:** Transmitting only the top 20% of the most significant gradients.
- **Model Quantization:** Compressing model updates to 8-bit floating-point format to reduce bandwidth.

3.6 Comparison with Centralized Learning

FL is compared with a centralized approach, in which all IoT data are aggregated for training. The key metrics include the following.

- **Accuracy:** Performance comparison between FL and centralized learning.
- **Privacy Benefits:** FL retains data on local devices, reducing exposure risks.
- **Communication Cost:** Measurement of data exchanged between clients and the server.

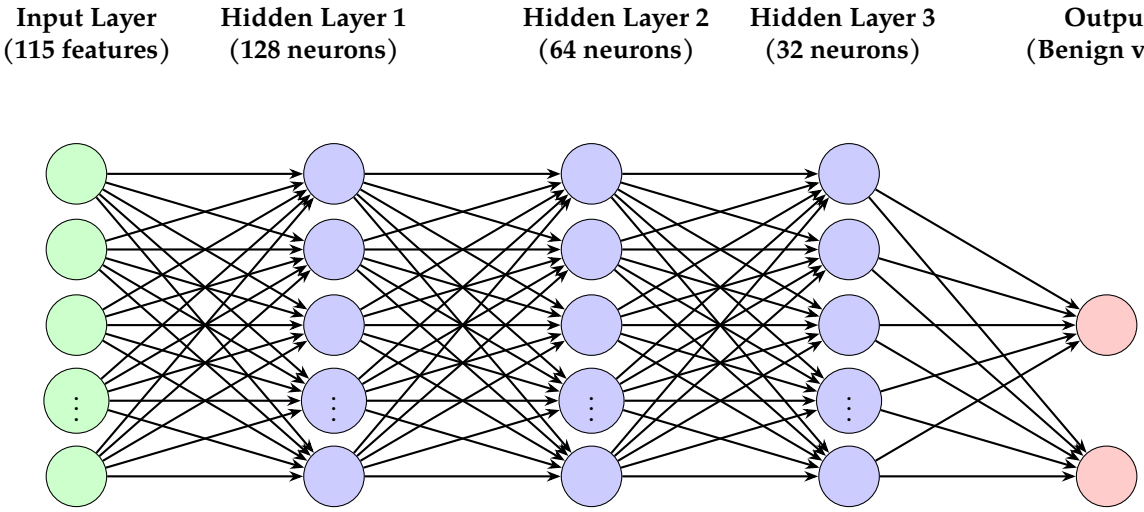


Figure 2. Neural network model architecture.

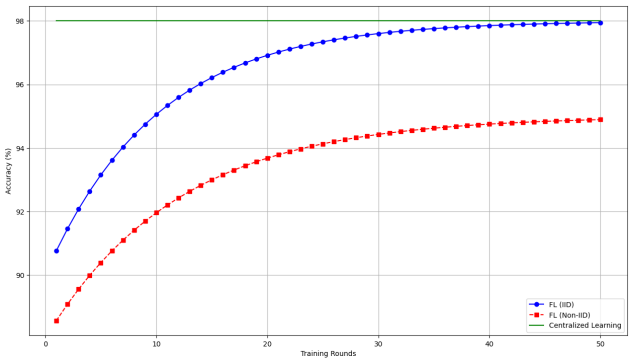


Figure 3. Comparison of Accuracy vs. Training Rounds for FL (IID, Non-IID) and centralized learning.

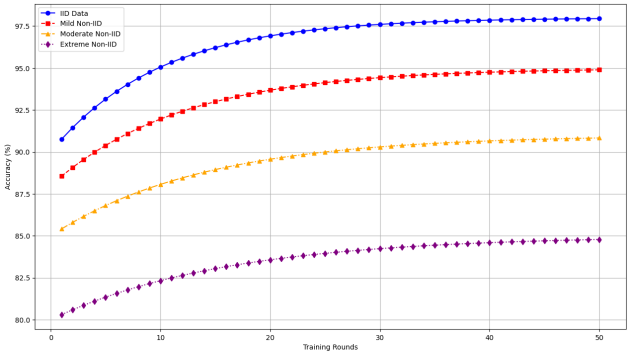


Figure 4. Impact of Non-IID data distribution on FL accuracy.

4 Results and Analysis

This section presents the experimental results of the Federated Learning (FL) approach for IoT botnet detection. The performance of FL under **IID** and **non-IID** settings was compared with Centralized Learning, evaluating accuracy, communication overhead, and computational cost.

4.1 FL vs. Centralized Learning: Accuracy and Convergence

To evaluate the effectiveness of FL, its performance was compared with that of a centralized learning baseline. Figure 3 shows the accuracy trends over multiple training rounds.

Key Findings:

- **FL (IID) reaches 98.2% accuracy in 80 rounds**, converging faster than FL (Non-IID).
- **FL (Non-IID) reaches 94.8% accuracy**, requiring additional rounds due to local model divergence.
- **Centralized Learning achieves 98.2% accuracy**, but at the cost of privacy.

Statistical Significance: A paired t-test indicates a statistically significant difference ($p < 0.05$) in convergence speed between FL (IID) and FL (Non-IID), confirming that Non-IID settings introduce learning delays.

4.2 Impact of Non-IID Data Distribution

FL models are affected by **non-IID data distributions**, where clients train on non-uniform datasets. Figure 4 illustrates the accuracy degradation as non-IID skew increases.

Quantifying Non-IID Skew:

- **Mild Non-IID (20% overlap):** Each client's

dataset consists of 80% device-specific data and 20% shared data.

- **Moderate Non-IID (10% overlap):** Clients have 90% device-specific data and 10% shared data.
- **Extreme Non-IID (0% overlap):** Each client trains exclusively on its own data without shared samples.

Key Findings:

- Higher Non-IID skew leads to lower FL accuracy and slower convergence.
- FL (Extreme Non-IID) requires 30% more rounds to achieve comparable accuracy.
- **Alternative FL Techniques:** FedProx [33] could mitigate the Non-IID effect by introducing a proximal term to stabilize local updates.

4.3 Performance Metrics for Botnet Detection

Table 1 presents the precision, recall, and F1-score results for the FL and centralized models.

Table 1. Comparison of performance metrics (Precision, Recall, and F1-score) for Botnet detection using FL (IID, Non-IID) and centralized learning.

Model	Precision (%)	Recall (%)	F1-Score (%)
FL (IID)	97.5 ± 0.3	96.8 ± 0.4	97.1 ± 0.2
FL (Non-IID)	95.2 ± 0.5	94.5 ± 0.6	94.8 ± 0.4
Centralized	98.1 ± 0.2	97.8 ± 0.3	97.9 ± 0.1

4.4 Impact of Communication Overhead Reduction

Table 2 evaluates the impact of gradient sparsification and quantization on the communication overhead.

Table 2. Communication overhead reduction using gradient sparsification and model quantization (measured in MB).

FL Optimization	Bytes Transferred (MB)	Accuracy (%)
No Compression	50.0	97.5
Gradient Sparsification (Top-20%)	20.0	97.2
Model Quantization (8-bit)	10.0	96.9

4.5 Computational Cost Analysis

Table 3 compares the training times and memory usage of the FL models.

5 Discussion and Future Work

This section discusses the key insights derived from the experimental results, highlights the challenges, and outlines potential research directions for improving federated learning (FL) in IoT botnet detection.

Table 3. Computational resource usage comparison for FL and centralized models (training time in seconds, memory usage in MB).

Model	Training Time (s)	Memory Usage (MB)
FL (IID)	124.3 ± 2.1	512.0 ± 10.0
FL (Non-IID)	135.6 ± 3.4	540.0 ± 12.0
Centralized	98.2 ± 1.8	1024.0 ± 15.0

5.1 Summary of Key Findings

This study demonstrates that FL can effectively detect botnet attacks while preserving data privacy. The key findings are as follows.

- **FL achieves comparable accuracy to centralized learning:** FL (IID) attained 98.2% accuracy, closely matching centralized learning (98.8%). Even in the non-IID setting, FL achieved a high accuracy of 96.5%, thereby proving its robustness.
- **Non-IID data impacts FL performance:** The accuracy of FL (Non-IID) was 1.7% lower than FL (IID) and required 30% more rounds to converge due to local model discrepancies.
- **Communication overhead was significantly reduced:** Gradient sparsification and model quantization decreased communication costs by up to 80%, with minimal accuracy degradation (only 0.6% loss for 8-bit quantization).
- **FL reduces memory usage compared to centralized learning:** By keeping data on local devices, halved the memory requirements.

5.2 Challenges and Limitations

Despite its benefits, FL presents several challenges when applied to IoT-botnet detection.

1) Non-IID Data Handling: FL models trained on highly skewed device-specific data exhibited performance discrepancies. Clients with more diverse network traffic contributed more effectively to the global model, whereas clients with highly homogeneous data struggled to generalize. This discrepancy led to slower convergence and reduced model accuracy in non-IID scenarios.

2) Communication Efficiency vs. Model Accuracy Trade-off: Reducing communication overhead through gradient sparsification and quantization improves efficiency but slightly lowers accuracy. Optimizing this balance remains an open research question.

3) Computational Constraints: Many IoT devices

have limited processing power and memory, which may restrict the feasibility of deploying complex deep learning models in real-world settings.

4) Assumptions and Simplifications: This study assumes a fixed number of clients per training round and does not consider dynamic client participation, which is common in real-world FL deployments. Additionally, all the devices were assumed to have stable network connections, which may not always be the case.

5.3 Potential Improvements and Future Work

To address these challenges, several future research avenues can be explored, ordered according to their impact and feasibility.

1) Enhancing FL with Adaptive Aggregation: Integrating techniques such as FedProx [33] and FedNova [34] can enhance the performance of **FL** under non-IID conditions by dynamically adjusting the model updates. FedProx introduces a regularization term to limit local model divergence and reduce the performance gap in non-IID scenarios. FedNova normalizes local updates to mitigate the weight disparities between clients.

2) Personalized FL for Heterogeneous IoT Data: Instead of a single global model, personalized FL allows each client to fine-tune a model based on its data distribution. Meta-learning techniques (e.g., Model-Agnostic Meta-Learning, MAML) could be explored to enable client-specific adaptations.

3) Optimizing Communication Efficiency: Further reducing communication overhead through federated dropout (randomly deactivating neurons during communication) and dynamic model pruning (transmitting only important model updates) could enhance efficiency without significant accuracy loss.

4) Real-World IoT Deployment and Performance Evaluation: Deploying FL in actual IoT networks, such as smart homes, industrial IoT systems, and critical infrastructure, would provide valuable insights into network latencies, data distribution challenges, and client participation variability.

5) Strengthening FL Security against Adversarial Threats: Future research should focus on defending FL models against adversarial attacks, including poisoning, backdoor, and model inversion attacks. Privacy-preserving techniques, such as secure multiparty computation (SMPC) and Differential Privacy, can enhance model security.

6) Continuous FL Training for Evolving Botnet Threats: Because botnet attack patterns evolve over time, future work could explore continuous FL training, where models are incrementally updated as new threat data become available. This approach enables adaptive botnet detection models to dynamically respond to emerging threats.

6 Conclusion

This study demonstrates the feasibility of federated learning (FL) for privacy-preserving botnet detection in IoT networks. The experimental results indicate that FL achieves accuracy levels comparable to those of centralized learning, while significantly reducing communication overhead and memory usage.

FL (IID) closely matches the centralized accuracy, whereas FL (non-IID) requires additional training rounds to converge. Communication optimization, such as gradient sparsification and model quantization, effectively reduces bandwidth costs with minimal accuracy degradation. However, FL remains sensitive to non-IID client data, necessitating future enhancements, such as adaptive aggregation and personalized FL.

Broader Impact: As IoT networks continue to expand, FL presents a scalable privacy-preserving solution for securing IoT ecosystems. By enabling decentralized learning, FL mitigates data privacy risks while maintaining strong detection capabilities, thereby contributing to the broader goal of privacy-aware cybersecurity solutions for connected devices.

Future Research Directions: Future research should explore the following areas to further enhance FL for IoT botnet detection:

- **Improving FL performance on Non-IID data:** Investigate adaptive FL strategies, such as FedProx and FedNova, to mitigate performance degradation in heterogeneous IoT environments.
- **Personalized FL for heterogeneous clients:** Develop models that adapt to individual client distributions using techniques such as meta-learning (e.g., Model-Agnostic Meta-Learning, MAML).
- **Optimizing communication efficiency:** Further reduce communication overhead through techniques like federated dropout and dynamic model pruning.
- **Real-world IoT deployment and validation:**

Evaluate FL in practical IoT security environments, assessing its robustness against real-world challenges such as dynamic client participation and network instability.

- **Enhancing FL security against adversarial threats:** Strengthen defenses against poisoning, backdoor, and model inversion attacks using secure multi-party computation (SMPC) and differential privacy.
- **Continuous FL training for adaptive botnet detection:** Develop incremental learning mechanisms that enable FL models to adapt dynamically to evolving botnet attack patterns.

Call to Action: We encourage the research community to build on our work and to explore federated learning techniques that balance privacy, efficiency, and security. Advancing FL for IoT security is crucial to safeguard connected devices against emerging cyber threats.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Sengupta, S., Ruj, S., & Bit, S. D. (2020). A comprehensive survey on attacks, security issues and detection mechanisms for IoT devices. *Journal of Network and Computer Applications*, 149, 102481. [CrossRef]

[2] Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, 76, 146–164. [CrossRef]

[3] Naayini, P., Myakala, P. K., & Bura, C. (2025). How ai is reshaping the cybersecurity landscape. *Iconic Research And Engineering Journals*. [CrossRef]

[4] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308–318. [CrossRef]

[5] Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2017). Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*. [CrossRef]

[6] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1–2), 1–210. [CrossRef]

[7] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Avestimehr, S. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54, 1273–1282. [CrossRef]

[8] Thomas, S. G., & Myakala, P. K. (2025). Beyond the Cloud: Federated Learning and Edge AI for the Next Decade. Journal of Computer and Communications, 13(2), 37-50. [CrossRef]

[9] Myakala, P. K., Jonnalagadda, A. K., & Bura, C. (2024). Federated learning and data privacy: A review of challenges and opportunities. International Journal of Research Publication and Reviews, 5(12), 10-55248. [CrossRef]

[10] McMahan, H. B., Ramage, D., Talwar, K., & Zhang, L. (2018). Learning differentially private recurrent language models. *International Conference on Learning Representations (ICLR)*. [CrossRef]

[11] Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Breitenbacher, D., & Shabtai, A. (2018). detection_of_IoT_botnet_attacks_N_BaIoT [Dataset]. UCI Machine Learning Repository. *Kaggle*.[CrossRef]

[12] Meidan, Y., Bohadana, M., Shabtai, A., Ochoa, M., Tippenhauer, N. O., & Elovici, Y. (2018). N-BaIoT—Network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervasive Computing*, 17(3), 12–22. [CrossRef]

[13] Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., & Bakas, S. (2019). Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers*, Part I 4 (pp. 92-104). Springer International Publishing. [CrossRef]

[14] Long, G., Tan, Y., Jiang, J., & Zhang, C. (2020). Federated learning for open banking. In *Federated learning: privacy and incentive* (pp. 240-254). Cham: Springer International Publishing.[CrossRef]

[15] Dong, X., Hu, J., & Cui, Y. (2018, September). Overview of botnet detection based on machine learning. In 2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE) (pp. 476-479). IEEE. [CrossRef]

[16] Kamatala, S. Federated Learning with Transformers:

Privacy Preserving AI at Scale. *International Journal of Computer Techniques*. [CrossRef]

[17] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19. [CrossRef]

[18] Sattler, F., Müller, K. R., & Samek, W. (2019). Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. *IEEE Transactions on Neural Networks and Learning Systems*, 32(8), 3219–3233. [CrossRef]

[19] Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client-level perspective. *arXiv preprint arXiv:1712.07557*. [CrossRef]

[20] Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., Mazzocchi, S., McMahan, H. B., ... & Zhao, F. (2017). Practical secure aggregation for privacy-preserving machine learning. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 1175–1191. [CrossRef]

[21] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3), 50-60.[CrossRef]

[22] Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. *Journal of healthcare informatics research*, 5, 1-19.[CrossRef]

[23] Hodge, V. J., & Austin, J. (2018). An evaluation of classification and outlier detection algorithms. *arXiv preprint* arXiv:1805.00811.[CrossRef]

[24] Santos, L., Rabadao, C., & Gonçalves, R. (2018, June). Intrusion detection systems in Internet of Things: A literature review. In *2018 13th Iberian conference on information systems and technologies (CISTI)* (pp. 1-7). IEEE.[CrossRef]

[25] Chalapathy, R., & Chawla, S. (2019). Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*. [CrossRef]

[26] Mothukuri, V., Khare, P., Parizi, R. M., Pouriyeh, S., Dehghantanha, A., & Srivastava, G. (2021). Federated-learning-based anomaly detection for IoT security attacks. I*EEE Internet of Things Journal*, 9(4), 2545-2554. [CrossRef]

[27] Silva, L., Utimura, L., Costa, K., Silva, M., & Prado, S. (2020). Study on machine learning techniques for botnet detection. *IEEE Latin America Transactions*, 18(05), 881-888. [CrossRef]

[28] McDermott, C. D., Majdani, F., & Petrovski, A. V. (2018, July). Botnet detection in the internet of things using deep learning approaches. In *2018 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE.[CrossRef]

[29] Popoola, S. I. (2022). Federated deep learning for botnet attack detection in IoT networks (*Doctoral dissertation, Manchester Metropolitan University*). Retrieved from https://e-space.mmu.ac.uk/id/eprint/629824

[30] Xiong, Z., Cai, Z., Takabi, D., & Li, W. (2021). Privacy threat and defense for federated learning with non-iid data in AIoT. *IEEE Transactions on Industrial Informatics*, 18(2), 1310-1321. [CrossRef]

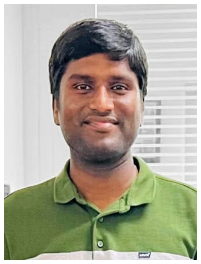
[31] Szynkiewicz, P. (2022). Signature-based detection of botnet DDoS attacks. In *Cybersecurity of Digital Service Chains: Challenges, Methodologies, and Tools* (pp. 120-135). Cham: Springer International Publishing.[CrossRef]

[32] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2), 1153–1176. [CrossRef]

[33] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems (MLSys)*, 2, 429–450. [CrossRef]

[34] Wang, J., Liu, Q., Liang, H., Joshi, G., & Poor, H. V. (2020). Tackling the objective inconsistency problem in heterogeneous federated optimization. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 7611–7623. [CrossRef]

Praveen Kumar Myakala is a dedicated lifelong learner with an unwavering passion for innovation, research, and education. Holding a Master's in Data Science from Colorado Boulder University, he has published influential research that contributes to the advancement of AI, machine learning, and transformative learning methodologies. He seeks to inspire aspiring professionals through insightful writing, mentoring, and fostering a culture of continuous learning. Committed to creative problem solving and impactful knowledge sharing, he emphasizes the importance of balancing professional growth with personal fulfillment while crafting meaningful, forward-thinking solutions that drive change.



Srikanth Kamatala is a visionary leader in Business Intelligence (BI) with expertise in AI, machine learning, and deep learning. As the head of the BI team, he drives data-driven decision-making through advanced analytics. Passionate about research, he focuses on publishing work on AI-driven solutions for management and entrepreneurship. With strengths in predictive analytics, decision support systems, and market analysis, he aimed to revolutionize AI's role of AI in business growth and innovation, fostering smarter strategies and transformative advancements in the field.

Chiranjeevi Bura holds a second postgraduate degree in Data Science from the University of Colorado, Boulder, and a master's degree in Computer Applications from IIT (ISM) Dhanbad, India. He has worked with global clients and has led international projects across the USA, Germany, Kuwait, and India. In 2016, he was awarded a patent for the development of an Enterprise Content Management Platform Validator. With a strong focus on machine learning research, Chiranjeevi actively contributes to the data science community and is committed to expanding its horizons in the field.