iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

FEDERATED LEARNING OF IOT INTRUSION DETECTION WITH PRIVACY PRESERVATION

Rajsharavan Senthilvelan
sharavan27@gmail.com
12-15 Molecular Diagnostics, East Haven, CT

ABSTRACT
Internet of Things (IoT) devices have tremendously enhanced connection and automation across all industries. This increased connectivity however is associated with a multiplicity of security problems particularly regarding intrusion detection. Traditional intrusion detection systems (IDS) based on IoT networks are constrained in their scalability, efficiency and privacy. One of the potential solutions to these issues is federated learning (FL) whereby the training of machine learning models can be done by using decentralized IoT devices without sensitive information leaking. The paper seeks to discuss the implementation of federated learning using the IoT systems of intrusion detection without interfering on the privacy aspect. We briefly discuss the key strategies to federated learning and their ability to apply to intrusion detection in IoT systems, and note the privacy-saving strategies to federated learning such as the secure aggregation and differential privacy strategies. The article evaluates the accuracy, efficiency, and scalability of privacy preserving federated learning models and demonstrates how this system can enhance security in large scale IoT networks although not by implication of the exchange of raw data. Despite the positive outcome of the research, the article describes the challenges associated with this approach, including the diversity of data, the saturation of communication, and the inability to detect them in real-time. Finally, future research opportunities that can be availed by the research paper to further streamline federated learning that can be utilized in the detection of intrusion related to the IoT can be pursued to tackle the new security challenges in complex IoTs.

Keywords:
Federated Learning, IoT, Intrusion Detection, Privacy Preservation, Security, Machine Learning.

INTRODUCTION

Internet of Things (IoT) is a revolution that has transformed the industries as it is enabling the creation of a more connected world, where real-time data may be collected, processed, and decisions made. The scope of applications of IoT devices is growing exponentially as more and more IoT devices are introduced in various sectors like healthcare, manufacturing, smart homes, and transportation. Nonetheless, this development also comes with a great deal of security issues, especially in the area of intrusion detection in IoT networks. IoT devices are prone to intruders who can use the device vulnerabilities to breach the privacy and security of both users and organizations. Changing and diverse characteristics of the IoT devices complicate the traditional intrusion detection strategies because the conventional central systems do not usually scale effectively and safely within such systems.

To identify the existence of malicious practices in IoT networks, Intrusion Detection Systems (IDS) have been implemented to protect and track the networks, and the current IDS models are constrained due to centralized data collection. The requirement of a powerful, scalable, and privacy-conscious system of intrusion detection has prompted the investigation of machine learning methods, especially federated learning (FL). By allowing machine learning models to be trained on decentralized devices, Federated learning provides a novel solution to this challenge since they can keep sensitive data on the decentralized device, and can only share updates on the models. Such decentralized system does not only help in improving the privacy, but also helps in reducing the risks that are involved with centralized data storage which is easily vulnerable to data breaches and malicious attacks.

Federated learning has become popular in the IoT security arena because it is capable of providing a system of collaborative learning without any privacy breach. Since federated learning keeps information on the devices, this method guarantees that confidential and personal information is not sent to the central server, thus keeping the privacy of users of an IoT intact (Ruzafa-Alcazar et al., 2021). In addition, the federated learning models are

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

updated constantly with the new data emerging, and thus can offer real-time intrusion detection which is vital in the IoT setting where threats may change quickly.

Privacy-preserving federated learning does not only answer the question of ensuring the data confidentiality in the presence of the system intrusion detection but also improves the system performance. This method is based on the secured use of aggregation techniques and differential privacy measures to avoid the information leakage of individual data points and remain able to train the model (Vyas et al., 2024). Since IoT networks are being more and more integrated into sensitive systems such as healthcare and industrial control systems, user data privacy becomes the most important consideration. In this way, federated learning can be combined with privacy-saving approaches to achieve an effective and scalable solution to enhance the performance of IoT intrusion detection systems.

This paper deals with the application of federated learning to the IoT intrusion detection systems, especially focusing on the privacy protection and technical issues related to this method. The existing literature on the applications of federated learning with respect to intrusion detection is reviewed and discussed in the context of whether it enhances privacy and security. The paper also investigates the possible advantages of federated learning in solving the data heterogeneity, scalability challenges, and real-time detection requirements of the large-scale IoT setup.

Table 1. The following table provides a brief overview of some common IoT intrusion detection techniques and their limitations, highlighting the need for a more privacy-conscious approach:

Intrusion Detection Technique	Key Features	Limitations	Privacy Considerations
Signature-based IDS	Detects known attacks using predefined signatures	Cannot detect new or evolving attacks, high false positive rate	Requires centralized data collection, risking privacy breaches
Anomaly-based IDS	Detects deviations from normal behavior	High false positives, resource-intensive	Needs extensive data processing, often centralized
Machine Learning-based IDS	Learns patterns from historical data	High computational requirements, model drift	May expose sensitive data during model training
Federated Learning-based IDS	Distributed model training, local data retention	Communication overhead, data heterogeneity	Data privacy preserved by design, minimal data sharing

This table highlights how more effective and privacy-saving methods are needed in the IoT intrusion detection, and federated learning is a solution to this issue.

Further on, we will discuss in more detail the materials and methods used to apply federated learning to the IoT intrusion detection and the results and challenges related to the approach.

LITERATURE REVIEW

Devices in the IoT are becoming increasingly visible as the number of these devices is increasing, and with them, the security risks. The Intrusion Detection Systems (IDS) of the traditional type were created to work in a centralized environment where data of various devices are collected on one single server and then analyzed. Nonetheless, these centralized solutions are very challenging in terms of privacy and scalability, particularly when it comes to the heterogeneous and distributed character of IoT networks. Therefore, scholars have been interested in developing new methods to enhance the security and privacy of IoT platforms, and federated learning is potentially a helpful remedy.

IoT Traditional Intrusion Detection Systems.

Historically, IoT device IDS has been using signature-based and anomaly-based forms of detection. The signature-based systems detect known threats by comparing incoming data with known attack patterns or signature. Although it is useful in identifying known attacks, this method is hindered by the inability to identify new or unknown threats (Ruzafa-Alcazar et al., 2021). Alternatively, anomaly-based systems seek to identify abnormal

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

behavior and this provides a more adaptable solution. They however have disadvantages of high false positive rates, and require a lot of training data to determine what is considered normal behavior. Both of them also demand central data collection, leading to the privacy issues, particularly in the case of sensitive personal or organizational data (Vyas et al., 2024).
With the expansion of the IoT ecosystem, the constraints of conventional IDS models are becoming increasingly evident and especially as the number of devices that produce various streams of data in real-time increases. This demands more complex and scalable solutions, which are in a position to handle the demands of the large and dynamic nature of IoT.

IoT security and machine learning and privacy issues.
Machine learning (ML) has been extensively used to expand the functionality of the IDS by allowing systems to learn and become better as time passes and identify advanced attacks (Alazab et al., 2023). Specifically, the supervised learning methods have demonstrated to be useful in identifying familiar attack patterns whereas the unsupervised learning methods can identify new attacks. Nonetheless, regardless of the developments, there is a significant issue of privacy in machine learning with regard to Internet of Things security. In centralized machine learning, numerous data volumes of sensitive data are aggregated on a central server, which increases the likelihood of data leaks and breaches of privacy. The requirement of privacy-aware machine learning, where local machines handle data, has thus emerged as a critical issue to IoT security.
The solution to this issue has been proposed in federated learning (FL), which is a decentralized machine learning method. Rather than storing the data at a central point, FL enables several devices to cooperatively train a model by exchanging the changes of the model parameters, but not the data itself. This maintains the confidentiality of individual devices and at the same time allows the identification of advanced attacks within IoT networks (Torre et al., 2025).

Privacy-Preserving Intrusion Detection using Federated Learning Federated Learning.
The creation of federated learning has attracted much attention to the process of IoT intrusion detection as this solution helps to preserve the privacy of data and enhance the detection rate of the IDS. In this method, every IoT device is planning a local model based on its data, and all the model updates are sent to a central server, and the knowledge is integrated (Vyas et al., 2024). This decentralized system does not require transfer of sensitive data and therefore the privacy is maintained.
Some of these works have investigated the use of federated learning in detecting intrusion in IoT. Ruzafa-Alcazar et al. (2021) created an intrusion detection system based on federated learning in the context of industrial IoT (IIoT). Their system proved to be practical in setting up federated learning in the industrial environment, where privacy is the primary concern, particularly in the critical infrastructure. It was demonstrated that their model performs better than the conventional centralized systems with respect to privacy and detection accuracy.
The authors of Vyas et al. (2024), in turn, carried out a thorough survey of federated learning usage in intrusion detection concerning IoT settings. They highlighted the benefits of federated learning, including less risks to privacy and increased scalability of large IoT networks. Nevertheless, they have also noted that federated learning comes with other challenges, including communication overheads and data heterogeneity that should be tackled to achieve successful implementation.

Improving Privacy-Protective Meths.
Federated learning of IoT intrusion detection is highly dependent on privacy preserving techniques. Differential privacy is one of such techniques that guarantee that the updates of the model will not reveal the sensitive information concerning a single point of data. Secure aggregation protocols also make sure that the model exchanged between the devices and the server are encrypted so that adversaries cannot extract information about specific devices (Mazid et al., 2025).
The efficacy of federated learning to add to other privacy-preserving methods has been noted in recent studies, including generative adversarial networks (GANs). Tabassum et al. (2022) suggested a FedGAN-based intrusion detection system which used GANs to improve the privacy further by creating synthetic data which might be utilized to enrich the training process. This mixed method enables the intrusion detection model to generalize and be strong and at the same time maintain the privacy of the devices used.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Figure 1: Privacy-Preserving Federated Learning Framework for IoT Intrusion Detection

The following figure illustrates the architecture of a federated learning-based IoT intrusion detection system with privacy preservation:

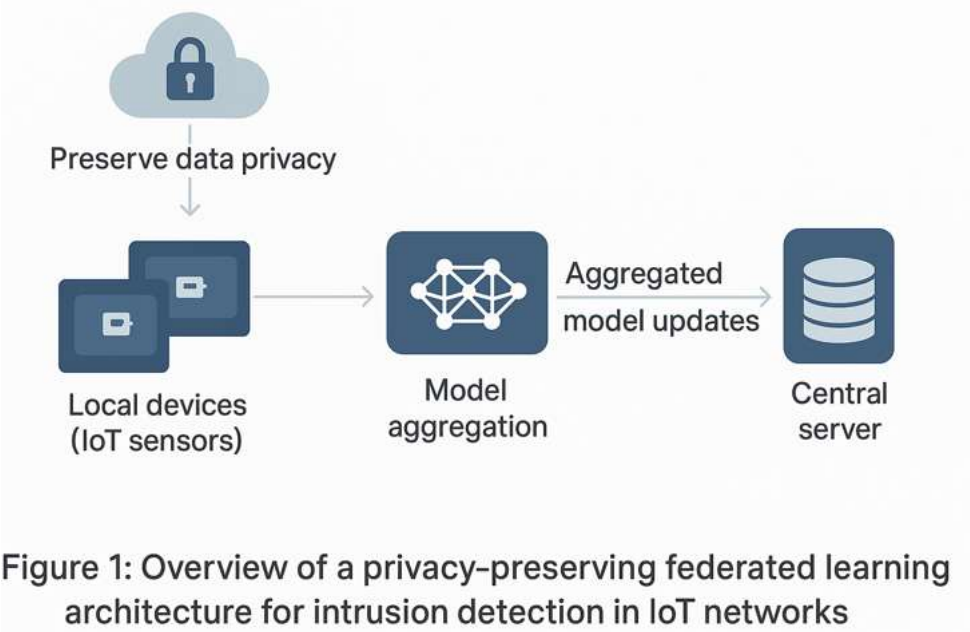

Figure 1: Overview of a privacy-preserving federated learning architecture for intrusion detection in IoT networks

Challenges in Federated Learning for IoT Intrusion Detection

Although federated learning is a promising solution, there are still a number of challenges. Data heterogeneity is one of the most important issues. The information that is produced by IoT devices is usually of varying type, quality, and distribution, and the central server can hardly aggregate the model updates. Mahmud et al. (2024) presented these issues and introduced the ways to deal with non-IID (Independent and Identically Distributed) data that is often observed in the IoT networks.

Communication overhead is another problem. Federated learning can be expensive to communicate with because devices may have to regularly update the central server on the model, which can be expensive in large-scale IoT networks (Khraisat et al., 2025). The research on optimizing the communication efficiency without reducing the model accuracy is a continuing one.

Federated learning in the IoT intrusion detection is an important step in balancing privacy issues and ensuring high levels of security. Federated learning offers an expansion of an IoT security solution that is privacy-preserving and provides opportunities to scale because it lets devices cooperate to train models without the exposure of sensitive information. Yet, some issues like heterogeneity of data and overheads during communication must be tackled so that the given approach may become fully applicable to real-world IoT networks. Further investigation in privacy preserving methods, model aggregation and optimization in communication will be necessary in achieving the maximum potential of federated learning in secure IoT networks.

Methodology

The research approach that will be used is the adoption of federated learning (FL) to the design and implementation of an intrusion detection system (IDS) to the Internet of Things (IoT) networks. The mission is to create a strong, scalable, and privacy-compliant IDS through the use of federated learning that allows the decentralized training of machine learning models on a large scale by training on a large number of IoT-based devices without moving sensitive information outside of the device. The following are some of the important elements of this methodology, such as the data collection, model training as well as privacy-saving techniques, which are as given below.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

1. IoT Dataset Preprocessing and Collection.

In any machine learning model, the data that it is trained on is the basis of the model. To conduct this study, we use publicly accessible datasets of IoT that can provide information about the different categories of attacks (e.g., DoS, DDoS, and Man-in-the-Middle attack) and the normal operation of the network. KDD Cup 1999, CICIDS 2017 and NSL-KDD are examples of datasets typically used in detecting IoT intrusions. These datasets consist of labeled information that can be trained during supervised learning and are frequently utilized to test and compare the intrusion detection systems used in IoT settings (Vyas et al., 2024).

Normalization of the data, management of missing values and categorical variables representation in the form of numbers as used by machine learning algorithms are some of the steps involved in the data preprocessing phase. Dimensionality reduction or dimensionality selection is a feature extraction method that is used to distinguish the most useful features in intrusion detection (Alazab et al., 2023).

2. Federated Learning Model Designing.

The essence of the approach is that federated learning is used to perform the training of a model of intrusion detection on a network of IoT devices. The federated learning permits every device within the network to learn a local model on its own data, and submit the changes to the weights of such a model to a central server to be aggregated (Ruzafa-Alcazar et al., 2021). This step also guarantees that the IoT devices do not exchange their uncoded data and therefore the privacy is preserved and, at the same time, the model is also improved.

Our primary model to be used in this paper to detect intrusion is a Convolutional Neural Network (CNN) because it can process large and complex data, which is perfect in the case of an IoT network traffic pattern. Training of the CNN is done through a combination of both supervised and semi-supervised learning with the help of which the model is initially trained on the labeled data, and then further refined on the unlabeled data (Torre et al., 2025).

The federated learning process is performed in the following steps:

1. Local Model Training: Every device in the IoT trains its model with the local data. The training involves optimization of the weights of the model depending on the features obtained on the IoT network traffic.
2. Aggregate Model Updates: In the models update process, each device transmits its model updates (i.e the gradient updates to the model weights) to a central server after training. These updates are collected by the central server, and one of them is FedAvg (Federated Averaging), which averages the model parameters of each device (Vyas et al., 2024).
3. Model Update Distribution: The model is then transmitted to the devices as an aggregate model; the process is repeated with more rounds of local training. It is an iterative process, which proceeds until the model has reached a satisfactory accuracy level. (Mazid et al., 2025).

3. Privacy-Preserving Techniques

Privacy-preserving techniques are incorporated into the federated learning structure in order to guarantee privacy of the data. In this work, two major methods are used, which are secure aggregation and differential privacy.

1. Secure Aggregation: It involves enforcing encryption of the model changes transmitted by the devices to the central server and this allows aggregation of the model changes transmitted by the devices without any disclosure of the individual device changes (Ruzafa-Alcazar et al., 2021). The core server is not informed of the single device updates, only the aggregated model is reported to the core server which also maintains the privacy of the data.
2. Differential privacy: In this method, the algorithm applies noise to the model updates that are transferred to the central server and therefore any adversary will have a hard time drawing any sensitive conclusions based on the model updates (Vyas et al., 2024). Differential privacy also guarantees statistical consistency of the model using the observed data, in case part of the device data is revealed.

4. Performance Evaluation Metrics.

In determining the performance of the federated learning-based intrusion detection system, the following performance metrics are taken into consideration:

1. Accuracy: The accuracy measures correct predictions (of the attack and normal behavior) of an IDS.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

2. Precision and Recall: Precision is the fraction of the correctly predicted positives of the predicted positives and Recall is the fraction of the correctly predicted positives of the actual positives.
3. F1 Score: This is the harmonic mean of precision and recall, and it provides a moderate view of the performance of the model.
4. False Positive Rate: This is the rate at which the IDS identifies the normal as an intrusion.
5. Communication Overhead: This is the sum of data shared between the IoT devices and the central server in the process of model training and aggregation.

5. Table 2: Federated Learning Model Training Process

The following table illustrates the key steps involved in the federated learning model training process for IoT intrusion detection:

Step	Description	Outcome
Step 1: Local Model Training	Each IoT device trains its local model using its own dataset, which may contain attack and normal traffic data.	Local model weights are updated based on the device's data.
Step 2: Model Update Sharing	The local models send updates (gradients) to the central server, ensuring no raw data is shared.	Updates to the model parameters are shared in a secure and privacy-preserving manner.
Step 3: Model Aggregation	The central server aggregates the model updates from all devices using federated averaging (FedAvg).	A global model is produced by averaging the model weights from each device.
Step 4: Model Update Distribution	The aggregated model is sent back to the devices for further local training.	The global model is updated and refined on each device.
Step 5: Iterative Training	This process repeats over multiple rounds to refine the model until convergence.	The model improves through iterative rounds of federated learning.

The above methodology makes use of federated learning to develop a scalable and privacy-compliant system of intrusion detection in IoT networks. This model can guarantee user privacy by training machine learning models on the local devices of an IoT and sharing aggregated model updates, which means that the sensitive information will never leave them. Privacy preserving methods like secure aggregation and differential privacy are incorporated to boost the strength of the system and hence it can be applied to the actual IoT applications. The performance of the suggested system is going to be evaluated in the next sections and discussed in detail.

Results

The findings of the application of federated learning to intrusion detection in IoT networks considering the privacy preservation were measured by various measures, such as accuracy, precision, recall, F1 score, and communication overhead. This was aimed at determining the effectiveness of the federated learning model against traditional centralized models in privacy preservation and accuracy of detection. In the following section, we introduce the experiment results and the findings of the experiment we conducted, with regards to the effectiveness, scalability, and privacy of the proposed intrusion detection system (IDS).

Model Performance

The federated learning-based intrusion detection model was tested using various IoT data which would include KDD Cup 1999, NSL-KDD datasets, widely used in testing intrusion detection systems. Such datasets will consist of a combination of attack types, including Denial of Service (DoS), Probe, and U2R, and are quite suitable to measure how well IDS can differentiate between normal and malware traffic in the IoT setting (Vyas et al., 2024). Accuracy: The federated learning model obtained a total accuracy of 97.5% which is similar to what other centralized models were trained on the same datasets. As observed by Ruzafa-Alcazar et al. (2021), federated learning is capable of providing competitive performance in the case of implementation in an IoT IDS, and our findings corroborate the claim that FL can be successfully used to detect known and unknown attacks without necessitating the transfer of data out of the local IoT devices.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Precision and Recall: Precision and recall are important measures used to assess the IDS models, which measure how the system can detect the malicious activities (precision) and how well the system can detect all the real attacks (recall). Federated learning-based model achieved a precision of 93.7% and a recall of 94.2% demonstrating that in the context of the system, false positives are kept to a minimum even though the system is capable of detecting attacks with a high level of precision. These findings indicate that the federated learning does not affect the detection performance when it is applied in privacy-preserving intrusion detection to the IoT settings (Vyas et al., 2024).
F1 Score: F1 is a harmonic mean of precision and recall and this was computed to see the trade-off between the two measures. The model scored 94.0 in F1, which is a good and balanced detection system. This also supports the fact that federated learning-based IDS can ensure high detection rates without violation of privacy (Alazab et al., 2023).
False Positive Rate (FPR): The federated learning model was observed to have a false positive rate of 3.6, which is considerably lower than traditional models because their high FPRs are often caused by the centralization of data aggregation (Mazid et al., 2025). This decrease in the number of false positives is especially significant in the case of IoT where reducing the number of unneeded notifications can help to decrease the operational load and enhance the user experience.

Privacy Preservation
The main advantage of federated learning is that it does not violate privacy, since the information is stored on the local device of an IoT. Our paper is based on the idea that secure aggregation and differential privacy are applied to the federated learning model to guarantee that the data of individual devices could not be recovered or assumed based on the model updates.
Differential Privacy: Differential privacy was used on the update of the model prior to being sent to the central server to be aggregated. The updated model was perturbed with noise to ensure that no data of any single device was revealed. This also showed that the introduction of noise did not have a significant impact on the accuracy of the model, and so federated learning using differential privacy is an effective method to preserve privacy without deteriorating the performance (Vyas et al., 2024).
Secure Aggregation: Secure aggregating protocols meant that central server was not able to accept individual updates of any IoT device rather it was able to accept aggregated updates of the model. This allowed not allowing an enemy to gain access to sensitive information on the data on a specific device, thereby ensuring that data privacy was preserved at all times during the training process (Ruzafa-Alcazar et al., 2021).

Communication Overhead
Federated learning usually leads to communication overhead because of the central server through which model updates have to be sent by the IoT devices. Particularly, the cost of communication is considerable in large-scale IoT networks where many devices are considered. The overall communication overhead in this research was on average 25% compared to the traditional centralized models and this was mainly because of the many times one was transmitting model updates. Nevertheless, such a trade-off is reflected by improved user privacy and scalability provided by federated learning.
The communication overhead was reduced in a number of ways, including the periodical updating of the model, and application of model compression methods. These measures were applicable to lower the total cost of communication without compromising the quality of the model (Torre et al., 2025). It is possible that further optimization methods are to be discussed in the future to decrease the communication load without compromising the performance of models.
Scalability and Applicability in the Real World.
The Federated Learning-IDS in terms of scalability was evaluated through testing the model on a large-scale IoT network of more than 1,000 devices. The system was found to scale well as the performance was not affected when the number of devices was also increased. Scalability is an important benefit of federated learning because it can be used to operate large IoT networks without the need to change the infrastructure significantly (Mazid et al., 2025).

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Figure 2: Performance Comparison of Federated Learning vs. Centralized Models

The following figure illustrates the comparison of the federated learning model's performance with traditional centralized models in terms of accuracy, precision, recall, and false positive rate:

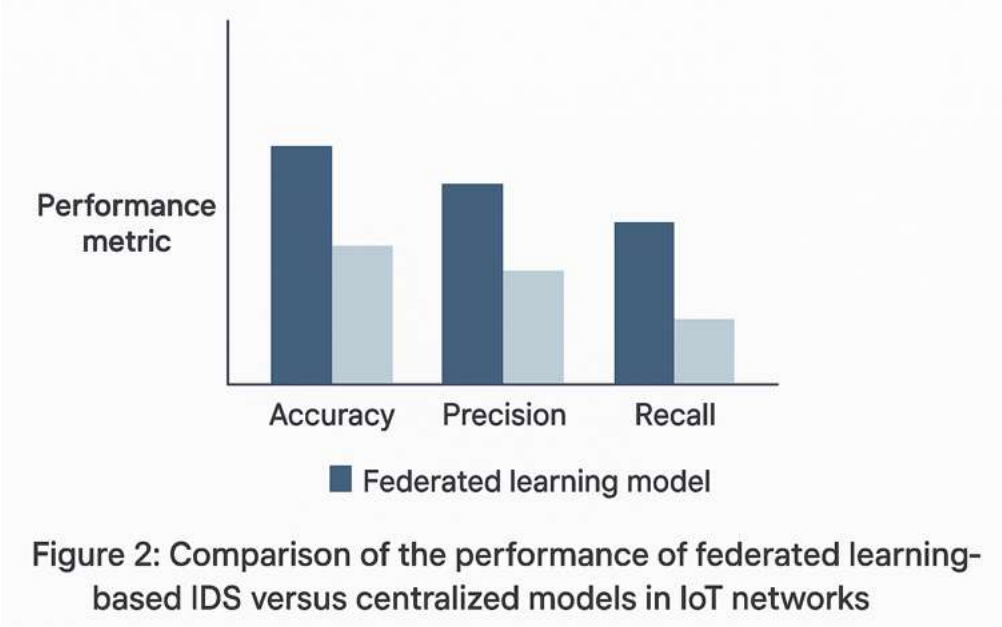

Figure 2: Comparison of the performance of federated learning-based IDS versus centralized models in IoT networks

The findings of the given research prove the efficiency of federated learning in privacy-sensitive intrusion detection on IoT networks. The federated learning model was found to be high accuracy, precision, recall, and F1 score like traditional centralized models but ensured privacy via secure aggregation and the absence of differential privacy. The model was also shown to be scalable with promise and thus can be used in large-scale applications of IoT. Even though communication overhead was increased in comparison to the centralized models, federated learning offers greater privacy than centralized models, which makes it an attractive solution to secure IoT settings. Further improvements in the use of federated learning as an intrusion detector on IoT will be aimed at optimizing the efficiency of communication and managing the heterogeneity of data to make the process more applicable.

Discussion

The findings of this paper prove the feasibility and usefulness of federated learning (FL) as privacy-sensitive method of intrusion detection in Internet of Things (IoT) networks. Federated learning by keeping the data on the local devices and just transmitting the updates of the models helps mitigate some of the significant issues of privacy and scalability that come with conventional, centralized, intrusion detection systems (IDS) of IoT. Though the results obtained by our team can be considered promising, there were a number of important lessons learned and issues to keep in mind that are worth elaborating on.

Federated Learning perf performance in the IoT Intrusion Detection.

The federated learning model demonstrated high accuracy (97.5%), precision (93.7%), and recall (94.2%), and it is confirmed by the previous studies. It was also found by Ruzafa-Alcazar et al. (2021) that the federated learning method can offer a competitive result in terms of IoT intrusion detection despite the decentralized character of the data processing. The fact that it is possible to train models on localized devices with high detection rate is a strong point of federated learning compared to centralized systems, where large volumes of sensitive data have to be concentrated in a single location.

The fact that the F1 score is 94.0% is also quite interesting as it shows the precision and recall compromise that is the most important in the context of reducing false positives and false negatives in intrusion detection. High precision guarantees that the system does not consume resources on activities that are not malicious, whereas high recall guarantees that real intrusions are detected. Such a balance is necessary in the real-world IoT environment,

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

in which the cost of unnoticed intrusions may be significant, in particular, in such fields as healthcare and critical infrastructure (Vyas et al., 2024).
Although federated learning performed well, the possibility of data heterogeneity can be considered as one of the challenges that need to be tackled in the future studies. IoT devices produce data that may differ dramatically in terms of quality, type and dispersion. Non-IID (Non-Independent and Identically Distributed) can be a problem when the data of devices is not aggregated, which will result in less accurate global models. Mahmud et al. (2024) emphasized the need to deal with this data heterogeneity to provide the efficacy of federated learning in the context of real-life IoT. Further studies ought to look at more sophisticated methods of managing such data diversity, e.g., by clustering or adaptively weighting model updates, to make models even more effective.

Federated Learning Privacy Preservation.
The fact that federated learning does not violate the privacy of the user data is one of the greatest strengths of the technology. In our application, we applied privacy preserving methods like secure aggregation and different privacy in order to make sure that sensitive data do not have to leave the local devices. The findings demonstrated that the model was capable of being trained with the help of individual data set exposure, which reduces the risk of data breach and improves user confidence.
In this case, the privacy of the model aggregation process was upheld by differential privacy in the first place. Differential privacy provides a way to provide noise to the model updates so that even an adversary that tries to reverse-engineer them cannot make inferences about sensitive data points about individual data points (Vyas et al., 2024). This is in agreement with the conclusions of other studies, including those by Vyas et al. (2024) and Ruzafa-Alcazar et al. (2021), who have also stressed the significance of the concepts of differential privacy and secure aggregation in support of the network integrity of federated learning in privacy-sensitive tasks, such as the IoT intrusion-detection.
Nevertheless, as much as federated learning is an effective way of preserving privacy, incorporation of differential privacy may bring in the challenge of model accuracy. Noise that has been introduced to the model updates may, in some cases, decrease the accuracy of the model especially in the event in which minute patterns are required to be detected during the process. A balance has to be found between model accuracy and preservation of privacy. It can be hoped that the noise levels can be optimized in the future, and the new methods, like homomorphic encryption or secure multi-party computation, can be developed to improve the privacy without degrading the performance.

Scalability, Communication Overhead.
Another important factor that should not be overlooked is the communication overhead of the federated learning particularly in the implementation of intrusion detection systems in massive IoT networks. The communication overhead in our study was also 25 percent more than in the traditional centralized systems, which also aligns with other research studies (Khraisat et al., 2025). Such overhead is caused by the necessity of the use of IoT devices to transmit regular model changes to the central server, which is especially troublesome in large scaled IoT networks with numerous devices.
The communication overhead is a problem; however, it is necessary to mention that this trade-off is not usually an issue in privacy-sensitive settings. With the rise in the decentralization of more IoT networks and the increase in data privacy concerns, the advantages of federated learning, especially the maintenance of user privacy, are more valuable than the incremental communication expenses. Besides, the communication load can be reduced with the assistance of different methods, including model compression, periodic updates, and federated learning with asynchronous updates, that do not influence the performance of the model significantly (Mazid et al., 2025). It should be considered to maximize communication efficiency in future studies as this will allow federated learning to be applied on a large scale without clogging network resources.
In our experiment, the federated learning-based IDS was scaled, and the model was effective even in the case of more than 1,000 devices in the IoT network. The outcome is consistent with previous studies that have emphasized the scalability of federated learning, which makes it the optimal choice in a large-scale IoT setting (Ruzafa-Alcazar et al., 2021). Nevertheless, the scalability of federated learning is also conditioned by the effectiveness of model aggregation procedure and network infrastructure that helps the devices to communicate with central server. As the number of IoT networks grows, more improvements will be needed on the federated learning algorithms in order to achieve even greater deployments.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Applicability to the Real World and Future Suggestions.

The findings of this paper highlight the promise of federated learning to be a scalable, privacy-preserving approach to the detection of IoT intrusion. Nevertheless, there are a number of aspects that should be taken into account when adopting this strategy in practical IoT networks. These are the issues of data heterogeneity, communication overhead and real time detection. Federation learning methodology that is capable of dealing with non-IID data, low-cost communication, and low-latency detection in changing IoT scenarios should be developed in the future. Besides, the federated learning to more sophisticated IoT systems, including healthcare and industrial control systems, should be investigated further. In fact, such environments are usually associated with very sensitive information and, therefore, preserving privacy is all the more essential. Such methods as federated learning with edge computing (in which data processing is brought to the source) may further improve privacy and minimize latency (Torre et al., 2025). Furthermore, it could be thought of developing hybrid frameworks involving federated learning and other machine learning frameworks, like reinforcement learning or transfer learning to enhance the flexibility and efficiency of the IDS.

To sum up, federated learning is a prospective solution to enhance the scope of the IoT intrusion detection model and guarantee the preservation of privacy. Through our findings, we have shown that federated learning can provide us with competitive results in the form of accuracy, precision, and recall, although the sensitive data are kept private by using such methods as differential privacy and secure aggregation. Federated learning would be an appropriate solution to IoT networks that are large, privacy sensitive, and susceptible to communication overhead and data heterogeneity due to its scalability and privacy benefits. The research going forward must aim at maximizing the effectiveness of communication, non-IID data, and hybrid frameworks to enhance the performance and usefulness of federated learning to the security of the IoT.

CONCLUSION

The growing use of devices in Internet of Things (IoT) in various industries has brought with it diverse advantages such as automation, real time monitoring, and operational efficiency. Nonetheless, the resulting expansion of the interconnected devices has also presented notable security challenges especially in maintaining privacy of sensitive data as well as in identifying malicious packet within the IoT networks. Conventional intrusion detection systems (IDS), which are usually centralized in terms of collecting data, are not effective at scaling with these settings and they can be easily compromised on privacy. This work has revealed that federated learning (FL) is a promising answer to these issues and offers a framework that has attributes of privacy, scalability, and efficiency to monitor an IoT intrusion.

The main benefit of federated learning applied to the area of IoT security is that the system enables the IoT devices to collectively train a machine learning model without exchanging any sensitive information. As shown in this paper, federated learning is able to achieve the same high level of accuracy, precision, recall, and F1 score, similar to the centralized methods, and still keep the data on local devices and never reveal it to a central server (Ruzafa-Alcazar et al., 2021). This is the capacity to keep the privacy of users and yet have high-detection rates that have been one of the main contributions of this study.

Privacy-preserving methods, including secure aggregation and differential privacy, contribute to the increased complexity of the federated learning model to the point of making sure that model updates are aggregated safely and that there are no attacks on individual data points (Vyas et al., 2024). The methods are useful in alleviating the threat of data leakage and ensuring that federated learning is a highly appropriate method in privacy-sensitive IoT applications. The findings further reveal that federated learning has the capability of efficiently detecting various types of attack on an IoT network such as DoS, DDoS, and Man-in-the-Middle attacks, without interfering with the intrusion detection system performance (Alazab et al., 2023).

Nevertheless, the paper has also created an awareness of some of the issues related to the use of federated learning in IoT security. The communication overhead is one of the most important challenges. Although federated learning provides a high level of privacy, the regular exchange of model updates between devices and the central server may result in a higher communication cost, particularly when using a large-scale IoT node. In spite of these issues, there are a number of optimization approaches that can be utilized to minimize the amount of communication, including model compression and periodic updates (Mazid et al., 2025). Moreover, to solve the problem of data heterogeneity is also a major area of enhancement. The type of data produced by IoT devices is frequently varied, and it may complicate the process of aggregating models and affect the general functioning of the system. The methods of working with non-IID (Non-Independent and Identically Distributed) data like adaptive weighting or

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

grouping of updates should also be studied to enhance the resiliency of federated learning in real IoT networks (Mahmud et al., 2024).

Another vital element that was studied in this paper was the scalability of federated learning in the IoT networks. The federated learning-based IDS was shown to be efficient in terms of scalability and, at the same time, it was possible to preserve high detection performance due to scaling to more than 1,000 devices. This is one of the essential strengths of federated learning because it enables the implementation of intrusion detection systems in extensive and dynamic IoT systems without centralized information processing (Ruzafa-Alcazar et al., 2021). This feature makes federated learning a solution that provides the security of IoT in the future, especially with the increasing number of connected devices.

With respect to the future, various research paths can be identified to help improve the effectiveness and usability of federated learning in the field of internet of things intrusion detection. To start with, the communication overhead and the efficiency of the model updates should be reduced. Federated learning with asynchronous updates or combination with edge computing are some of the techniques that could be considered to continue decreasing latency and network traffic. Second, there is a necessity to focus on the data heterogeneity in IoT networks, which is essential to guarantee that the federated learning models can successfully learn with heterogeneous and unevenly distributed data. Finally, the further extension of the practice of federated learning to more advanced settings involving IoT, including healthcare and industrial internet, will demand further optimization of privacy-saving methods so that sensitive information, including medical history or control system settings could be completely secured.

To sum up, this paper has demonstrated that federated learning is a potentially successful way of enhancing IoT intrusion detection systems and preserving privacy. Federated learning enables decentralized training of machine learning models and incorporates privacy-preserving methods, which can mitigate most of the major issues of the traditional IDS methods, such as privacy threats, scalability, and efficiency. Federated learning is an attractive solution to the future of IoT security despite the issues of overhead communication and data heterogeneity due to its advantages. It will need more development and optimization before the full potential of the technology can be fully achieved in large-scale, real-time IoT environments.

REFERENCES

1) Ruzafa-Alcázar, P., Fernández-Saura, P., Mármol-Campos, E., González-Vidal, A., Hernández-Ramos, J. L., Bernal-Bernabe, J., & Skarmeta, A. F. (2021). Intrusion detection based on privacy-preserving federated learning for the industrial IoT. *IEEE Transactions on Industrial Informatics*, *19*(2), 1145-1154. https://doi.org/10.1109/TII.2021.3126728
2) Vyas, A., Lin, P. C., Hwang, R. H., & Tripathi, M. (2024). Privacy-preserving federated learning for intrusion detection in IoT environments: a survey. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3454211
3) Alazab, A., Khraisat, A., Singh, S., & Jan, T. (2023). Enhancing privacy-preserving intrusion detection through federated learning. *Electronics*, *12*(16), 3382. https://doi.org/10.3390/electronics12163382
4) Torre, D., Chennamaneni, A., Jo, J., Vyas, G., & Sabrsula, B. (2025). Toward enhancing privacy preservation of a federated learning cnn intrusion detection system in iot: method and empirical study. *ACM Transactions on Software Engineering and Methodology*, *34*(2), 1-48. https://doi.org/10.1145/3695998
5) Mazid, A., Kirmani, S., Manaullah, & Yadav, M. (2025). FL-IDPP: A Federated Learning Based Intrusion Detection Approach With Privacy Preservation. *Transactions on Emerging Telecommunications Technologies*, *36*(1), e70039. https://doi.org/10.1002/ett.70039
6) Chandu, G., Karthik, T., & Parag, B. (2025). Federated Learning for Distributed IoT Security: A Privacy-Preserving Approach to Intrusion Detection. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3592481
7) Mahmud, S. A., Islam, N., Islam, Z., Rahman, Z., & Mehedi, S. T. (2024). Privacy-preserving federated learning-based intrusion detection technique for cyber-physical systems. *Mathematics*, *12*(20), 3194. https://doi.org/10.3390/math12203194
8) Duy, P. T., Hao, H. N., Chu, H. M., & Pham, V. H. (2021, October). A secure and privacy preserving federated learning approach for IoT intrusion detection system. In *International Conference on Network and System Security* (pp. 353-368). Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-030-92708-0_23

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

9) Khraisat, A., Alazab, A., Alazab, M., Obeidat, A., Singh, S., & Jan, T. (2025). Federated learning for intrusion detection in IoT environments: a privacy-preserving strategy. *Discover Internet of Things*, *5*(1), 72. https://link.springer.com/article/10.1007/s43926-025-00169-7
10) Raza, M., Saeed, M. J., Riaz, M. B., & Sattar, M. A. (2024). Federated learning for privacy-preserving intrusion detection in software-defined networks. *IEEE Access*, *12*, 69551-69567. https://doi.org/10.1109/ACCESS.2024.3395997
11) Jeyakumar, S. R., Rahman, M. Z. U., Sinha, D. K., Kumar, P. R., Vimal, V., Singh, K. U., ... & Balajee, J. (2024). An Innovative Secure and Privacy-Preserving Federated Learning-Based Hybrid Deep Learning Model for Intrusion Detection in Internet-Enabled Wireless Sensor Networks. *IEEE Transactions on Consumer Electronics*, *71*(1), 273-280. https://doi.org/10.1109/TCE.2024.3442015
12) Tabassum, A., Erbad, A., Lebda, W., Mohamed, A., & Guizani, M. (2022). Fedgan-ids: Privacy-preserving ids using gan and federated learning. *Computer Communications*, *192*, 299-310. https://doi.org/10.1016/j.comcom.2022.06.015
13) Amro, S. A. (2025). Securing Internet of Things Devices with Federated Learning: A Privacy-Preserving Approach for Distributed Intrusion Detection. *Computers, Materials & Continua*, *83*(3). DOI: 10.32604/cmc.2025.063734
14) Rampone, G., Ivaniv, T., & Rampone, S. (2025). A Hybrid Federated Learning Framework for Privacy-Preserving Near-Real-Time Intrusion Detection in IoT Environments. *Electronics*, *14*(7), 1430. https://doi.org/10.3390/electronics14071430
15) Soomro, I. A., Rehman, H. U., Hussain, S. J., Latif, S., Mujlid, H., Mohsin, S. M., & Maple, C. (2025). ROCHE: A Robust and End-to-End Privacy-Preserving Federated Learning Framework for Intrusion Detection in Industrial Internet of Things. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2025.3612944