F1-Guided Softmax Aggregation for Federated Intrusion Detection in Non-IID Environments

1st Muhamamd Mustaqeem
dept. of Information and Communication Engineering
Islamia University of Bahawalpur
Bahawalpur, Pakistan
mmustaqem786@gmail.com

2nd Muhammad Dawood Sajid
dept. of Information and Communication Engineering
Islamia University of Bahawalpur
Bahawalpur, Pakistan
dawoodsajid239@gmail.com

3rd Sana Tariq
dept. of Information and Communication Engineering
Islamia University of Bahawalpur
Bahawalpur, Pakistan
Sana.tariq@iub.edu.pk

4th Aoun Muhammad
dept. of Information and Communication Engineering
Islamia University of Bahawalpur
Bahawalpur, Pakistan
aoun.muhammad@iub.edu.pk

***Abstract*—Federated learning is a way to allow privacy-preserving intrusion detection systems (IDS) without sharing raw data. But the traditional methods, such as Federated Average(FedAvg), do not work well on Non-Independent and Identically Distributed(non-IID) and imbalanced datasets. This study presents an F1-Guided Softmax Aggregation method, which is achieved by assigning weights to the clients using the F1-scores to enhance the global model. The results obtained from the experiments conducted on the CI-CIDS2017 data set demonstrate that the performance of the proposed model is more accurate, precise, recall, and with better convergence stability. The proposed algorithm outperforms the other two, FedAvg and Scaffold algorithms, with almost 3 percent F1 score and reduces the number of false negatives effectively.**

***Index Terms*—Federated Learning, Intrusion Detection System, FedAvg Algorithm, Scaffold Algorithm, F1-guided Softmax Aggregation.**

I. Introduction

With the rapid growth of the Internet of Things (IoT) devices, it boosted the number of cyber threats, meanwhile its need is to combat with the speed of the growth [1]. Internet of Things (IoT) brought in many attacks to this world, with its quick growth rate it requires quick response to combat in the real time. For this, Intrusion Detection System (IDS) has play a vital role of network security and safety. Tradional IDS accession is very hard to handle as it depends on centralized machine learning models which trains data locally on servers and then collect their sensitive and formal information at their central server to evaluate. This leads to a lot of issues that concerns about privacy, security and other regulatory compliances.

Federated Learning is one of the key concepts to fulfill these requirements in a decentralized learning system. In Federated learning, multiple clients take actively part in training of model but this training will ensure all data safety and security measures by sharing only parameters instead of sharing their raw data. However, despite of its privacy measures, this actually degrade the performance in real world scenarios of non-identically distributed clients which commonly known as non-IID data [2]. This issue is especially true with IDS with the changes in the nature of network traffic, attack distribution, and data imbalance with raw data.

Furthermore, the datasets of Intrusion Detection are genetically imbalanced, meaning that in most situations, the number of attacks is low compared to the number of normal traffic; with most of the content being usual network activities [3]. During these type of settings, traditional performance metrics of evaluation like accuracy can be misleading. Instead of using accuracy, we adapt F1-score which balance the precision-recall performance that makes it more reliable metrics of evaluation.

In order to tackle the instability, this paper present a new aggregation approach called F1-Guided Softmax Aggregation for Federated Intrusion Detection. The proposed method will help us to evaluate client by using f1-score and mold this score to aggregation weight in softmax function, which facilitates the global model to line up high quality client updates while keeping non-IID environment. This research outlines the gap between the FedAvg basic algorithm, Scaffold algorithm and the F1-Guided Softmax Aggregation for Federated intrusion detection system in non-IID environment based on the realistic scenario of training with CICIDS2017 dataset without relying on any intruder identify.

DOI: https://doi.org/10.67535/tsp.000003.001
Published under the ICACNC 2026 Proceedings

II. Related Work

The intrusion detection system is a solution that could be used to monitor network traffic and to be alerted to any suspicious traffic.Traditional IDS are signature based which can detect known attack patterns. The unusual behaviour is however anomaly based flags. Both methods are effective and have their pros and cons. It means that everything can be sent to one point in the network to be analysed [4]. In recent years, the operation of IDS was revolutionized by machine learning. Machine learning based systems don't need to be set once more, but they can be trained from data and may improve or change over time to meet a new threat. Currently, the most popular dataset that is used in this field of research is CICIDS2017.

Based on the deep learning method, CICIDS2017 has been illustrated by Jose and Jose (2023) and Kurniabudi et al. (2021) showed that selecting features is an effective method to increase the accuracy of detection [5]. The following studies illustrate the improvement in the performance of the IDS using ML. But there is still a problem with these models as they require a huge amount of centralized data. There is no way to gather all of the data without infringing on privacy.

To address privacy concerns, federated learning is addressed. Every client can train their own locally and upload a Model weights file to the central server, all models can be aggregated together and the aggregation of all models is a global Model. But, it is not necessary to share raw data [6].

Researchers are aware of the potential of applying FL to IDS. Lazzarini et al. (2023) used FL for intrusion detection in IoT networks with successful detection of threats without sharing raw data [7]. Fedorchenko et al. (2022) studied a comprehensive comparison of FL-based IDS approaches, which confirmed the advantages of FL-based approaches in terms of privacy and scalability [8]. Recently, Buyuktanir et al. (2025) carried out a survey of the developments of FL for IDS on different platforms [9].

In federated learning the biggest challenge is Non-IID data. This approach has different clients to hold data with different statistical properties. In intrusion detection systems, mostly one client sees normal traffic and others face different attacks [10]. This phenomenon causes local models to drift in different directions.

Non-IID data is the most challenging aspect of federated learning. This way, there are various clients to store data of various statistical property. Most of one client experiences normal traffic and other clients are attacked in various ways causing local drift in the case of intrusion detection systems [10]. Given the analysis by Zeng et al. (2023), the model divergence is observed for Non-IID distributions in FL, and the unstable convergence is observed in other distributions [11]. Moreover, Khudhair et al. (2025) applied FedAvg to the Non-IID CICIDS2017 dataset, and obtained that the FedAvg's performance rate is low in distribution and can cause low accuracy and slow convergence [12].

TABLE I
Comparison of Aggregation Methods

Method	Aggregation Basis	Handles Non-IID	Performance-Aware
FedAvg	Equal Averaging	No	No
Scaffold	Control Variates	Yes	No
FedProx	Proximal Term	Yes	No
Proposed Method	F1-Guided Softmax	Yes	Yes

As opposed to conventional approach focusing mainly on clients drift or optimization issues, a conventional approach is used in the proposed which explicitly incorporates the client detection performance in the aggregation process. Despite the fact that FedProx stabilizes training in heterogeneous environments, it does not take the client's performance into the aggregation. With the proposed F1-Guided Softmax Aggregation, the updates given at the client end are weighted by their F1-scores and hence the method is more effective in learning in the non-IID cases where there are imbalanced datasets.

III. Methodology

A simulation framework for our F1 guided- SoftMax aggregation is developed based on the real world network scenarios for provisionsally knowing its function under the condition of the heterogeneous data environment for federated intrusion detection system work. Our global model, which can be implemented on decentralized clients, was a light weight Multi-layer Perception (MLP). These layers comprise of an input layer with 80+ features of the dataset, two hidden layers with the ReLU activation and an output layer that displays the output values from 0 to 1 with 0 signifying a benign and 1 indicating an attack, thus, the basic pattern and higher-level abstractions are also noted. The methodology is based on these: Dataset, System Model, Client-Side Training, Server-side aggregation strategy. Make a Very Simple overview of the general model.

A. DataSet Description

The dataset used in this study is CICIDS2017, which is a popular dataset for intrusion detection systems (IDS). It includes 5 days of real world network traffic and attacks like DoS, DDoS, Brute Force, Botnet, Port Scanner, Web Attack. The dataset is made up of over 80 network traffic features in CSV files that are flow-based.

B. System Model

The suggested federated learning framework involves a central server and a number of clients that have local

traffic data sets. Clients train intrusions models locally without sharing raw data, records updates sent to the server, and combines the updates to form a global model. The data distribution is not IID which is an indication of real world situations and better collaborative IDS training.

C. Non-IID Data Partitioning

Data was non-IID partitioned in CICIDS2017 for setting up a realistic federated intruder detection setup. The training data was pre-processed and normalized, then split into three clients. The samples of benign and attack class were initially classified and then partitioned class-wise among the clients. A set of distinct subsets of the full dataset was given for each client, leading to a different distribution of local data characteristic of each client, but no loss of class representation.This partitioning is representative of realistic federated learning scenarios with various traffic and attack patterns observed across different clients.

D. Client-Side Training and Evaluation

Each client receives the global model and executes local training on their private datasets. To address class imbalance in intrusion detection system, class weights are applied and aggregated during training to ensure class balance learning between benign and attack. After Local training, each client figures out its model on validation subsets and compute its f1 score. The F1-score is chosen as basic evaluation metric due to its effectiveness in handling imbalance classification. To ensure numerical stability and avoid zero contribution from client, a lower bound is applied. Each client then sends their updated model weights with their computed f1-score to the server.

E. F1-Guided SoftMax Aggregation

The proposed technique allocates an adaptive weight to each client in contrast to FedAvg, which takes out uniform averaging updates from the clients. The first step in normalizing the F1-score obtained from the clients, using min-max normalization, to compare with other obtained F1-scores.

$$\tilde{f}k = \frac{f_k - f\min}{f_{\max} - f_{\min} + \epsilon} \tag{1}$$

In this equation,
f_k = F1 score of client$k$
ϵ = small constant to prevent division by zero

Using Softmax function, normalized scores are then transformed into aggregation weights using the temperature parameter (τ):

$$w_k = \frac{\exp(\tilde{f}k/\tau)}{\sum j = 1^K \exp(\tilde{f}_j/\tau)} \tag{2}$$

The temperature parameters are used to control the "better distribution" and with lower values of (τ) having higher performance clients.

Last, the overall parameters was updated by Using weighted aggregation:

$$W^{(t+1)} = \sum_{k=1}^{K} w_k \cdot W_k^{(t)} \tag{3}$$

Where;
$\mathrm{W}_k^{(t)}$ = local model weights of client(k) at round(t),

w_k = corresponding aggregation weight

F. Implemented Algorithms

The proposed framework is tested with FedAvg and Scaffold in non-IID intrusion detection problems. The major drawback of FedAvg is that it is simple and widely adopted for federated learning but it suffers from client drift, unstable convergence due to the fact that updates from all participating clients are equally averaged, irrespective of the quality of the updates.

Scaffold is a federated learning algorithm that aims to minimize the drift experienced by clients and at the same time uses control variates at both the server and client levels. The following correction terms ensure that the model is more likely to converge and less prone to instability when operating in Non-IID conditions, however they also introduce higher computational complexity and communication burden in training.

The proposed F1-Guided Softmax Aggregation method proposes adding weights for aggregation in Non-IID environments based on the clients' F1 scores. This method ensures that only high-quality clients are selected, mitigates the effect of low-quality updates, deals with data imbalance, and maintains the stability of federated learning performance.

G. Evaluation Metrics

The proposed federated intrusion detection framework evaluation is made using several metrics, such as accuracy, precision, recall and the f1-score. These metrics is suitable when discussing the correctness of model when working with imbalanced distribution of data.

- Accuracy: The overall correctness of model. [13]

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

- Precision: The proportion of correctly predicted attacks among all instances. [14]

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

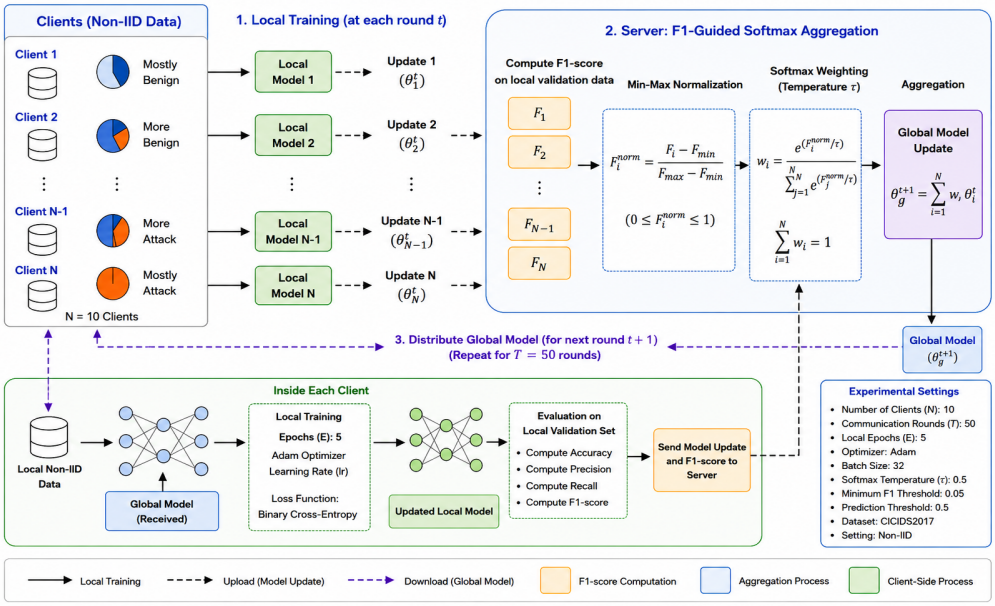

Fig. 1. Model Preview of F1 Guided Softmax Aggregation

• Recall: The proportion of actual attack instances that are correctly identified by model. [15]

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

• F1-Score: The harmonic mean of precision and recall, provides balanced evaluation of both metrics. [16]

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

IV. Results and Discussion

The simulation was performed on the 7th Generation Intel Core i7 HQ on high performance workstation with 16GB of RAM. The experiments were realistically implemented with python in a controlled environment called virtual environment (venv).

TABLE II
Experimental Hyperparameters

Parameter	Value
Dataset	CICIDS2017
Number of Clients	10
Communication Rounds	50
Local Epochs	5
Batch Size	32
Optimizer	Adam
Learning Rate	0.001
Hidden Layer 1	64 Neurons
Hidden Layer 2	32 Neurons
Activation Function	ReLU
Output Activation	Sigmoid
Validation Split	20%
Softmax Temperature (τ)	0.5
Random Seed	42

The experimental setup used in the study was summarized in Table-II. The proposed method for computing the aggregation uses a softmax temperature $\tau = 0.5$, which is found to be a good compromise between the need to highlight high-quality clients and to have a stable aggregation.

A. Confusion Matrix for Algorithms

• Confusion Matrix for FedAvg

To give a better understanding of the types of errors, the confusion matrix in the Figure 3 is provided for the FedAvg model.

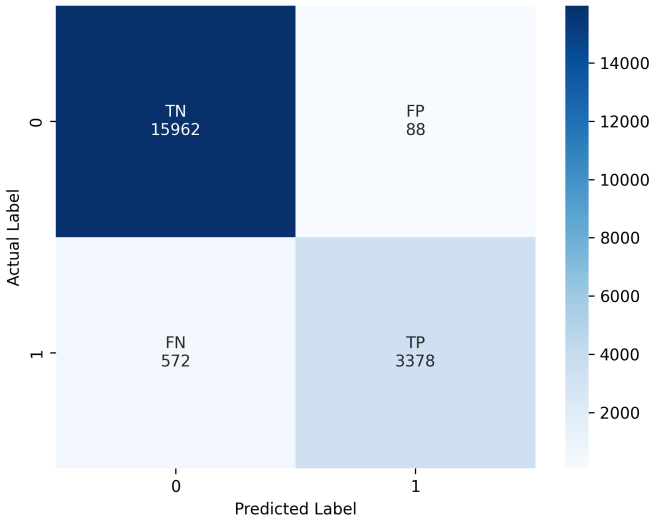

Fig. 2. FedAvg Confusion Matrix

FedAvg has a considerable number of false positives (88) and importantly, 572 false negatives meaning that whenever trying to detect a real attack, the model has stumbled by missing that attack for more than 650 attacks.

- Confusion Matrix For Scaffold

The confusion matrix for Scaffold also shows many false positives (81) and moreover, 554 false negatives, which means it also missed more than 630 real attacks.

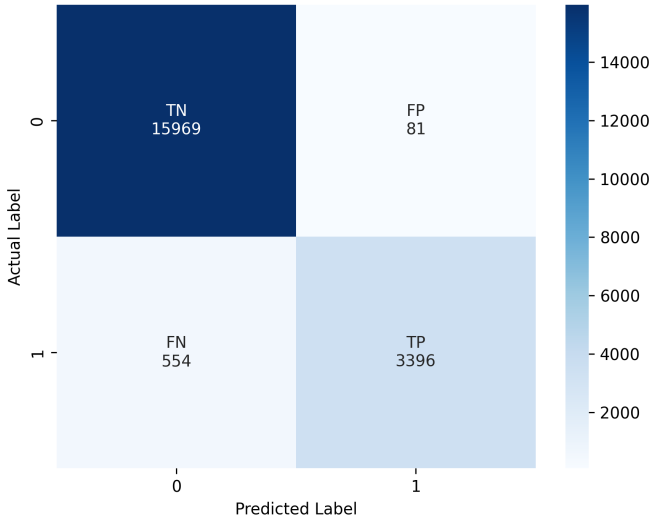

Fig. 3. Scaffold Confusion Matrix

- Confusion Matrix For F1-guided Softmax

The confusion matrix of F1-guided Softmax has been implemented and gaves us a more in-depth analysis of error type. And also detect more attack as compared to others.

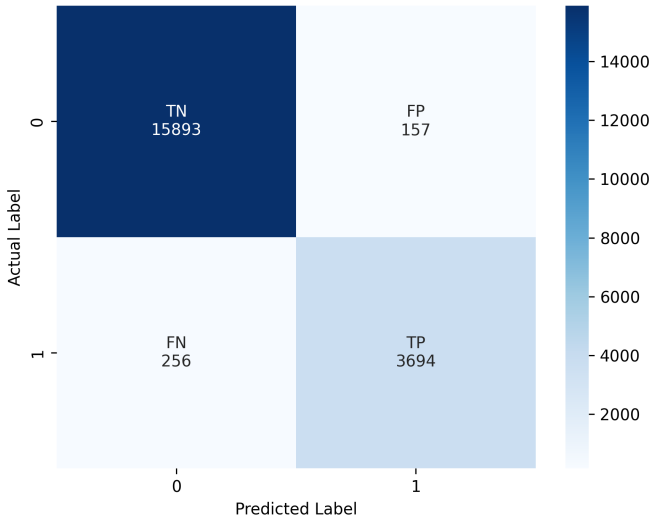

Fig. 4. F1-Guided Softmax Confusion Matrix

- Ablation Analysis

The ablation study illustrates the effectivness of the proposed framework from both components. Softmax weighting results in performance over FedAvg and F1-based weighting makes attack detection further with performance. An integration of both mechanisms cements its effect, showing the astonishing benefits of the proposed strategy of F1-Guided Softmax Aggregation.

TABLE III
Ablation Study of the Proposed Aggregation Method

Method	Accuracy (%)	Recall (%)	F1-score (%)
FedAvg	96.7	85.5	91.1
Softmax Only	97.1	88.9	92.4
F1 Only	97.4	90.8	93.2
F1 + Softmax (Proposed)	97.9	93.5	94.7

V. Performance Comparison

The CICIDS2017 dataset is used for evaluating the performance of proposed framework in a non-IID setting. Each client has his own subset of data, to check if there is any difference between the clients (N=10) Experiments are run for 50 communication rounds (T=50) run on a local training pool, 5 epochs per round (with Adaptive Moment Estimation(Adam) optimizer). Finally, the proposed approach considered is specially effective for intrusion detection in non-IID scenarios which is typically found in the real world as it provides superior number of accurate detection (better true positive rate and lower false negative rate) than FedAvg and Scaffold.

TABLE IV
Performance Comparison

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
FedAvg	96.7	97.5	85.5	91.1
Scaffold	96.8	97.7	86.0	91.5
Proposed F1-Guided Softmax	97.9	95.9	93.5	94.7

Table IV indicates that the method proposed in this paper is the best overall method of all the methods evaluated. Specifically, when performing an attack traffic detection task under imbalanced and non-IID data, the proposed approach showed the improvement in recall and F1-Score.

Firstly a convergence behaviour verification of the compared methods is done in the same context of heterogeneous environment so as to understand their stability. The results clearly depicts the difference between baseline FedAvg method, Scaffold algorithm and proposed F1-Guided Softmax Aggregation approach.

From the results, it is obvious that the proposed method has a better Trade-off between precision and recall rate that gives a high value of F1-Score. The model allows for handling the priorities based on the performance of F1 in the specific locations they occupy, effectively handling minority attack patterns and enhancing the ability to

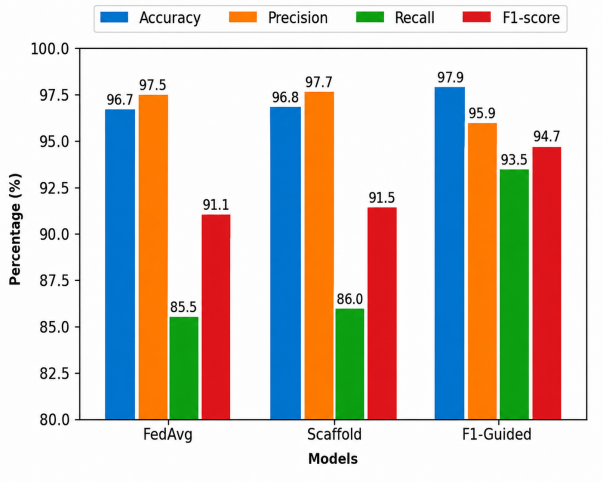

Fig. 5. Comparison Model Performance on Non-IID Dataset

detect.

VI. Conclusion

This paper suggested a F1-Guided Softmax Aggregation approach to Non-IID Federated intrusion detection. The proposed approach differs from previous ranking methods like FedAvg and Scaffold, which do not take into account task-specific performance, but rather the F1 score is used to incorporate this as part of the ranking process. This is done in a manner that helps add weighting to every client update based on the quality of the update, and that can respond to differing (and imbalanced) data distributions.

Experimental results obtained while testing on CICIDS2017 data set show that the proposed method can outperform the recalls by achieving a positive gain of 6 percent or better than both on false negative and true positive rate, which are crucial parameters for intrusion detection system. The proposed method is more straightforward than other complex methods, such as Scaffold; it provides better precision/recall rate with steady convergence like other advanced methods. Results show the strength of the performance aware aggregation in federated learning and its use in real world cyber security tasks.

VII. Future Work

Potential future steps for the project may include testing the proposed protocol in next larger FL networks with different number of clients, building it up to scale and transferring it to the real world application. The framework is applicable for testing the data distribution for the non-IID case (for example Dirichlet partitioning case) and can also be applied to test its survivability in realistic network scenarios. Moreover, it is claimed that the improvement of aggregation process can be made by incorporating supplemental performance metrics, such as loss prediction confidence in order to make aggregation process more effective to have an effective hybrid weighting method. Furthermore, deploying approaches respecting privacy and more efficient in communication was to make the framework appear more suitable for a large scale and real world deployment of intrusion detection systems.

References

[1] W.-C. Chung, C.-A. Lo, Y.-H. Lin, Z.-H. Chen, and C.-L. Hung, "Decentralized federated learning with non-iid data: Challenges, trends, and future opportunities," *ACM Computing Surveys*, vol. 58, no. 8, pp. 1–41, 2026.

[2] A. Karunamurthy, K. Vijayan, P. R. Kshirsagar, and K. T. Tan, "An optimal federated learning-based intrusion detection for iot environment," *Scientific Reports*, vol. 15, no. 1, p. 8696, 2025.

[3] M. Devi, P. Nandal, and H. Sehrawat, "Federated learning-enabled lightweight intrusion detection system for wireless sensor networks: A cybersecurity approach against ddos attacks in smart city environments," *Intelligent Systems with Applications*, vol. 27, p. 200553, 2025.

[4] M. A. Bilal, I. Ul Islam, S. Idrees, M. Qasim, M. J. Khan, and J. Khan, "Dataset-centric evaluation of federated intrusion detection models in iot networks," *Scientific Reports*, 2026.

[5] K. Kurniabudi, D. Stiawan, D. Darmawijoyo, M. Y. B. Idris, B. Kerim, and R. Budiarto, "Important features of cicids-2017 dataset for anomaly detection in high dimension and imbalanced class dataset," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 9, no. 2, pp. 498–511, 2021.

[6] R. T. Khudhair, G. C. Hock, and A. B. Abu-Bakar, "A comparative analysis of federated learning on non-iid data for an intrusion detection system," *Emerging Trends in Engineering and Sustainability*, vol. 1, no. 1, pp. 37–44, 2025.

[7] R. Lazzarini, H. Tianfield, and V. Charissis, "Federated learning for iot intrusion detection," *Ai*, vol. 4, no. 3, pp. 509–530, 2023.

[8] E. Fedorchenko, E. Novikova, and A. Shulepov, "Comparative review of the intrusion detection systems based on federated learning: Advantages and open challenges," *Algorithms*, vol. 15, no. 7, p. 247, 2022.

[9] B. Buyuktanir, Ş. Altinkaya, G. Karatas Baydogmus, and K. Yildiz, "Federated learning in intrusion detection: advancements, applications, and future directions," *Cluster Computing*, vol. 28, no. 7, p. 473, 2025.

[10] T. D. Nguyen, A. Alazab, A. Khraisat, and T. Jan, "Feature reduction in federated learning for intrusion detection in iot networks," *Cybersecurity*, vol. 9, no. 1, p. 102, 2026.

[11] D. Zeng, S. Liang, X. Hu, H. Wang, and Z. Xu, "Fedlab: A flexible federated learning framework," *Journal of Machine Learning Research*, vol. 24, no. 100, pp. 1–7, 2023.

[12] K. M. Khudhair, R. H. Jabbar, M. H. Hasan, D. K. Muhsen, A. F. Kreah, and A. H. Najim, "Behavioral biometrics-based intrusion detection in online banking using lstm networks," in *2025 International Conference on Electrical Engineering and Informatics (ICEEI)*. IEEE, 2025, pp. 1–7.

[13] 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 Computing Surveys*, vol. 57, no. 1, pp. 1–38, 2024.

[14] U. Islam, F. Din, A. U. Haq, and I. Ali, "Federated security framework: A heterogeneous federated learning architecture for privacy-preserving intrusion detection in iot networks," *Transactions on Emerging Telecommunications Technologies*, vol. 37, no. 4, p. e70402, 2026.

[15] A. Ubaida, K. Z. Haider, M. A. Rasheed *et al.*, "Intrusion detection based on federated learning–a review," *Spectrum of Engineering Sciences*, vol. 4, no. 4, pp. 372–393, 2026.

[16] M. A. Q. Riyadi and A. M. Dewi, "Federated learning for privacy-preserving iot intrusion detection under extreme non-iid conditions," *Jurnal Ilmu Komputer dan Informasi*, vol. 19, no. 1, pp. 107–120, 2026.