

Federated Learning-Based Intrusion Detection Systems for Transportation IoT: A Comparative Study on Efficiency, Privacy, and Scalability

Farzana Anjum G¹, Divya Rani², Aparna Singh³, Dr. Nirmala S⁴

^{1,2,3}Student, Dept. of Computer Science Engineering, AMC Engineering College, Karnataka, India

⁴Professor, Dept. of Computer Science Engineering, AMC Engineering College, Karnataka, India

Abstract - With the widespread integration of Internet of Things (IoT) devices in transportation systems, maintaining data confidentiality and system reliability has become increasingly critical. Traditional intrusion detection systems (IDS), typically built on conventional machine learning models, face challenges including data privacy risks, high communication overhead, and reduced responsiveness in real-time scenarios. Federated Learning (FL) offers a decentralized approach that enables local model training across distributed devices, preserving data privacy while leveraging collective intelligence. This paper provides a comparative analysis of four FL-based IDS frameworks developed for transportation IoT environments. Each study introduces unique methodologies—ranging from lightweight fine-tuning techniques to feature optimization strategies and real-time deployments on edge hardware. The analysis focuses on evaluating detection accuracy, system efficiency, scalability, and deployment feasibility. The comparative findings highlight critical design considerations and outline potential pathways for developing effective, privacy-preserving IDS solutions tailored for the transportation sector.

Key Words: Federated Learning, Intrusion Detection System (IDS), Transportation IoT, Edge Computing, Cybersecurity, Privacy Preservation, Distributed Learning

1. INTRODUCTION

In recent years, smart transportation systems have rapidly adopted Internet of Things (IoT) technologies, enabling advanced functionalities such as automated driving, real-time traffic monitoring, and vehicle-to-everything (V2X) communication. While these developments enhance operational efficiency and safety, they also expose systems to a wide array of cyber threats. Traditional IDS frameworks often rely on centralized data collection and model training, which can lead to significant drawbacks, including compromised privacy, communication bottlenecks, and limited scalability in dynamic vehicular environments.

In order to overcome these obstacles, researchers have explored Federated Learning (FL), a distributed machine learning framework that allows distributed devices to

collaboratively develop models without exchanging raw data. This approach maintains user privacy while enhancing adaptability and responsiveness across heterogeneous networks. In the context of transportation IoT, FL-based IDS architectures have shown potential in detecting complex and evolving attack patterns while operating efficiently on edge devices.

This study presents a comparative review of four contemporary FL-IDS models that aim to strengthen cybersecurity in transportation networks. These comprise systems engineered for resource-constrained vehicular nodes, models optimized through hybrid server-edge learning, and frameworks integrating deep learning integrated with feature selection algorithms. By analysing each approach in terms of accuracy, deployment strategy, and system requirements, this paper offers significant understanding of the current landscape of FL-IDS and identifies potential directions for future research.

2. LITERATURE SURVEY

The implementation of Federated Learning (FL) with Intrusion Detection Systems (IDS) has emerged as a significant advancement in the cybersecurity landscape of Transportation IoT (T-IoT). Several recent studies propose FL-based frameworks to mitigate evolving cyber threats while safeguarding data privacy and reducing centralized dependencies.

In the study by Bhavsar et al. [4], an FL-IDS system optimized for vehicular environments is introduced. This model employs a blend of logistic regression and convolutional neural networks (CNNs) and is deployed on real-time edge platforms such as Raspberry Pi and Jetson Xavier. By utilizing embedded devices for local model training and a central aggregator for global model refinement, this architecture successfully maintains data privacy while maintaining high detection accuracy. Experiments conducted using the NSL-KDD and Car-Hacking datasets demonstrated a performance gain over centralized IDS models, achieving accuracies of up to 99%.

Lazzarini et al. [2] propose a lightweight FL-based IDS framework for IoT systems that utilizes a shallow artificial

neural network (ANN) and federated averaging (FedAvg) to perform both binary and multiclass classification on datasets like ToN_IoT and CICIDS2017. Their study compares centralized learning to FL approaches, indicating that FL can maintain comparable accuracy while significantly reducing privacy risks. The authors also evaluate advanced aggregation methods (FedAvgM, FedAdam, and FedAdagrad), finding that FedAvg and FedAvgM generally perform better in heterogeneous environments.

In their study, Akinie et al. [3] present a hybrid FL-IDS framework optimized for resource-constrained edge environments such as Connected and Autonomous Vehicles (CAVs). The proposed system leverages server-side pre-training and on-device fine-tuning to reduce computational and memory overhead. Using a transfer learning-inspired approach, only the classification head is updated on the edge, preserving energy while maintaining model adaptability. The system achieved memory savings of up to 42% and training time reduction of up to 75%, all while sustaining detection accuracy levels up to 99.2%.

Another remarkable contribution is by Karunamurthy et al. [1], who integrate a Chimp Optimization Algorithm with FL to improve feature selection and detection performance in IoT-based networks. Their approach allows distributed training using deep learning classifiers while simultaneously reducing dimensionality to enhance classification accuracy. Using the MQTT dataset, the proposed solution accomplishes a detection accuracy of 95.59%, outperforming traditional machine learning-based IDS methods. Their work also emphasizes scalability and robustness against complex cyberattacks, including DDoS and brute force attacks.

Collectively, these studies demonstrate the potential of FL in safeguarding T-IoT networks while offering varying strategies for overcoming resource constraints, data heterogeneity, and communication bottlenecks. The utilization of edge devices for local inference, novel aggregation and optimization techniques, and hybrid training strategies are recurring themes that indicate the direction of future advancements in this domain.

3. PROBLEM STATEMENT

The increasing reliance on Internet of Things (IoT) devices in transportation systems—such as connected and autonomous vehicles (CAVs), smart traffic control units, and roadside sensors—has introduced a new layer of complexity and vulnerability to modern cyber-physical infrastructures. These systems are frequently targeted by sophisticated cyber threats, including denial-of-service attacks, data tampering, spoofing, and unauthorized access. Traditional Intrusion Detection Systems (IDS), which rely on centralized data collection and training, are ill-suited for such environments due to their inability to scale, maintain data privacy, and deliver real-time responses. Centralized

approaches also give rise to critical privacy challenges and suffer from communication overhead, latency issues, and model drift due to non-uniform data distribution across devices.

Although Federated Learning (FL) has developed into a viable alternative by permitting collaborative learning without raw data sharing, existing FL-IDS frameworks still face challenges related to computational constraints on edge devices, feature selection complexity, model accuracy degradation, and inefficient aggregation strategies. There is a lack of standardized evaluation across diverse architectures, which makes it difficult to determine the most efficient and scalable FL-IDS approach for Transportation IoT environments.

4. PROPOSED SOLUTION

This research undertakes a comprehensive comparative study of four recent Federated Learning-based IDS frameworks that have been developed for securing Transportation IoT systems. The primary goal is to identify the architectural elements, learning strategies, and optimization techniques that enable accurate, efficient, and privacy-preserving threat detection in distributed vehicular and infrastructure-based environments. By analysing models that employ diverse methodologies—such as hybrid server-edge learning, shallow neural networks, convolutional deep learning classifiers, and metaheuristic-based feature selection—this study aims to uncover patterns and trade-offs that influence model performance under real-world constraints.

The selected frameworks have been evaluated using benchmark datasets that reflect the complexities of traffic-related cyber threats, including NSL-KDD, CICIDS2017, ToN-IoT, and MQTT datasets. Each system offers a distinct approach to handling non-IID data, limited computational resources, and the demand for minimal communication overhead between clients and the central aggregator. Through this comparative analysis, the research seeks to highlight the benefits and drawbacks of current FL-IDS solutions and propose a direction for future development that emphasizes robustness, adaptability, and real-time feasibility in Transportation IoT applications. By synthesizing these insights, the study facilitates the growth of secure, decentralized IDS models capable of defending against evolving cyber threats while ensuring data confidentiality and operational efficiency across vehicular networks.

5. ARCHITECTURE DIAGRAM

The architecture of a Federated Learning-based Intrusion Detection System for Transportation IoT is designed to ensure security while preserving privacy and maintaining system efficiency. At the heart of this architecture lies a

central aggregator, typically located on a server or cloud platform, which is responsible for coordinating the learning process. This component does not gather raw data from devices; rather, it aggregates model parameters sent by distributed clients to generate an updated global model.

Each edge device, such as a connected vehicle, roadside unit (RSU), or embedded IoT sensor, collects local network traffic data and creates a local replica of the IDS model using this data. These models may be built using shallow ANNs, CNNs, or hybrid ML-DL combinations, based on the specific framework. Importantly, only the updated model weight updates or gradients are shared back to the central aggregator—ensuring that sensitive raw data remains on the edge environment.

In some implementations, such as the one proposed by Akinie et al., a hybrid server-edge architecture is employed. Here, the heavier pre-training is conducted on the server using proxy datasets, and only lightweight fine-tuning is performed on edge devices, which helps reduce memory and computation loads. Bhavsar et al.'s FL-IDS uses real-time data collection from CAVs and employs CNNs for deep learning-based classification. Meanwhile, Karunamurthy et al. enhance this architecture with a Chimp Optimization Algorithm to select the most relevant features before training, thereby reducing the dimensionality and improving classification efficiency.

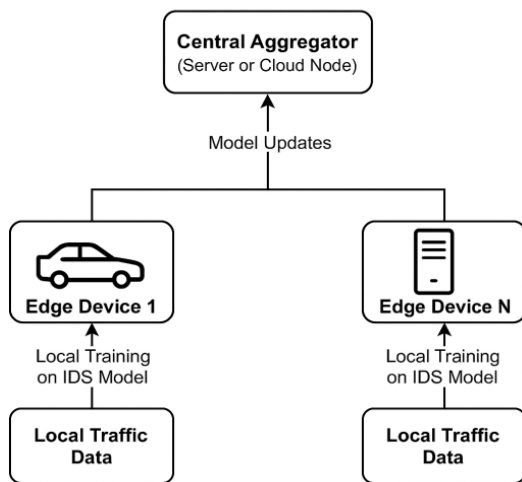


Fig -1: Architecture of a Federated Learning-Based Intrusion Detection System (FL-IDS) for Transportation IoT environments

6. METHODOLOGY

The methodology adopted in this research involves a comparative evaluation of four Federated Learning-based Intrusion Detection System (FL-IDS) frameworks designed for Transportation IoT environments. Each selected study introduces a distinct architectural model and learning strategy, focusing on enhancing intrusion detection

performance while addressing challenges such as data privacy, resource constraints, and real-time adaptability. The primary objective of this methodological approach is to assess the effectiveness, scalability, and practicality of each framework in terms of detection accuracy, communication efficiency, and deployment feasibility.

In contrast, the model proposed by Lazzarini et al. utilizes a shallow artificial neural network and compares various aggregation algorithms such as FedAvg, FedAdam, and FedAdagrad to evaluate their performance in binary and multiclass classification tasks. The evaluation is carried out on standard benchmark datasets like ToN-IoT and CICIDS2017. Karunamurthy et al. propose an enhancement to traditional FL-IDS models by integrating a Chimp Optimization Algorithm to perform feature selection, thereby reducing the complexity of the input space and improving the detection performance of deep learning classifiers.

Each framework was examined through a common set of metrics—accuracy, false positive rate, computational load, communication efficiency, and adaptability to distributed network environments. The datasets used by the authors were also noted, as they vary in size, data type, and structure, offering perspective on how each model functions under various traffic conditions. For example, models tested on high-dimensional datasets like CICIDS2017 and ToN-IoT reveal the scalability of the framework, while lightweight datasets such as NSL-KDD help evaluate real-time performance in constrained environments.

Through this comparative approach, the study aims to identify key architectural and algorithmic choices that contribute to better intrusion detection in Transportation IoT. Rather than proposing a single unified model, this methodology emphasizes analysing the benefits and drawbacks of diverse implementations to derive generalized recommendations for future FL-IDS design in vehicular and smart infrastructure networks.

7. RESULTS AND DISCUSSION

This research analyses and compares four Federated Learning-based Intrusion Detection Systems (FL-IDS) frameworks developed for Transportation IoT environments. Each model emphasizes a different aspect of improving intrusion detection, such as privacy preservation, resource optimization, aggregation strategies, or feature selection. To evaluate these frameworks systematically, parameters such as detection accuracy, processing efficiency, scalability, and privacy measures were considered.

The FL-IDS proposed by Bhavsar et al. demonstrates excellent real-world deployment capabilities by integrating both logistic regression and CNN classifiers into an edge-server FL framework. Their experiments on NSL-KDD and

Car-Hacking datasets achieved accuracies of up to 99%, validating the effectiveness of deep learning models combined with federated learning for vehicular networks. However, real-time performance was found to be highly dependent on the edge device's computational capacity.

In contrast, the model by Akinie et al. addresses resource constraints more aggressively by proposing a hybrid server-edge architecture. Pre-training is carried out on a central server, while lightweight fine-tuning occurs on edge devices. This approach achieves up to 99.2% detection accuracy while reducing memory usage by 42% and training time by 75%, making it highly practical for Connected and Autonomous Vehicles (CAVs) with limited resources.

Meanwhile, the framework developed by Lazzarini et al. focuses on evaluating different aggregation algorithms in a federated setting. Their use of FedAvg and FedAvgM showed stable and reliable performance, while FedAdam and FedAdagrad slightly underperformed in non-IID environments. Their experiments demonstrated that FL models could achieve comparable performance to centralized models on the ToN-IoT and CICIDS2017 datasets, proving the feasibility of FL for distributed IoT security without raw data exchange.

Table -1: Comparative analysis of different Federated Learning-based Intrusion Detection Systems for Transportation IoT in terms of architecture, focus area, datasets, accuracy, optimization techniques, resource usage, scalability, and privacy preservation.

Aspect	Bhavsar et al. (FL-IDS)	Akinie et al. (FedFT)	Lazzarini et al. (FedAvg Study)	Karunamurthy et al. (Chimp FL-IDS)
FL Architecture	Edge-server with CNN & LR	Hybrid with fine-tuning	Multi-aggregation FL	FL with Chimp-based feature selection
Focus	Real-time edge deployment	Resource-efficient & scalable	Aggregation method analysis	Dimensionality reduction
Datasets	NSL-KDD, Car-Hacking	Custom CAV dataset	ToN98*--IoT, CICIDS2017	MQTT
Accuracy	Up to 99%	Up to 99.2%	~95-97%	95.59%
Optimization	CNN-based feature extraction	Pre-training + client fine-tuning	FedAvg, FedAdam, etc.	Chimp Optimization Algorithm

	n	tuning		
Resources	Moderate	42% less memory, 75% faster	Low to moderate	Reduced computation
Scalability	Good with few clients	High scalability	Handles heterogeneity well	Moderate, focus on feature efficiency
Privacy	Model updates only	Strong privacy, low device load	Full FL privacy	Full privacy + local feature selection

8. CONCLUSION

The increasing deployment of IoT devices in transportation systems has introduced new vulnerabilities that demand efficient, scalable, and privacy-preserving security mechanisms. Federated Learning (FL) offers a promising solution by enabling collaborative model training across distributed nodes without requiring the exchange of raw data. This research provided a comparative study of four recent FL-based Intrusion Detection Systems (FL-IDS) designed specifically for Transportation IoT environments. The analysis revealed that different architectures prioritize different aspects of system design, ranging from real-time deployment feasibility, as demonstrated by Bhavsar et al., to resource optimization through hybrid server-edge learning models proposed by Akinie et al. Similarly, the exploration of multiple aggregation algorithms by Lazzarini et al. emphasized the importance of aggregation choice in maintaining model stability under heterogeneous data conditions, while Karunamurthy et al.'s feature selection approach underscored the role of dimensionality reduction in improving detection efficiency.

Overall, the findings indicate that no single FL-IDS framework provides a universal solution; rather, the optimal design depends heavily on the target deployment environment, available computational resources, data characteristics, and security requirements. Future endeavours may concentrate on developing adaptive FL-IDS models that combine lightweight model architectures, intelligent feature selection, dynamic aggregation strategies, and robust mechanisms for handling non-IID and imbalanced data. By integrating these advancements, it is possible to build intrusion detection systems that not only safeguard smart transportation networks but also align with the evolving needs of privacy, scalability, and real-time threat detection.

9. FUTURE WORK

While the current study has highlighted the strengths and limitations of existing Federated Learning-based Intrusion Detection Systems for Transportation IoT, several avenues remain open for further exploration and enhancement. One important direction is the development of adaptive FL frameworks capable of dynamically selecting aggregation strategies based on the nature of participating clients' data distribution. This would allow models to maintain high accuracy even in highly non-IID and heterogeneous environments, which are typical in large-scale transportation networks.

Another area for future research lies in optimizing communication overhead in federated settings. Although FL reduces raw data transmission, frequent model updates can still strain bandwidth, especially in vehicular networks where connectivity is intermittent. Techniques such as model compression, sparsification, and asynchronous updates could be further explored to minimize network congestion without compromising model performance.

Moreover, integrating blockchain or secure multiparty computation (SMPC) protocols into FL-IDS frameworks can enhance trustworthiness and transparency among collaborating devices. Such integration would not only strengthen the privacy guarantees but also ensure that malicious participants cannot compromise the training process.

In addition, expanding the scope of FL-IDS models to detect emerging attack types, such as adversarial attacks against deep learning models and coordinated multi-vehicle threats, will be critical to maintaining system resilience. Future studies could also explore multi-modal learning approaches that incorporate various sensor modalities—such as LiDAR, radar, and camera feeds—along with network traffic data to improve the overall detection capabilities of autonomous transportation systems.

Lastly, real-world deployments and longitudinal studies in operational CAV ecosystems are necessary to validate FL-IDS models under practical constraints like mobility, varying device resources, and real-time latency requirements. Such studies would help bridge the gap between theoretical advancements and practical implementations in smart transportation infrastructures.

ACKNOWLEDGEMENT

First and foremost, we extend our sincere thanks to our guide, Dr. Nirmala S for her invaluable guidance, constant encouragement, and insightful feedback throughout the study. Her expertise and suggestions have been instrumental in shaping this project.

We are also deeply thankful to our Head of Department, Dr. V. Mareeswari for providing a conducive environment for learning and research.

We are deeply grateful to our institution, AMC Engineering College, for providing the necessary resources and infrastructure to carry out this project successfully. We also acknowledge the faculty members of our Department of Computer Science for their support and motivation.

REFERENCES

- [1] 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. 8696, 2025. [Online]. Available: <https://doi.org/10.1038/s41598-025-93501-8>
- [2] R. Lazzarini, H. Tianfield, and V. Charissis, "Federated Learning for IoT Intrusion Detection," *AI*, vol. 4, no. 3, pp. 509–530, Jul. 2023. [Online]. Available: <https://doi.org/10.3390/ai4030028>
- [3] R. Akinie, N. K. Gyimah, M. Bhavsar, and J. Kelly, "Fine-Tuning Federated Learning-Based Intrusion Detection Systems for Transportation IoT," *arXiv preprint, arXiv:2502.06099*, 2025. [Online]. Available: <https://arxiv.org/abs/2502.06099>
- [4] M. H. Bhavsar, Y. B. Bekele, K. Roy, J. C. Kelly, and D. Limbrick, "FL-IDS: Federated Learning-Based Intrusion Detection System Using Edge Devices for Transportation IoT," *IEEE Access*, vol. 12, pp. 52215–52230, Apr. 2024. [Online]. Available: <https://doi.org/10.1109/ACCESS.2024.3386631>