A Federated LSTM Autoencoder Framework for Privacy-Preserving Intrusion Detection in V2X Networks

B. Vishwanath
Department of ECE JNTUH, University College of Engineering, Science & Technology Hyderabad, Telangana, India
vishwanathb782@gmail.com (corresponding author)

P. Chandrasekhar Reddy
Department of ECE JNTUH, University College of Engineering, Science & Technology Hyderabad, Telangana, India
drpcreddy@jntuh.ac.in

Received: 2 July 2025 | Revised: 8 August 2025, 3 September 2025, and 7 September 2025 | Accepted: 9 September 2025

 DOI: https://doi.org/10.48084/etasr.13121

ABSTRACT

The fast growth of Vehicle-to-Everything (V2X) networks requires privacy-preserving Intrusion Detection Systems (IDSs) for effective operation. The proposed Federated Long Short-Term Memory Autoencoder (Fed-LSTM-AE) framework allows distributed vehicular clients to perform collaborative anomaly detection through model parameter sharing without exchanging raw data. The framework enables each client to create its own LSTM-based autoencoder model of normal traffic patterns while sharing only model parameters with a central server through federated learning to maintain data privacy and improve system scalability. Experiments using the VeReMi dataset show that Fed-LSTM-AE achieves better performance than the centralized LSTM, one-dimensional Convolutional Neural Network (1D CNN), Random Forest, and Isolation Forest baseline methods in terms of detection accuracy, F1-score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC) metrics. The framework shows strong detection performance against various attack types while achieving efficient federated training convergence and maintaining stability under non-Independent and Identically Distributed (non-IID) data conditions. The results demonstrate Fed-LSTM-AE's suitability for real-world V2X deployments because it maintains privacy protection while being adaptable and communication-efficient.

Keywords-V2X security; federated learning; LSTM autoencoder; vehicular networks; privacy preservation; anomaly detection

I. INTRODUCTION

The fast development of Intelligent Transportation Systems (ITS) together with Vehicle-to-Everything (V2X) communication technologies has revolutionized connected and autonomous mobility systems. The transportation ecosystem benefits from V2X through its four communication modalities, which include Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), Vehicle-to-Pedestrian (V2P), and Vehicle-to-Network (V2N) communication, for smooth information sharing between different entities [1-6]. The connected nature of these systems enables substantial improvements in road safety, traffic efficiency, and real-time decision-making [7]. However, the open and distributed nature of vehicular networks creates multiple security vulnerabilities, which endanger the reliability and integrity of these systems. The V2X network faces numerous cyberattacks, including GPS spoofing, Sybil attacks, message injection and replay attacks, and Denial-of-Service (DoS) exploits. These threats compromise both data integrity and network availability, posing direct risks to physical safety that can lead to vehicle accidents and worsen traffic congestion [8, 9].

The deployment of effective and scalable intrusion detection mechanisms becomes more difficult because vehicular environments present high mobility, strict latency requirements, and intermittent connectivity. Signature-based Intrusion Detection Systems (IDSs) and centralized anomaly detection frameworks fail to address the distinctive challenges of V2X networks because of their traditional, static security solutions. Deep learning techniques have shown promising results in traffic pattern modeling and anomaly detection using Long Short-Term Memory (LSTM) networks [10, 11] and Convolutional Neural Networks (CNNs) [12, 13]. The current practice of centralized model training for these systems requires large-scale data storage and collection, which worsens

privacy and latency issues in vehicular environments. These approaches fail to address the fundamental decentralized characteristics and heterogeneity of vehicular data. Moreover, they often fail to recognize the natural diversity and distributed characteristics of vehicular data.

To address these challenges, this study proposes a Federated LSTM Autoencoder (Fed-LSTM-AE) framework for anomaly-based intrusion detection in V2X networks. The proposed system uses federated learning and deep sequential modeling to enable collaborative and privacy-preserving training of anomaly detectors across distributed vehicular nodes. Each vehicle independently trains an LSTM autoencoder on its local benign traffic data, transmitting only the learned model parameters, never raw data, to a central server. The server then aggregates these parameters via a federated averaging strategy, yielding a global model that generalizes across heterogeneous clients while safeguarding privacy and minimizing bandwidth consumption.

The proposed method achieves high efficacy and robustness through experiments with the VeReMi dataset [14], which serves as a standard benchmark for V2X misbehavior detection. The proposed method is evaluated against four state-of-the-art baselines, including a centralized LSTM classifier, a one-dimensional CNN (1D CNN), Random Forest, and Isolation Forest models. The performance evaluation uses standard metrics which include accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). The paper also includes a per-attack analysis to show how the proposed detection model performs on different attack types, including GPS spoofing, position injection, Sybil attacks, and replay attacks.

The main contributions of this study are as follows:

- The proposed Fed-LSTM-AE framework represents a novel solution for anomaly-based intrusion detection in V2X networks, which fulfills both privacy preservation and scalability requirements.
- The development of a comprehensive federated learning architecture implemented using the Flower framework and TensorFlow, capable of simulating decentralized training across vehicular clients.
- Extensive benchmarking against four competitive IDS baselines on the VeReMi dataset, demonstrating superior performance in both aggregate and per-attack detection metrics.
- An analysis of each attack demonstrating the model's ability to handle different types of adversarial scenarios.

II. METHODOLOGY

The proposed Fed-LSTM-AE framework for anomaly detection in V2X networks includes a detailed design and operational workflow that addresses the specific requirements of vehicular environments. The framework is specifically designed to handle decentralized data generation and real-time detection while maintaining privacy standards and supporting the geographical scalability of vehicular nodes. The methodology achieves its objectives by integrating four main components into a unified system, including a local LSTM autoencoder for learning benign traffic patterns, a federated learning protocol for privacy-preserving collaborative model training, a reconstruction-based anomaly detection mechanism, and a realistic federated simulation environment built using the VeReMi dataset. The following discussion presents each component with structured pseudocode to ensure methodological transparency and reproducibility.

A. Local Model

Each vehicular client trains a local LSTM autoencoder to learn the temporal structure of benign V2X communication data. The autoencoder consists of two parts: an encoder that transforms time-series inputs [14] into a constant-size latent vector and a decoder that tries to rebuild the original sequence from this vector. Let the input to client i and be a sequence of T time-ordered feature vectors:

$$X^{(i)} = \left\{x_1^{(i)}, x_2^{(i)}, \dots, x_T^{(i)}\right\}, \;\; x_t^{(i)} \in \mathbb{R}^d \tag{1}$$

where each $x_i^{(t)}$ is a d-dimensional vector containing features such as position, speed, heading, and message timestamp at time t. The encoder processes the input sequence using stacked LSTM layers to capture temporal dependencies and compress it into a hidden representation h [15]. The decoder is initialized using the final hidden and cell states of the encoder and reconstructs the input sequence:

$$\hat{X}^{(i)} = \{\hat{x}_1^{(i)}, \hat{x}_2^{(i)}, \dots, \hat{x}_T^{(i)}\} \tag{2}$$

The model is trained to minimize the reconstruction loss using the Mean Squared Error (MSE) between the original and reconstructed sequences:

$$\mathcal{L}_{rec}^{(i)} = \frac{1}{T} \sum_{t=1}^{T} \left\| x_t^{(i)} - \hat{x}_t^{(i)} \right\|_2^2 \tag{3}$$

This unsupervised training strategy [16] allows the model to learn a compact representation of benign behavior without requiring labeled attack data. Since the model is trained solely on normal traffic, it is expected to yield low reconstruction error on benign inputs and higher error on anomalous or adversarial sequences. The training process is executed locally on each client, with no raw data shared externally. Once local training is complete, each client contributes to the collaborative model optimization process via federated learning, as described in the next subsection.

B. Federated Learning Protocol

The implementation of the synchronous Federated Averaging (FedAvg) protocol enables collaborative learning between multiple vehicular clients while maintaining data privacy [17, 18]. The system design allows clients to maintain their local V2X traffic data while sending trained model parameters to the central server. This approach minimizes bandwidth usage and protects sensitive vehicular data from exposure because it eliminates the need to share raw data. The server-side coordination process is described in Algorithm 1. The server initiates each communication round by choosing random clients for participation. These selected clients receive the current global model parameters and perform local training using their private datasets. After local training, clients send

their updated model parameters back to the server, which aggregates them using a weighted average based on the number of samples in each local dataset. This results in an updated global model that reflects the combined knowledge of the participating clients.

Algorithm 1: Server-side federated coordination
Input: θ_0 (initial global model parameters), K (total number of clients), T (number of training rounds, C (fraction of clients selected per round) $(0 < C \leq 1)$
Output: θ_T (final trained global model)
1: Initialize $\theta = \theta_0$
2: For each round t = 1 to T do:
3: Randomly select a subset S of $\lceil C \times K \rceil$ clients
4: For each client i in S (in parallel):
5: Send current global model θ to client i
6: Receive locally updated model θ_i from client i
7: End for
8: Aggregate all client updates: $\theta \leftarrow \sum(|\mathcal{D}_i|/|\mathcal{D}|) \cdot \theta_i$ where $|\mathcal{D}_i|$ is the size of client i's data, and $|\mathcal{D}|$ is the total data across all selected clients
9 : End for
10: Return θ as θ_T

This aggregation process ensures that clients with more data influence the global model proportionally more, thereby improving generalization across heterogeneous data distributions. The local training executed by each client is described next in Algorithm 2.

Algorithm 2: Client-side local training
Input: Client ID i, global model parameters θ
Output: Updated local model parameters θ_i
1: Initialize $\theta_i = \theta$
2: Load local dataset $\mathcal{D}_i$
3: For each epoch from 1 to E do:
4: Divide $\mathcal{D}_i$ into mini batches of size B
5: For each mini batch b in $\mathcal{D}_i$:
6: Compute reconstruction loss $\mathcal{L} = \text{AutoencoderLoss}(b, \theta_i)$
7: Update θ_i using gradient descent
8: End for
9: End for
10: Return θ_i

The key operation during local training is the calculation of the reconstruction loss for each mini batch, which quantifies how well the model replicates input sequences. This is detailed in Algorithm 3. The autoencoder loss is computed as the MSE between the input and reconstructed sequences. It serves as both the training loss during local optimization and the anomaly score during inference.

Algorithm 3: Autoencoder loss computation
Input: Mini batch b of V2X sequences, model parameters θ
Output: Batch loss $\mathcal{L}_{rec}$
1: For each sequence X in batch b:
2: Encode $X \rightarrow$ latent vector h using the LSTM encoder
3: Decode $h \rightarrow$ reconstructed sequence $\hat{X}$ using LSTM decoder
4: Compute squared error: loss = $\|X - \hat{X}\|^2$
5: End for
6: Compute the average loss across the batch: $\mathcal{L}_{rec} = (1/b) \times$ sum of all individual losses
7: Return $\mathcal{L}_{rec}$

The loss $\mathcal{L}_{rec}$ is minimized during local training and later used in the global inference phase to detect deviations from benign traffic patterns. After multiple rounds of federated updates, the global model captures a generalized representation of normal V2X behavior. This enables robust detection of malicious sequences, as described in the next subsection.

C. Anomaly Detection

Once the global LSTM autoencoder has been trained through multiple rounds of federated learning, it is deployed to participating vehicles or centralized infrastructure for real-time inference [19-21]. The primary role of the autoencoder during inference is to evaluate whether a given V2X communication sequence conforms to the learned benign patterns. This is accomplished by computing the reconstruction loss for each input sequence. Let $X' = \{x'_1, x'_2, \ldots, x'_T\}$ be a new sequence of observed V2X features, where $x'_t \in \mathbb{R}^d$ is the feature vector at time t. The trained autoencoder attempts to reconstruct this sequence, resulting in $\hat{X}' = \{\hat{x}'_1, \hat{x}'_2, \ldots, \hat{x}'_T,\}$. The anomaly score is then calculated using the MSE over the sequence:

$$\mathcal{L}_{rec}(X') = \frac{1}{T} \sum_{t=1}^{T} \|x'_t - \hat{x}'_t\|_2^2 \quad (4)$$

A threshold τ is established to differentiate between normal and anomalous behavior. This threshold is typically selected empirically by analyzing the distribution of reconstruction losses on a held-out validation set containing only benign samples. In this work, we set τ to the 95th percentile of the validation loss distribution, which effectively captures the upper bound of normal variations while minimizing false positives. The final anomaly decision rule is defined as:

$$\text{Anomaly}(X') = \begin{cases} 1, & \text{if } \mathcal{L}_{rec}(X') > \tau \text{ (anomalous)} \\ 0, & \text{otherwise (benign)} \end{cases} \quad (5)$$

This threshold-based approach is inherently unsupervised and does not rely on labeled attack data, which are often

unavailable or insufficient in real-world V2X scenarios. Moreover, since the model is trained only on normal data, it is well-suited to detect zero-day attacks and previously unseen anomalies that deviate from learned benign behavior. This detection mechanism is later evaluated in our experiments across multiple attack categories, including GPS spoofing, position falsification, Sybil identity fraud, and replay attacks. Before presenting those results, the next section describes how the VeReMi dataset is partitioned to simulate a realistic federated environment across heterogeneous vehicular clients.

D. System Integration

The Fed-LSTM-AE framework requires an operational framework that combines end-to-end training with aggregation and inference operations in a federated V2X network. The modular and scalable architecture provides both privacy compliance and practical deployment of the methodology described earlier. The system consists of two main components, including vehicular clients that represent individual vehicles or roadside units, and a central server that manages global training and aggregation. The VeReMi dataset is divided into non-Independent and Identically Distributed (non-IID) subsets, which each client receives separately to represent its unique localized communication behaviors under different spatial and temporal conditions. The established setup duplicates actual heterogeneous V2X environments.

The clients perform local LSTM autoencoder training using only benign data because vehicles do not possess attack data labels in real-world scenarios. The clients send model parameter updates to the server following each training round instead of sharing raw data, which protects privacy and minimizes communication costs. The central server executes the federated learning protocol by sending the global model to random client subsets during each round and then gathering their weight updates to perform FedAvg based on dataset sizes. The process continues until the model reaches convergence.

The trained global model serves as a deployment tool for inference purposes, either at the edge through client deployment for decentralized real-time anomaly detection or at the server level. The LSTM autoencoder uses reconstruction loss to evaluate new V2X sequences during inference and identifies anomalies when sequences exceed a predefined threshold, thus enabling detection of zero-day attacks without needing labeled adversarial data. The implementation utilizes the Flower framework for federated learning, together with TensorFlow for model training. The simulated environment allows clients to function autonomously while supporting parallel operations. The VeReMi dataset undergoes preprocessing to create normalized fixed-length sequences before being distributed across up to ten client nodes. The testing process uses attack data exclusively, whereas training operates only on benign data.

The integrated system provides strong, privacy-preserving anomaly detection capabilities for V2X networks when operating under actual federated system limitations. The following section explains the experimental setup, which includes model configurations, evaluation metrics, baseline comparisons, and performance assessment procedures.

III. EXPERIMENTAL SETUP

This section presents the experimental protocol designed to evaluate the proposed Fed-LSTM-AE framework under conditions representative of real-world federated learning in V2X networks. The experimental configuration emulates decentralized and privacy-preserving environments through partitioned, non-IID client datasets and distributed training processes. The evaluation methodology comprises five core components: dataset preprocessing, simulation of the federated learning environment, model configuration and training strategy, comparative baseline methods, and the set of performance metrics used for both global and per-attack analyses.

To validate the Fed-LSTM-AE framework under realistic federated V2X conditions, we use the VeReMi dataset, which provides time-stamped Cooperative Awareness Messages (CAMs) from simulated urban traffic, including both benign and multiple attack types. Data preprocessing involves selecting key features (position, speed, heading, timestamp), grouping CAMs into overlapping sequences of 20-time steps, and applying min–max normalization. The dataset is split into 10 non-overlapping, non-IID client subsets, each simulating a unique vehicle's communication log. Only benign data are used for client-side training, whereas attack samples are reserved for evaluation, reflecting real-world constraints.

Federated learning is simulated with the Flower framework using 10 virtual clients. Experiments are run on a workstation with an NVIDIA RTX 3090 GPU, 128 GB RAM, and AMD Ryzen 7950X CPU, using Python 3.10, TensorFlow 2.11, and Flower 1.4. Each client trains locally for five epochs per round, and in each round, five randomly selected clients participate. The server aggregates models using FedAvg, weighted by sample size, over 30 communication rounds.

The LSTM autoencoder architecture consists of two encoder LSTM layers (64 and 32 units) and two decoder LSTM layers (32 and 64 units), with a latent vector size of 32 and dropout of 0.2 after each layer. Training uses the Adam optimizer (learning rate 1e-3), MSE loss, and batch size 64. Anomaly detection is based on the 95th percentile of reconstruction losses on benign validation data.

The study performed a thorough evaluation of four leading baseline models through the same training process with preprocessed VeReMi data. The baselines consist of four models, which are the centralized LSTM classifier, the 1D CNN classifier, the Random Forest classifier, and the Isolation Forest detector. The centralized models (LSTM, CNN, Random Forest) were trained in a supervised manner on a consolidated dataset containing labeled benign and attack samples. In contrast, the proposed Fed-LSTM-AE framework and the Isolation Forest baseline operate in an unsupervised setting, utilizing only benign data for training. We assess performance using accuracy, precision, recall, F1-score, and ROC-AUC and also provide per-attack analysis for all adversarial scenarios in VeReMi. This protocol ensures a rigorous and reproducible evaluation of Fed-LSTM-AE under heterogeneous, privacy-preserving federated learning conditions.

IV. RESULTS AND DISCUSSION

The proposed Fed-LSTM-AE framework undergoes empirical evaluation using the VeReMi dataset under federated learning restrictions. The framework undergoes evaluation against four baseline models, which include a centralized LSTM classifier, a 1D-CNN classifier, a Random Forest classifier, and an Isolation Forest detector. The evaluation metrics for model performance include F1-score, precision, recall, ROC-AUC, and Area Under the Precision–Recall Curve (PR-AUC). The evaluation process includes aggregate results and per-attack performance analysis to provide a detailed assessment of model robustness across different adversarial scenarios. The experiments are performed under federated conditions, which mimic actual V2X network environments with non-IID client data.

A. Overall Detection Performance

The proposed Fed-LSTM-AE framework demonstrates its overall anomaly detection performance in Table I, which includes four baseline models. The evaluation of all models used identical preprocessed inputs, which included normalized, fixed-length V2X communication sequences extracted from the VeReMi dataset.

TABLE I. AGGREGATE PERFORMANCE COMPARISON OF DETECTION MODELS ON THE VEREMI DATASET

Method	F1-score	Precision	Recall	ROC-AUC	PR-AUC
Fed-LSTM-AE (proposed)	0.963	0.997	0.931	0.966	0.979
Centralized LSTM	0.928	0.995	0.869	0.945	0.963
1D CNN	0.943	0.994	0.896	0.967	0.978
Random Forest	0.929	0.944	0.915	0.953	0.967
Isolation Forest	0.667	0.627	0.712	0.715	0.778

The evaluation results show that Fed-LSTM-AE outperforms all baseline approaches in every assessment metric. The framework reaches an F1-score of 0.963, which demonstrates a strong equilibrium between precision and recall. The precision score of 0.997 indicates an extremely low false-positive rate, which is essential for safety-critical vehicular applications. The supervised models, LSTM and 1D-CNN, achieve competitive results but require labeled attack data and centralized training, which proves impractical for privacy-sensitive and bandwidth-limited vehicular environments. The unsupervised Isolation Forest demonstrates inferior performance compared to other methods because it operates without labels and produces poor results in F1-score and precision metrics. The results indicate that Fed-LSTM-AE provides effective anomaly detection in decentralized, label-free systems, making it a practical and scalable solution for real-time V2X network intrusion detection. The model's discriminative power is further analyzed in Figure 1, which shows the reconstruction score distributions and ROC curve for Fed-LSTM-AE. The left plot demonstrates how benign and attack score densities separate distinctly through a threshold that minimizes false-positive occurrences. The ROC curve shows an AUC value of 0.970, which demonstrates the model's strong classification confidence.

B. Per-Attack Performance

To assess the robustness of the proposed framework across different adversarial scenarios, detection performance was disaggregated by attack type. Table II reports the F1-scores for each model against the four primary attack classes in the VeReMi dataset: GPS spoofing, position injection, replay attack, and Sybil attack.

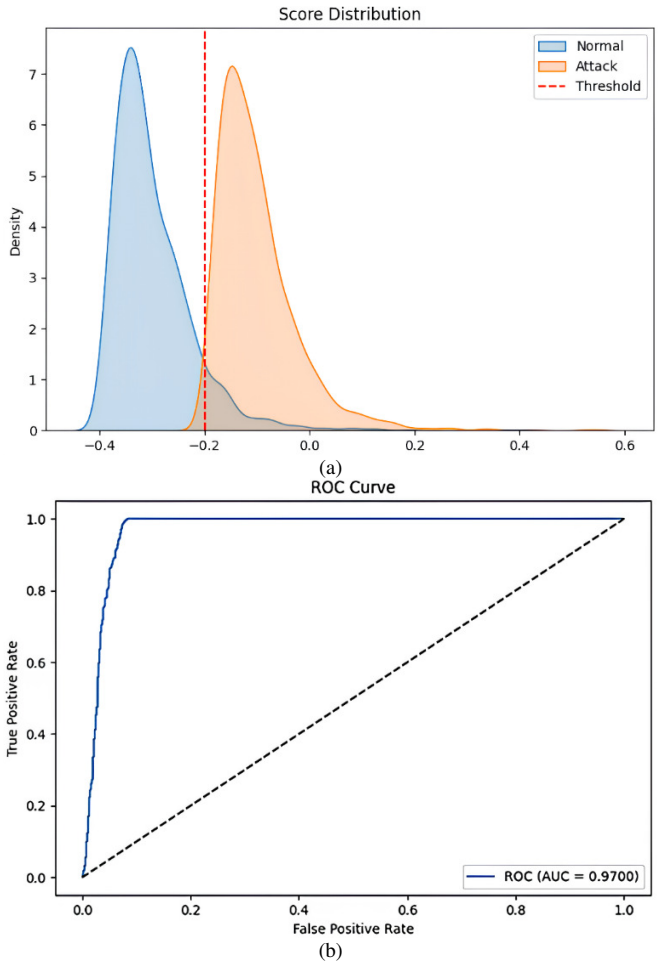

Fig. 1. Fed-LSTM-AE framework results: (a) reconstruction score distribution, (b) ROC curve.

TABLE II. PER-ATTACK F1-SCORE COMPARISON ACROSS EVALUATED MODELS

Model	GPS spoofing	Position injection	Replay attack	Sybil attack
Fed-LSTM-AE (proposed)	0.945	0.915	0.901	0.930
Centralized LSTM	0.912	0.889	0.873	0.899
1D CNN	0.883	0.851	0.837	0.871
Random Forest	0.842	0.801	0.779	0.808
Isolation Forest	0.796	0.765	0.721	0.782

As shown in Figure 2, the proposed Fed-LSTM-AE achieves the highest F1-scores across all four attack types, demonstrating strong generalization capabilities under heterogeneous threat conditions. Notably, the model performs

particularly well in detecting GPS spoofing (0.945) and Sybil attacks (0.930), which are considered among the most deceptive and disruptive forms of misbehavior in vehicular networks due to their exploitation of spatial or identity-based inconsistencies.

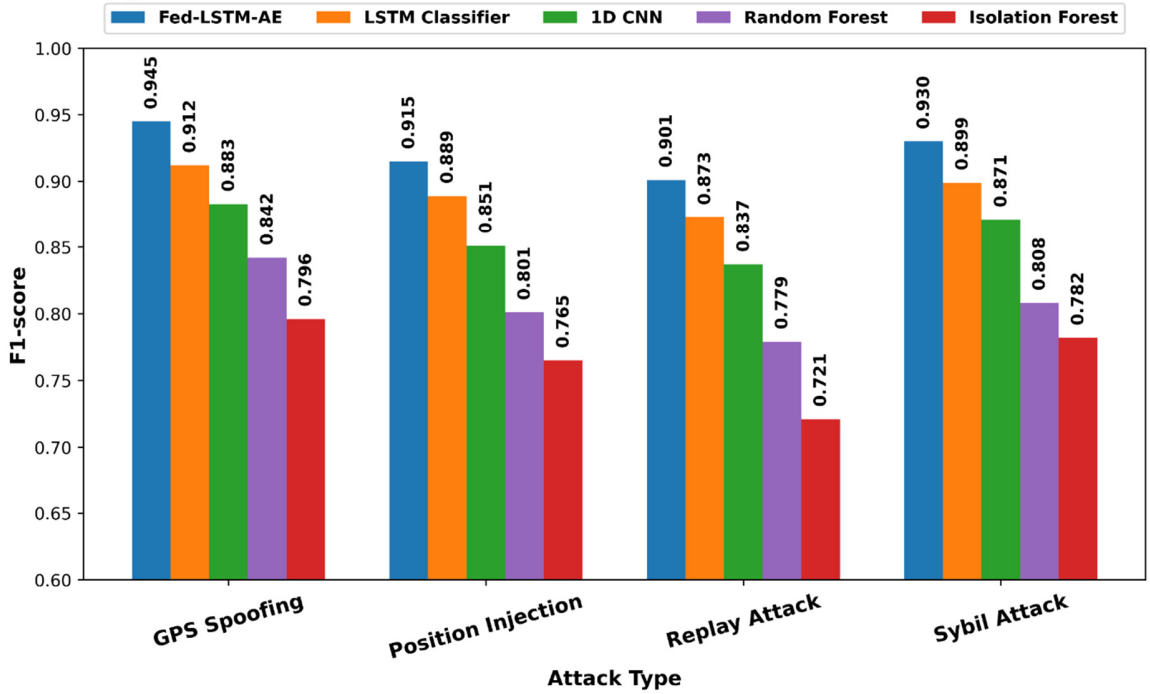

Fig. 2. F1-score comparison of different detection methods across various attack types.

Despite being trained exclusively on benign data, the Fed-LSTM-AE framework sustains high detection performance across all categories, highlighting its capacity for zero-day anomaly detection. This stands in contrast to supervised baselines (e.g., LSTM and CNN), which rely on labeled attack samples and centralized training infrastructure—conditions that may be infeasible or undesirable in real-world deployments. The relatively lower and less consistent F1-scores exhibited by Isolation Forest further emphasize the limitations of static feature-based anomaly detection and reinforce the importance of leveraging temporal representations through sequential modeling architectures such as LSTM autoencoders. These results affirm that the Fed-LSTM-AE framework is not only accurate at an aggregate level but also robust across distinct attack vectors, making it a viable candidate for practical, privacy-preserving intrusion detection in V2X networks.

The baseline models produced performance results that matched typical outcomes found in V2X security studies. Authors in [22] show that deep learning models (LSTMs, CNNs) outperform traditional machine learning methods (Random Forest) in detecting complex anomalies, as observed in this study. Authors in [23] and [24] also show that deep learning-based sequential models perform better than ensemble methods for vehicular network data analysis. Building on these findings, the Fed-LSTM-AE framework demonstrates superior performance compared to the baseline models according to the experimental results. The results also demonstrate that federated learning maintains privacy protection while reaching the same accuracy levels as conventional methods. The system achieves performance levels that match or surpass traditional centralized methods because of its distributed vehicular node knowledge acquisition process, which follows the privacy-preserving machine learning trends for the Internet of Vehicles (IoV) as described in [25].

V. CONCLUSION

The study presented a novel Federated Long Short-Term Memory Autoencoder (Fed-LSTM-AE) system for Vehicle-to-Everything (V2X) network anomaly detection through the combination of deep sequential learning with privacy-preserving federated learning. The Fed-LSTM-AE system learns normal traffic patterns locally through decentralized processing and detects anomalies through reconstruction loss to identify new attacks in real time. Experiments conducted using the VeReMi dataset demonstrate that the Fed-LSTM-AE framework surpasses all state-of-the-art supervised and unsupervised baselines by achieving high F1-scores, precision, and recall across multiple attack types, while also ensuring efficient convergence in non-Independent and Identically Distributed (non-IID) federated environments. The framework demonstrates strong performance and practical deployment potential for connected vehicle systems because it maintains privacy protection, requires minimal communication, and adapts to different environments. Future work will focus on integrating heterogeneous model architectures for different client hardware while incorporating differential privacy and secure aggregation mechanisms to enhance trust and compliance. Moreover, testing within real-world or digital twin vehicular environments will be conducted to evaluate system performance under changing conditions.

REFERENCES

[1] W. A. Mansouri, S. Asklany, S. H. Othman, and A. A. Darem, "New Scheduling Scheme in Cellular V2X Communication," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14096–14101, June 2024, https://doi.org/10.48084/etasr.7275.

[2] Z. Hussain, A. ur R. Khan, H. Mehdi, and S. M. A. Saleem, "Analysis of Device-to-Device Communication over Double-Generalized Gamma Channels," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3265–3269, Aug. 2018, https://doi.org/10.48084/etasr.2230.

[3] D. R. K. Raja, Z. A. Abas, C. S. Akula, Y. D. Kumar, G. H. Kumar, and V. Eswari, "Artificial intelligence powered internet of vehicles: securing connected vehicles in 6G," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 35, no. 1, pp. 213–221, July 2024, https://doi.org/10.11591/ijeecs.v35.i1.pp213-221.

[4] D. Wang, A. Qiu, Q. Zhou, and H. D. Schotten, "A Survey on the Role of Artificial Intelligence and Machine Learning in 6G-V2X Applications." arXiv, June 12, 2025, https://doi.org/10.48550/arXiv.2506.09512.

[5] G. H. Kumar, D. R. Kumar Raja, H. D. Varun, Navyashree, Abhishek, and S. Nandikol, "Optimizing Spatial Efficiency Through Velocity-Responsive Controller in Vehicle Platooning," in *2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions*, Bengaluru, India, 2024, pp. 1–5, https://doi.org/10.1109/CSITSS64042.2024.10816902.

[6] E. Farsimadan, L. Moradi, and F. Palmieri, "A Review on Security Challenges in V2X Communications Technology for VANETs," *IEEE Access*, vol. 13, pp. 31069–31094, 2025, https://doi.org/10.1109/ACCESS.2025.3541035.

[7] H. Alabdouli, M. S. Hassan, and A. Abdelfatah, "Enhancing Route Guidance with Integrated V2X Communication and Transportation Systems: A Review," *Smart Cities*, vol. 8, no. 1, Feb. 2025, Art. no. 24, https://doi.org/10.3390/smartcities8010024.

[8] R. Sedar, C. Kalalas, F. Vázquez-Gallego, L. Alonso, and J. Alonso-Zarate, "A Comprehensive Survey of V2X Cybersecurity Mechanisms and Future Research Paths," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 325–391, 2023, https://doi.org/10.1109/OJCOMS.2023.3239115.

[9] Z. Pethő, T. M. Kazár, Z. Szalay, and Á. Török, "Quantifying Cyber Risks: The Impact of DoS Attacks on Vehicle Safety in V2X Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 11, pp. 18591–18600, Nov. 2024, https://doi.org/10.1109/TITS.2024.3436840.

[10] M. S. Abood *et al.*, "An LSTM-Based Network Slicing Classification Future Predictive Framework for Optimized Resource Allocation in C-V2X," *IEEE Access*, vol. 11, pp. 129300–129310, 2023, https://doi.org/10.1109/ACCESS.2023.3332225.

[11] A. R. Abdellah, A. Muthanna, M. H. Essai, and A. Koucheryavy, "Deep Learning for Predicting Traffic in V2X Networks," *Applied Sciences*, vol. 12, no. 19, Oct. 2022, Art. no. 10030, https://doi.org/10.3390/app121910030.

[12] F. Khanmohammadi and R. Azmi, "Time-Series Anomaly Detection in Automated Vehicles Using D-CNN-LSTM Autoencoder," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 8, pp. 9296–9307, Aug. 2024, https://doi.org/10.1109/TITS.2024.3380263.

[13] C. Xu, H. Wu, Y. Zhang, S. Dai, H. Liu, and J. Tian, "A Real-Time Complex Road AI Perception Based on 5G-V2X for Smart City Security," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, Jan. 2022, Art. no. 4405242, https://doi.org/10.1155/2022/4405242.

[14] "VeReMi dataset." Github.io. [Online]. Available: https://veremi-dataset.github.io/.

[15] V. Elangovan, W. Xiang, and S. Liu, "A Real-Time C-V2X Beamforming Selector Based on Effective Sequence to Sequence Prediction Model Using Transitional Matrix Hard Attention," *IEEE Access*, vol. 11, pp. 10954–10965, 2023, https://doi.org/10.1109/ACCESS.2023.3241130.

[16] S. Hossain, S.-M. Senouci, B. Brik, and A. Boualouache, "A privacy-preserving Self-Supervised Learning-based intrusion detection system for 5G-V2X networks," *Ad Hoc Networks*, vol. 166, Jan. 2025, Art. no. 103674, https://doi.org/10.1016/j.adhoc.2024.103674.

[17] S. Lee, K. Koufos, C. Maple, and M. Dianati, "Application of Unsupervised Learning in Implementation of Joint Power and Index Modulation Access in V2X Systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 5, pp. 1308–1321, Oct. 2023, https://doi.org/10.1109/TCCN.2023.3276992.

[18] A. Gupta and X. Fernando, "Federated Reinforcement Learning for Collaborative Intelligence in UAV-Assisted C-V2X Communications," *Drones*, vol. 8, no. 7, July 2024, Art. no. 321, https://doi.org/10.3390/drones8070321.

[19] Y. T. Gebrezgiher, S. R. Jeremiah, S. Gritzalis, and J. H. Park, "VAE-Based Real-Time Anomaly Detection Approach for Enhanced V2X Communication Security," *Applied Sciences*, vol. 15, no. 12, June 2025, Art. no. 6739, https://doi.org/10.3390/app15126739.

[20] I.-S. Na, A. Haldorai, and N. Naik, "Federal Deep Learning Approach of Intrusion Detection System for In-Vehicle Communication Network Security," *IEEE Access*, vol. 13, pp. 2215–2228, 2025, https://doi.org/10.1109/ACCESS.2024.3521661.

[21] S. K. Kwon, J. H. Seo, J. Y. Yun, and K.-D. Kim, "Driving Behavior Classification and Sharing System Using CNN-LSTM Approaches and V2X Communication," *Applied Sciences*, vol. 11, no. 21, Nov. 2021, Art. no. 10420, https://doi.org/10.3390/app112110420.

[22] S. A. Abdel Hakeem and H. Kim, "Advancing Intrusion Detection in V2X Networks: A Comprehensive Survey on Machine Learning, Federated Learning, and Edge AI for V2X Security," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 8, pp. 11137–11205, Aug. 2025, https://doi.org/10.1109/TITS.2025.3558849.

[23] M. Almehdhar *et al.*, "Deep Learning in the Fast Lane: A Survey on Advanced Intrusion Detection Systems for Intelligent Vehicle Networks," *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 869–906, 2024, https://doi.org/10.1109/OJVT.2024.3422253.

[24] W. Aljabri, Md. A. Hamid, and R. Mosli, "Lightweight and Adaptive Data-Driven Intrusion Detection System for Autonomous Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 2, pp. 2282–2292, Feb. 2025, https://doi.org/10.1109/TITS.2024.3509459.

[25] M. Alharthi, F. Medjek, and D. Djenouri, "Ensemble Learning Approaches for Multi-Class Intrusion Detection Systems for the Internet of Vehicles (IoV): A Comprehensive Survey," *Future Internet*, vol. 17, no. 7, July 2025, Art. no. 317, https://doi.org/10.3390/fi17070317.