

Received 3 November 2025; revised 29 March 2026 and 12 May 2026; accepted 18 May 2026. Date of publication 28 May 2026; date of current version 3 June 2026.

The associate editor coordinating the review of this article and approving it for publication was S. R. Pandey.

Digital Object Identifier 10.1109/TMLCN.2026.3697341

SAFE-ADVENT: Privacy-Preserving and Resilient Distributed Federated Learning for DDoS Attacks Detection in 5G-V2X Networks

BADRE BOUSALEM^{1,2}, FAHDAH ALALYAN³ (Graduate Student Member, IEEE), VINICIUS F. SILVA⁴, SYLVAIN CHERRIER², WAEEL JAAFAR³ (Senior Member, IEEE), AND RAMI LANGAR^{2,3} (Member, IEEE)

¹Metanext, Paris 92400, France

²LIGM-CNRS, UMR 8049, University Gustave Eiffel, Marne-la-Vallée 77454, France

³Software and IT Engineering Department, École de Technologie Supérieure (ÉTS), Montreal, QC H3C1K3, Canada

⁴Computer Science Department, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais 31270-901, Brazil

CORRESPONDING AUTHOR: R. Langar (rami.langar@etsmtl.ca)

This work was supported in part by the ANR Intelligent Orchestrated Security and Privacy-Aware Slicing for 5G and Beyond Vehicular Networks (5G-INSIGHT) Project under Grant ANR-20-CE25-0015; in part by the Innovation for Defence Excellence and Security (IDEaS) Program from the Department of National Defence (DND), Canada; and in part by Metanext.

ABSTRACT Wireless communication technology is a vital enabler for 5G vehicle-to-everything (5G-V2X) communications, which are vulnerable to cyberattacks such as distributed denial of service (DDoS). DDoS detection requires sophisticated artificial intelligence (AI)-based approaches, whose main challenge is data privacy. Secured peer-to-peer (P2P) federated learning (FL) techniques could overcome this challenge by training detection models locally while sharing model parameters for secure aggregation. However, such a strategy scales poorly and incurs a high communication cost as the number of FL clients grows. In such a context, we propose SAFE-ADVENT, a secure P2P FL strategy that reduces the communication cost, incorporating transfer learning and client selection within the P2P FL system to detect DDoS attacks. By using data collected from our 5G-V2X testbed, under offline scenarios, SAFE-ADVENT converged faster than deep learning, centralized FL and SAC-based P2P FL benchmarks. Through live traffic testing and under adversarial attacks, SAFE-ADVENT outperformed all benchmarks in most scenarios. Finally, SAFE-ADVENT proved to be scalable by preserving high performance when increasing the number of FL agents.

INDEX TERMS 5G-V2X, cybersecurity, attack detection, DDoS, FL.

I. INTRODUCTION

THE advent of technologies such as vehicle-to-everything (V2X), smart infrastructure, autonomous vehicles, and the internet of things (IoT) has made wireless communication a critical enabler for their development and diffusion [1]. Additionally, a rapid deployment of applications such as virtual reality (VR) and artificial intelligence (AI) has culminated in high volumes of data [2]. As a consequence, in the past decades, wireless communication technologies have undergone numerous changes [3]. With the evolution of cellular networks from their first-generation (1G) to the fifth-generation (5G) and beyond (B5G) [4], an

extensive set of devices is now supported, facilitating access to computational resources and uninterrupted connectivity for smart and autonomous operations [1]. The growth of data communication is also eased by achieving high throughput and supporting virtualized connections. Different applications can be supported by the 5G architecture, also ensuring predefined network performance requirements [5].

To enhance mobile clients, 5G-V2X-based systems introduced new applications and network services focused on quality of experience (QoE) improvement, also ensuring safety for vehicles on the road. Such an environment requires seamless, high-performance communication

standards between vehicle users and the roadside infrastructure [6], [7], [8], [9], [10].

The 5G new radio (NR) technology is crucial to enable the deployment of 5G-V2X-based systems, since it includes the *Uu* interface for vehicle-to-network (V2N) communication. For vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-pedestrian (V2P), 5G NR also includes the PC5-based *sidelink* interface [11]. 5G networks in a “pure” setup, also defined as “stand alone” (SA), allow the use of techniques such as network slicing, software defined networking (SDN) and network function virtualization (NFV), enabling efficient network resource management and sharing between multiple applications and services, with respect to the predefined quality of service (QoS) prerequisites [12], [13].

However, several challenges come with the use of 5G NR. Indeed, the number of threats inherent to the complexity of 5G-based systems makes it hard to define system boundaries. Moreover, it is crucial to raise awareness of threats, given the rapid deployment of 5G [14]. Softwarization, virtualization, and cloudification are critical solutions for ensuring optimal network performance. However, they also introduce security vulnerabilities [15]. Therefore, the adoption of security measures is essential in 5G-based environments [7], [16].

In the 5G-V2X context, standards must also consider all security threats inherent to vehicles; otherwise, the automobile and network industries would not be able to successfully deploy such standards [17]. An example is distributed denial of service (DDoS) attacks, which may flood 5G-V2X interfaces with unexpected packets using protocols such as the transmission control protocol (TCP), the user datagram protocol (UDP), and the Internet control message protocol (ICMP). The main consequences of such attacks include poor communication performance (e.g., a decrease in the throughput or an increase in the communication delay), as well as service interruptions, thus preventing the exchange of users’ data and key control messages related to traffic management, which may lead to accidents or vehicular failures.

Assessing DDoS attacks in 5G-V2X is not simple since the communication protocol used for flooding and/or the number of attack sources and other network-related variables may constantly change. The current literature includes numerous studies focusing on DDoS attacks, but it does not yet provide full coverage of the 5G-V2X context. The use of AI-based approaches like machine learning (ML) would help define a detailed, real-time profile inherent to malicious users, ensuring privacy, security, and threat detection in 5G and 5G-V2X-based systems [7], [8], [9], [10].

Nevertheless, a challenge that remains is achieving secure and private knowledge sharing between all ML-based agents [5], [7]. An alternative to provide data privacy in such a context is centralized federated learning (FL), a distributed ML technique.

Google initially introduced centralized FL as a hierarchical training architecture comprising a central server and multiple distributed participants. Each participant trains its model locally while sharing its parameters with the central server for aggregation. By sharing model updates instead of explicitly sharing or accessing training data, centralized FL achieves its goal of ensuring privacy-preserved collaborative training, thus making it more suitable than conventional ML-based approaches. Nevertheless, it is prone to flaws such as the single point of failure and the lack of balance between different data distributions in local trainers [16], [18].

In response to the flaws mentioned above, decentralized FL-based approaches were proposed to remove the centralized server and provide protection against semi-honest participants by securing the communication of model updates. These approaches ensure privacy protection while allowing local trainers to benefit from insights derived from other trainers’ data and enhancing communication efficiency. One of the proposed methods, called peer-to-peer (P2P) FL with secure average computation (SAC), applies averaging and n -out-of- n secret partitioning to counter semi-honest participants. In this approach, while model updates are transmitted among participants in a masked form via secret partitioning, their true values are protected from inference by other peers, and raw training data is not shared or accessed by other participants. The main drawback lies in the potentially high communication cost in large-scale systems, despite its positive results in terms of accuracy [16], [18], [19], [20].

Eliminating the centralized server alone does not fully mitigate security concerns. Different risks remain such as adversarial manipulation, which leads to misclassification and inefficient convergence. Thus, beyond decentralization, it is crucial to ensure that model updates remain resilient against adversarial influence. One such threat comes from semi-honest (honest-but-curious) participants who adhere to protocol execution but attempt to infer sensitive information from exchanged updates [16], [18], [19], [20].

To circumvent all these limitations, we first proposed in [21] a framework for cyberattack detection, integrating client selection and transfer learning with a SAC-based P2P FL approach. While the original work demonstrated its feasibility in a simulated environment, this paper extends that foundation by implementing and validating the framework, here named *SAFE-ADVENT* (Secure Average computation-based **F**ederated Learning for DDoS Attacks **D**etection in 5G-**V**ehicle-to-Everything **N**etworks), on a 5G-V2X testbed. This extension enables an in-depth evaluation of its performance in a realistic vehicular network setting. The results highlight our approach’s robustness, adaptability, and suitability for practical 5G-V2X applications.

The core contribution of this work lies in the practical application and rigorous systems-level evaluation of secure distributed learning in highly constrained, adversarial

vehicular environments. Specifically, our key contributions are summarized as follows:

- 1) We design and deploy SAFE-ADVENT, a fully decentralized and privacy-preserving cyberattack detection framework tailored for edge environments. Rather than introducing new fundamental ML algorithms, this work focuses on the critical networking challenge of efficiently deploying SAC within bandwidth-constrained V2X settings. To this end, we incorporate a dynamic, performance-driven agent selection mechanism that reduces the $\mathcal{O}(TN^2)$ communication overhead inherent to conventional SAC-based P2P FL, where T denotes the total number of training rounds and N the total number of FL agents. By restricting aggregation to a subset of selected agents $K_t \leq N$ at each round, the communication complexity is reduced to $\mathcal{O}(TK_t^2)$, yielding a significant improvement when $K_t < N$.
- 2) We integrate a transfer learning mechanism that ensures non-selected (or recovering) vehicles remain synchronized with the global detection state. This architecture structurally mitigates the impact of adversarial data poisoning and severe non-independent and identically distributed (non-IID) distributions without requiring central orchestration or raw payload inspection.
- 3) We provide a rigorous empirical validation on a physical 5G-V2X testbed. We demonstrate that SAFE-ADVENT significantly reduces the actual network communication footprint compared to standard SAC while maintaining $> 90\%$ DDoS detection accuracy under active adversarial poisoning and scalability tests.

Table 1 shows a list of the most used acronyms in this paper.

The remainder of the paper is structured as follows. Section II reviews the related works, while Section III describes the system model. Sections IV and V detail the design and implementation of SAFE-ADVENT, respectively, followed by a presentation of our experimental results in Section VI. Finally, Section VII concludes this paper.

II. RELATED WORKS

ML has strong potential to address privacy, scalability, and real-time processing challenges in 5G-V2X networks. However, existing works still have important limitations for detecting large-scale cyber threats such as DDoS attacks in dynamic vehicular environments [7], [16], [22], [23], [24]. Hussain et al. [25] proposed a convolutional neural network (CNN)-based DDoS detection model using cellular call detail record (CDR) data, but evaluated it only with synthetically generated attacks. Yakan et al. [22] proposed a long short-term memory (LSTM)-based FL misbehavior detection model at the 5G edge, using the *VeReMi-extension* and *MixAll* datasets covering 19 V2X misbehavior attack

TABLE 1. List of most used acronyms.

Acronym	Description
5G	Fifth-Generation
5G-V2X	5G Vehicle-to-Everything
AI	Artificial Intelligence
AV	Autonomous Vehicle
B5G	Beyond 5G
CDR	Call Detail Record
CN	Core Network
CNN	Convolutional Neural Network
DDoS	Distributed Denial of Service
DL	Deep Learning
DP	Differential Privacy
FL	Federated Learning
ICMP	Internet Control Message Protocol
IDS	Intrusion Detection System
IID	Independent and Identically Distributed
IMSI	International Mobile Subscriber Identity
LSTM	Long Short-Term Memory
MAC	Medium Access Control
MEC	Mobile Edge Computing
ML	Machine Learning
MVNO	Mobile Virtual Network Operator
NR	New Radio
O-RAN	Open Radio Access Network
OAI	OpenAirInterface
P2P	Peer-to-Peer
PCAP	Packet Capture
PRR	Packet Reception Rate
RAN	Radio Access Network
RIC	RAN Intelligent Controller
RF	Radio Frequency
RL	Reinforcement Learning
SA	Stand Alone
SAC	Secure Average Computation
SDN	Software Defined Networking
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UE	User Equipment
UPF	User Plane Function
USRP	Universal Software Radio Peripheral
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
VNF	Virtual Network Function
XGBoost	Extreme Gradient Boosting

types. Despite using different methods and datasets, both studies relied on non-real-time data, thus limiting their applicability to dynamic 5G-V2X environments.

Other works also addressed DDoS detection in 5G-V2X. Asensio et al. [26] proposed a reinforcement learning (RL)-based zero-touch service management (ZSM) framework for mobile edge computing (MEC)-enabled 5G-V2X, where slices are allocated according to attack severity. However, the evaluation relied on the offline *VeReMi* dataset, and the testbed was not clearly identified as physical or emulated. Han et al. [27] used feature selection to reduce detection time on the *InSDN* and *CICIDS2017/2018* datasets, but their simulated SDN setup did not include actual 5G infrastructure or 5G-based behavior. Anbalagan et al. [28] proposed a CNN-based intrusion detection system (IDS) for a simulated 5G-V2X network, with hyperparameter optimization and an autonomous vehicle (AV) isolation mechanism. However, it showed higher latency with more AVs and lacked secure alert communication. Overall, these studies relied on offline or simulated settings, limiting feasibility in real-time and adversarial 5G-V2X environments.

In the FL context, Abouaomar et al. [18] proposed a federated deep reinforcement learning (FDRL) method for open radio access network (O-RAN) slicing in 6G,

supporting mobile virtual network operator (MVNO) collaboration but overlooking privacy, adversarial, and cost aspects. Wijethilaka et al. [29] proposed an FL-enabled security orchestrator (FLeSO) for 5G network slice security, but it suffers from scalability and accuracy issues with many slices. Le et al. [20] designed an adaptive privacy-preserving FL (Ada-PPFL) scheme with differential privacy, enabling client-level protection and adversarial robustness, but excessive noise degrades prediction accuracy. Turina et al. [30] explored federated split learning with privacy gains through distance correlation, but it incurs runtime overhead and limited attack modeling.

Other works applied centralized FL to 5G DDoS detection. Maiga et al. [31] proposed a hybrid FL-based model combining extreme gradient boosting (XGBoost) for feature selection with a deep neural network for detection in 5G core virtual network functions (VNFs). Although the model achieved high accuracy, it was not evaluated under non-IID data, and detection time was reported only for the proposed method, limiting cross-study comparison. Sheikhi et al. [32] explored an unsupervised FL-based model using autoencoders for DDoS detection in 5G core networks (CNs). Their setup trained each client on a single attack type, indirectly addressing non-IID data challenges, but considered only two clients, raising concerns about scalability and generalizability.

Despite the achieved results, none of the above works explicitly evaluated adversarial robustness or communication overhead reduction, leaving gaps in assessing scalability and resilience in real 5G-V2X scenarios. Liu et al. [33] proposed a secure FL framework for 5G that integrates blockchain and local differential privacy (DP) to defend against poisoning and inference attacks. Smart contracts identify unreliable participants, while noise added to local updates preserves privacy. However, the framework assumes a trusted central aggregator, creating a single point of failure and ignoring communication cost. Attanayaka et al. [16] proposed a SAC-based P2P FL framework for anomaly detection in O-RAN. Their hierarchical clustered design reduces communication overhead compared to conventional P2P FL, but relies on static client clusters, limiting adaptability in dynamic and mobile environments. To further address overhead in decentralized FL, we proposed in [34] a secure P2P FL approach for B5G O-RAN, focusing on overhead reduction. FL was orchestrated at near-real-time radio access network (RAN) intelligent controllers (RICs) instead of user equipments (UEs). Although aligned with O-RAN use cases, this design limits the capture of fine-grained device-level attack data required in V2X environments. Moreover, it does not consider compromised FL agents and was evaluated only through simulations and an emulated testbed with virtual UEs. Hence, its real-world applicability remains limited.

Building upon our earlier work in [21], which introduced a SAC-based P2P FL framework for cyberattack detection integrating client selection and transfer learning, this study

extends the approach by deploying and validating it, now called SAFE-ADVENT, on a physical 5G-V2X testbed. Unlike [34], where FL was executed at the level of near-real-time RICs, this work moves training directly onto robot vehicles and performs P2P aggregation over emulated PC5/Uu interfaces (cf. Section V). Furthermore, (i) the model is trained in TensorFlow, converted to *LiteRT* [35], and exported as a compact *.tflite* file for on-device inference, reducing communication overhead per federated round and minimizing RAM and latency on the vehicles; (ii) the standard *LiteRT* converter with default graph optimizations is used, maintaining compatibility with post-training quantization (FP16/INT8) to ensure inference on resource-constrained hardware; (iii) explicit wireless channel and adversarial models are introduced, the latter accounting for compromised clients (cf. Section III); (iv) feature selection and engineering are made explicit using *Argus flow features* and short-window entropy; and (v) *SAFE-ADVENT* is validated on a physical 5G-V2X testbed under both online adversarial and scalability scenarios (cf. Section V).

Table 2 compares all works above using nine criteria: 1) Use of FL, 2) use of distributed FL, 3) privacy preservation, 4) adversarial resistance, 5) reduction of communication overhead, 6) validation environment, 7) context, 8) strengths, and 9) limitations. The last row presents our work, highlighting secure FL for 5G-V2X with physical testbed validation and adversarial resilience.

As summarized in Table 2, *SAFE-ADVENT* is positioned to jointly cover the full set of practical requirements for 5G-V2X DDoS detection that are only partially addressed by prior works, namely: (i) FL with *distributed* (P2P) execution, (ii) privacy preservation through secure aggregation, (iii) explicit resilience against adversarial behaviors, (iv) communication-overhead reduction mechanisms, and (v) validation in a realistic 5G-V2X context that includes a physical testbed component. In addition, *SAFE-ADVENT* targets scalability and malicious-agent handling in adversarial vehicular settings, which are typically not evaluated together in existing solutions.

III. SYSTEM MODEL

This section presents a detailed view of our system model. Subsection A outlines the main concepts inherent to SAC-based FL. Subsection B introduces the Adversarial model, and Subsection C describes the channel model. Finally, Subsection D presents the problem formulation and quantitative performance metrics that lead to the design of *SAFE-ADVENT*.

A. SAC-BASED FL

Introduced by [19], in essence, the input/output of SAC-based decentralized FL approaches is the same as centralized FL approaches based on an aggregation server, i.e., both average the participants' model updates into a global model. However, their transmissions within the FL system differ. In particular, SAC relies on two mechanisms, namely

TABLE 2. Comparison of related works.

Ref.	FL	Dist. FL	Priv.-Preserv.	Adv. Resist.	Reduced Comm. Overhead	Valid. Env.	Context	Strengths	Limitations
[21]*	✓	✓	✓	—	✓	<i>S</i>	5G/B5G cyberattack detection.	Secure and efficient P2P FL.	No lightweight deployment; No Argus and entropy-based feature selection; Based on a public dataset.
[16]	✓	✓	✓	—	✓	<i>S</i>	O-RAN	SAC-based P2P FL anomaly detection.	Static clustering; Limited adaptability to dynamic environments.
[18]	✓	✗	✗	—	—	<i>S</i>	O-RAN for 6G slicing	Better robustness, service level agreement (SLA) satisfaction, handles non-IID across MVNOs.	Centralized, lacks explicit privacy/adversarial focus, no cost analysis.
[22]	✓	✗	✗	—	—	<i>S + ET</i>	5G-FL for V2X Edge	LSTM-based misbehavior detection.	Offline and online scenarios based on offline datasets.
[25]	✗	✗	✗	—	—	<i>S</i>	5G cyber physical system (CPS)	CNN-based DDoS detection with cellular CDR data.	Significant false positives in sudden activity bursts.
[26]	✗	✗	✗	—	—	<i>UT</i>	5G-MEC-enabled V2X	RL-based DDoS detection; Automated resource allocation.	High orchestration overhead; Evaluated only on an offline dataset.
[27]	✗	✗	✗	—	—	<i>S</i>	SDN IDS	ML-based DDoS detection with feature selection; Reduced detection time.	SDN simulation without 5G network considerations.
[28]	✗	✗	✗	—	—	<i>S</i>	5G-V2X IDS for AVs	CNN-based detection with hyperparameter optimization; AV isolation mechanism.	Latency increases with number of AVs; No secure communication for alerts.
[29]	✓	✗	✗	—	—	<i>S + ET</i>	5G network slicing	Improves slice security.	Accuracy drops with more slices; Requires many rounds; No adversarial/cost analysis.
[20]	✓	✗	✓	✓	—	<i>S</i>	FL Security	Client-level privacy protection; Adversarial robustness.	Prediction accuracy degrades with excessive noise.
[30]	✓	✗	✓	✗	✗	<i>S + ET</i>	Federated split learning	Good convergence; Privacy gain via distance correlation.	Runtime overhead; Limited tested attack model.
[31]	✓	✗	✗	—	—	<i>S</i>	5G-FL for VNFs	Hybrid XGBoost-FL for DDoS detection; XGBoost enables feature selection.	Does not address non-IIDness across VNFs; Based on offline dataset.
[32]	✓	✗	✗	—	—	<i>S + ET</i>	5G-FL for Core	FL unsupervised DDoS detection; Autoencoder enables feature extraction.	Trained on two FL agents only.
[33]	✓	✗	✓	✓	—	<i>S</i>	5G-FL with Blockchain	Blockchain and DP protect against poisoning and inference attacks.	Increased communication and computation costs with number of FL agents.
[34]	✓	✓	✓	—	✓	<i>S + ET</i>	B5G O-RAN	Secure P2P FL for B5G O-RAN; Reduced communication cost.	Simulated/emulated testbed; Virtual UEs only; No adversarial attacks.
Our Work	✓	✓	✓	✓	✓	<i>S* + ET + PT</i>	5G-FL for V2X	Secure FL for 5G-V2X; Handles scalability and adversarial attacks; Detects malicious agents.	Higher overhead than centralized FL; Limited number of physical UEs in testbed.

*Original feasibility was demonstrated in a simulated environment in [21].

S = Simulation ; *UT* = Unspecified Testbed ; *PT* = Physical Testbed ; *ET* = Emulated Testbed.

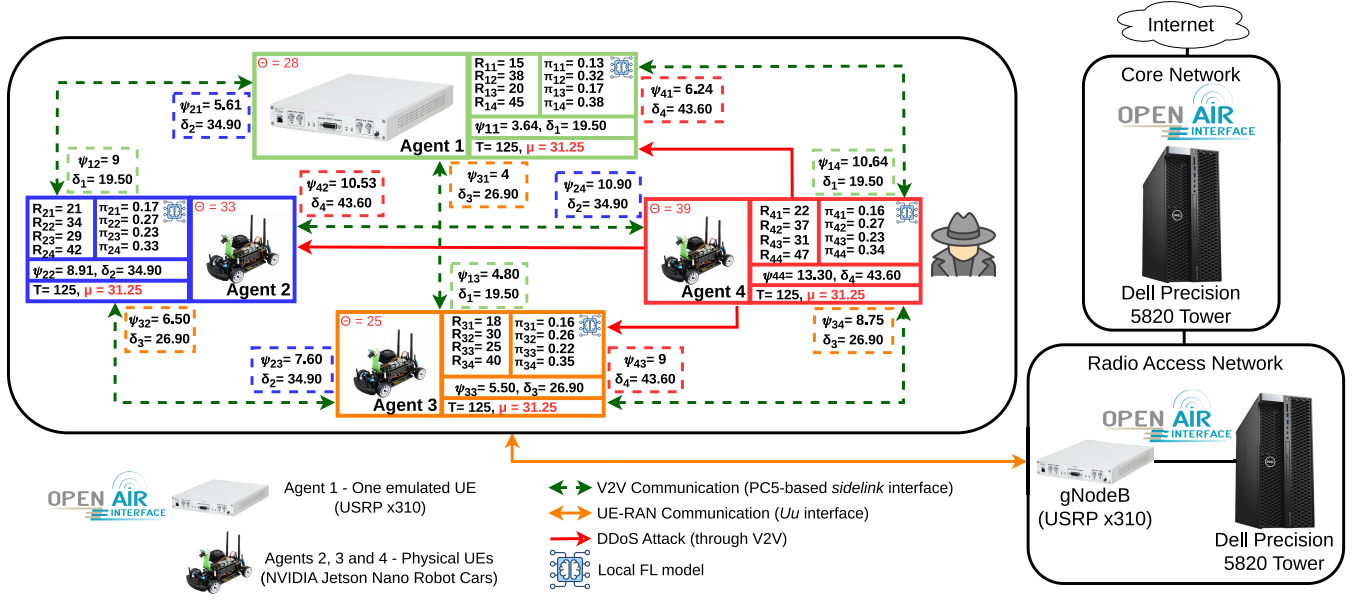


Fig. 1. 5G-V2X testbed: Main hardware and software modules. Example scenario with four agents, with Agent 4 as a malicious UE.

lightweight n -out-of- n secret partitioning and secure multi-party average calculation that leverage the partitioning method.

In the secret partitioning method, each $Agent_j$ ($j \in \{1, \dots, N\}$) generates N positive random numbers $\{R_{j1}, \dots, R_{jN}\}$, which are then used to compute the percentage distributions, denoted as $\pi_{j1}, \dots, \pi_{jN}$:

$$\pi_{ji} = \frac{R_{ji}}{\sum_{k=1}^N R_{jk}}, \quad (i, j) \in \{1, \dots, N\}^2, \quad (1)$$

where N represents the total number of Agents in the P2P FL environment. The percentage distributions are then used to obtain N partial weights $\{\psi_{j1}, \dots, \psi_{jN}\}$ given by

$$\psi_{ji} = \theta_j \times \pi_i, \quad (i, j) \in \{1, \dots, N\}^2, \quad (2)$$

where θ_j denotes the model update of $Agent_j$, and $\Theta^{(t)} = \{\theta_1, \dots, \theta_N\}$ contains the parameters of all trained models in the current round t .

For instance, suppose $Agent_1$ holds $\theta_1 = 28$. This value is divided into partial weights $\psi_{11} = 3.64$, $\psi_{12} = 9$, $\psi_{13} = 4.80$, and $\psi_{14} = 10.64$, as depicted in Fig. 1. These partitioned weights, $\{\psi_{j1}, \dots, \psi_{jN}\}$, are then applied in the next step, i.e., the multi-party average calculation. In particular, each $Agent_j$ retains its own partial weight ψ_{jj} and transmits ψ_{ji} to $Agent_i \forall i \in \{1, \dots, N\}$ and $i \neq j$. For illustration, $Agent_1$ keeps $\psi_{11} = 3.64$, shares $\psi_{12} = 9$ with $Agent_2$, and forwards $\psi_{13} = 4.80$ to $Agent_3$ while $\psi_{14} = 10.64$ to $Agent_4$. Subsequently, each $Agent_j$ computes its subtotal δ_j as follows:

$$\delta_j = \sum_{i=1}^N \psi_{ji}, \quad j \in \{1, \dots, N\}. \quad (3)$$

TABLE 3. Communication cost and complexity comparison of SAFE-ADVENT and all benchmarks.

Parameter	DL	Cent. FL	SAC P2P FL	SAFE-ADVENT
Comm. Complexity	$\mathcal{O}(TN)^\dagger$	$\mathcal{O}(TN)$	$\mathcal{O}(TN^2)$	$\mathcal{O}(TK_t^2)$, $K_t \leq N$
Comm. Cost [‡]	Eq. (10)	$\Theta(N+1)T$	$2\Theta(N(N-1)T)$	Eq. (9)

Finally, each agent $Agent_j$ aggregates the subtotals to obtain the total T and computes the average weight, μ , according to the following equations:

$$T = \sum_{j=1}^N \delta_j \text{ and } \mu = \frac{T}{N}. \quad (4)$$

B. ADVERSARIAL MODEL

This section outlines the adversarial attacker assumptions used to evaluate SAFE-ADVENT. We consider data-plane threats that are realistic in 5G-V2X networks, such as: (i) *Network-level flow injection* aimed at resource exhaustion and evasion; (ii) *Compromised clients* that *poison local training*; and (iii) *Honest-but-curious* peers as a privacy threat. Control-plane forgery (breaking the message's integrity), radio frequency (RF) jamming, and Sybil attacks are beyond our scope [36], [37].

1) SETTING AND EXTREME CASE

We assess robustness in a highly challenging non-IID setting, where in each round t , a vehicle j experiences a local attack rate $\alpha_{j,t} \in [0.2, 0.4]$. During experiments, the adversary injects a total budget of $B_{inj} = 200,000$ noisy or misleading packet samples into the *live* traffic stream, distributed over rounds and vehicles approximately in

TABLE 4. Data collection parameters.

Parameter	Values
Number of malicious UEs	0 (benign TCP and UDP traffic only), 2, 3, 4
Malicious traffic type	TCP SYN only, UDP only, Mixed (TCP SYN and UDP simultaneously)
Number of malicious packets per second (per traffic type)	1,000
Source port range for malicious packets	100-65,535 (increased after each packet sent)
Destination port for malicious packets	80, 22, 443, 53 (one port assigned to each malicious UE)
Benign TCP and UDP traffic throughput	5 Mbps
Data collection duration per scenario	100 seconds

TABLE 5. Summary of experimental and hyperparameter settings.

Hardware & Software Environment	
UEs	NVIDIA Jetson Nano (4GB RAM, Quad-core)
CN/RAN Server	Dell Precision 5820 (Intel Xeon, 128 GB RAM)
Software Stack	Ubuntu 20.04, Python 3.8, TensorFlow 2.x, OAI
Radio Interfaces	USRP X310 (Uu), Wi-Fi Ad-Hoc (PC5)
ML Hyperparameters	
Local Data Split	80% Training / 20% Validation
Global Evaluation	Aggregated Local Validation
Optimizer	ADAM
Learning Rate	10^{-4}
Batch Size	100
Local Episodes (FL models)	10
Total Rounds (FL models)	8
Global Epochs (DL Benchmark)	80
Activation & Reg.	ReLU, L1 Regularization

proportion to their observed traffic volumes. These conditions match the robustness experiment reported in Table 10.

2) ADVERSARY CLASSES AND GOALS

Three classes of adversarial entities are considered:

- *Network-level injectors*, responsible for synthesizing and mixing attack flows into normal traffic, focusing on resource exhaustion (reducing packet

TABLE 6. SAFE-ADVENT's performance: Live traffic testing round.

Metric	Car 1	Car 2	Car 3	Emulated UE
Accuracy (%)	95.32	94.90	94.75	95.20
Precision (%)	95.20	94.90	94.70	95.20
Recall (%)	95.40	95	94.80	95.30
F1-Score (%)	95.35	94.95	94.78	95.25
Commun. Cost (MB)	11.3	15.6	21.7	13.1
Average Accuracy (%)	95.04			
Average F1-Score (%)	95.08			

reception rate—PRR—and throughput) and targeted evasion (causing attack flows to be misclassified as benign);

- *Compromised FL clients*, which perform local training on corrupted data and transmit biased updates to degrade the global model;
- *Honest-but-curious peers*, which comply with the learning protocol while attempting to extract information from other peers' updates, posing a privacy threat rather than an integrity one.

The primary objectives of these adversaries are: (i) reducing recall or true positive rate at a fixed low false positive rate, corresponding to the system's safety operating point, and (ii) slowing or destabilizing convergence of the FL model under non-IID conditions. These assumptions define the threat landscape against which we evaluate SAFE-ADVENT.

3) KNOWLEDGE, CAPABILITIES, AND LIMITS

The adversary is treated as a *black-box* with respect to model parameters, having no internal access beyond what an ordinary UE receives (i.e., the disseminated $\Theta^{(t)}$). It can generate and inject flows at the data plane over the emulated transport (cf. Section III-C), and a bounded fraction β of clients may be compromised in poisoning experiments.

4) ATTACK CONSTRUCTION: INJECTION, TARGETED EVASION AND DATA POISONING

Let x denote the UE's flow feature vector, comprising Argus rates, counts, flags, and short-window entropy. We generate modified samples $x' = \mathcal{T}(x; \eta)$, where η represents bounded perturbation factors. The transformation keeps all fields valid (no illegal protocol values) and shifts statistics toward benign-like behavior. Three main operations are applied:

- Header fields and flag ratios are perturbed within valid sets, and inter-arrival patterns are jittered;
- Byte and packet rates, as well as average packet sizes are scaled within predefined bounds;
- Short-window diversity is adjusted by resampling destination ports and size histograms to manipulate entropy.

For *targeted evasion*, the injected samples are presented to the detector during inference as benign-like flows. For *data poisoning*, compromised clients include a fraction $\rho_{j,t} \leq \alpha_{j,t}$

TABLE 7. Performance evaluation (Live traffic, IID scenarios).

Scenario	Approach	Detected DDoS (%)	Detected Benign (%)	Detected Attacker IP Address
Benign Only	Centralized FL	0.00%	100.00%	—
	SAC-based P2P FL	0.00%	100.00%	—
	DL Approach	0.00%	100.00%	—
	SAFE-ADVENT	0.00%	100.00%	—
Attack TCP Only	Centralized FL	100.00%	0.00%	10.1.1.1
	SAC-based P2P FL	100.00%	0.00%	10.1.1.1
	DL Approach	100.00%	0.00%	10.1.1.1
	SAFE-ADVENT	100.00%	0.00%	10.1.1.1
Attack UDP Only	Centralized FL	100.00%	0.00%	10.1.1.2
	SAC-based P2P FL	100.00%	0.00%	10.1.1.2
	DL Approach	100.00%	0.00%	10.1.1.2
	SAFE-ADVENT	100.00%	0.00%	10.1.1.2
Attack TCP & UDP	Centralized FL	100.00%	0.00%	10.1.1.3
	SAC-based P2P FL	100.00%	0.00%	10.1.1.3
	DL Approach	93.42%	6.58%	10.1.1.3, 10.1.1.8*
	SAFE-ADVENT	100.00%	0.00%	10.1.1.3
Benign & Attack TCP	Centralized FL	87.24%	12.76%	10.1.1.4
	SAC-based P2P FL	86.37%	13.63%	10.1.1.4, 10.1.1.7*
	DL Approach	84.79%	15.21%	10.1.1.4, 10.1.1.8*, 10.1.1.9*
	SAFE-ADVENT	93.06%	6.94%	10.1.1.4
Benign & Attack UDP	Centralized FL	88.12%	11.88%	10.1.1.5
	SAC-based P2P FL	87.46%	12.54%	10.1.1.5, 10.1.1.8*
	DL Approach	85.93%	14.07%	10.1.1.5, 10.1.1.8*, 10.1.1.9*
	SAFE-ADVENT	94.36%	5.64%	10.1.1.5
Benign & Attack TCP & UDP	Centralized FL	91.49%	8.51%	10.1.1.6
	SAC-based P2P FL	89.33%	10.67%	10.1.1.6, 10.1.1.7*
	DL Approach	88.66%	11.34%	10.1.1.6, 10.1.1.8*, 10.1.1.9*
	SAFE-ADVENT	95.72%	4.28%	10.1.1.6

*Misclassified IP addresses: These indicate IP addresses that were incorrectly classified as an attacker or benign.

of transformed samples in local training with incorrect labels, aiming to increase false negatives (reduce recall) while maintaining a low false positive rate.

C. CHANNEL MODEL

We use an emulated PC5-based *sidelink* interface (supporting unicast, groupcast, and broadcast) for peer-to-peer exchanges, while keeping an emulated 5G *Uu* interface as a fallback channel. In our testbed, the *sidelink* is emulated over Wi-Fi operating in ad-hoc mode, while the *Uu* interface is implemented using open-source 5G software and universal software radio peripheral (USRP) cards. The emulation framework keeps the send/receive patterns, message ordering, round timing, integrity checks, and timeouts specified by the learning protocol [38], [39].

1) ROUND TIMING AND NEIGHBOR DISCOVERY

Training proceeds in rounds of duration Δt . Each round starts with an admission window Δt_{aw} , during which nodes announce their availability and measure link quality to neighboring peers through short probe exchanges. Only responding nodes within Δt_{aw} are considered eligible for selection in that round [39].

2) LINK QUALITY AND ELIGIBILITY

For each candidate j , a link quality score $\ell_j \in [0, 1]$ is computed from Wi-Fi statistics (e.g., PRR and, when available, received signal strength indicator–RSSI/signal-to-noise ratio–SNR) during the interval Δt_{aw} . Nodes with $\ell_j < \ell_{min}$ are not selected to minimize mid-round dropouts. The selection process may also weight ℓ_j jointly with performance metrics (a_j, f_j) , as described in [39].

3) LOSS AND CONTENTION

We employ carrier sense multiple access with collision avoidance (CSMA/CA) as the medium access control (MAC) mechanism, where collisions and backoffs are reflected in the link quality metric ℓ_j . Within a round, residual losses for peer j are modeled as IID packet drops with probability $p_j = 1 - \text{PRR}_j$. Groupcast corresponds to MAC-layer multicast without ACKs, making it less reliable than unicast. When high reliability is required (e.g., for control messages or SAC shares), unicast is employed. For model dissemination, we favor groupcast or unicast-to-many with a small redundancy factor [39].

TABLE 8. Performance evaluation (Live traffic, Non-IID 30%–60% Attack rate scenarios).

Scenario	Approach	Detected DDoS (%)	Detected Benign (%)	Detected Attacker IP Address
Benign Only	Centralized FL	0.00%	100.00%	–
	SAC-based P2P FL	0.00%	100.00%	–
	DL Approach	0.00%	100.00%	–
	SAFE-ADVENT	0.00%	100.00%	–
Attack TCP Only	Centralized FL	100.00%	0.00%	10.1.1.1
	SAC-based P2P FL	100.00%	0.00%	10.1.1.1
	DL Approach	100.00%	0.00%	10.1.1.1
	SAFE-ADVENT	100.00%	0.00%	10.1.1.1
Attack UDP Only	Centralized FL	100.00%	0.00%	10.1.1.2
	SAC-based P2P FL	100.00%	0.00%	10.1.1.2
	DL Approach	100.00%	0.00%	10.1.1.2
	SAFE-ADVENT	100.00%	0.00%	10.1.1.2
Attack TCP & UDP	Centralized FL	100.00%	0.00%	10.1.1.3
	SAC-based P2P FL	100.00%	0.00%	10.1.1.3
	DL Approach	92.34%	7.66%	10.1.1.3
	SAFE-ADVENT	100.00%	0.00%	10.1.1.3
Benign & Attack TCP	Centralized FL	89.45%	10.55%	10.1.1.4
	SAC-based P2P FL	90.32%	9.68%	10.1.1.4, 10.1.1.7*
	DL Approach	80.24%	19.76%	10.1.1.4, 10.1.1.8*, 10.1.1.9*
	SAFE-ADVENT	94.58%	5.42%	10.1.1.4
Benign & Attack UDP	Centralized FL	88.77%	11.23%	10.1.1.5, 10.1.1.8*
	SAC-based P2P FL	90.94%	9.06%	10.1.1.5
	DL Approach	78.46%	21.54%	10.1.1.5, 10.1.1.8*, 10.1.1.9*
	SAFE-ADVENT	95.16%	4.84%	10.1.1.5
Benign & Attack TCP & UDP	Centralized FL	90.02%	9.98%	10.1.1.6, 10.1.1.8*
	SAC-based P2P FL	91.64%	8.36%	10.1.1.6
	DL Approach	71.53%	28.47%	10.1.1.7*, 10.1.1.8*, 10.1.1.9*
	SAFE-ADVENT	94.92%	5.08%	10.1.1.6

*Misclassified as an attacker (false positive) or as a benign user (false negative).

4) COMMUNICATION PATTERN

Within the selected set $\mathcal{A}_{\text{sel}}$ (of size S), peers exchange masked SAC shares over Wi-Fi using unicast or groupcast, and subsequently reconstruct the aggregate. The aggregated model $\Theta^{(t)}$ is then transmitted once to the non-selected peers (of size $N-S$), preferably via groupcast to minimize redundant transmissions. When *sidelink* quality degrades, dissemination can fall back to the emulated Uu interface [38], [39].

5) TIMEOUTS AND CHURN

Secure aggregation operates with a timeout τ_{SAC} . If certain SAC shares are missing when τ_{SAC} expires, the corresponding senders are excluded, and aggregation resumes with the remaining subset, ensuring the round remains bounded in time. Late joiners perform inference-only until the next admission window [40].

D. PROBLEM FORMULATION AND QUANTITATIVE PERFORMANCE METRICS

To rigorously define the main goals of SAFE-ADVENT, we formulate the DDoS attack detection task within the

5G-V2X network as a constrained distributed optimization problem.

Let $\mathcal{D} = \bigcup_{j=1}^N D_j$ represent the global training dataset, where D_j is the local training dataset comprising the network flows captured by *Agent_j*. Furthermore, let U_j denote the local validation dataset held by *Agent_j*.

The primary goal is to learn an optimal global anomaly detection model, parameterized by Θ^* , that minimizes the aggregate empirical loss $\mathcal{L}(\Theta)$ exclusively across all agents' local training data. This objective can be formulated as follows:

$$\Theta^* = \arg \min_{\Theta} \sum_{j=1}^N \frac{|D_j|}{|\mathcal{D}|} \mathcal{L}_j(\Theta; D_j), \quad (5)$$

where $\mathcal{L}_j$ denotes the local classification loss (e.g., categorical cross-entropy for binary DDoS detection) computed by *Agent_j*.

In a standard P2P FL environment, solving (5) using the SAC protocol requires $\mathcal{O}(N^2)$ communication exchanges per round. However, in highly dynamic 5G-V2X environments, the available PC5 *sidelink* bandwidth is strictly bounded. Thus, our optimization is subject to a rigid communication

TABLE 9. Performance evaluation (Live traffic, Non-IID 20%–40% Attack rate scenarios).

Scenario	Method	Detected DDoS	Detected Benign	Detected Attacker IP Address
Benign Only	Centralized FL	0.00%	100.00%	–
	SAC-based P2P FL	0.00%	100.00%	–
	DL Approach	0.00%	100.00%	–
	SAFE-ADVENT	0.00%	100.00%	–
Attack TCP Only	Centralized FL	100.00%	0.00%	10.1.1.1
	SAC-based P2P FL	100.00%	0.00%	10.1.1.1
	DL Approach	100.00%	0.00%	10.1.1.1
	SAFE-ADVENT	100.00%	0.00%	10.1.1.1
Attack UDP Only	Centralized FL	100.00%	0.00%	10.1.1.2
	SAC-based P2P FL	100.00%	0.00%	10.1.1.2
	DL Approach	100.00%	0.00%	10.1.1.2
	SAFE-ADVENT	100.00%	0.00%	10.1.1.2
Attack TCP & UDP	Centralized FL	100.00%	0.00%	10.1.1.3
	SAC-based P2P FL	100.00%	0.00%	10.1.1.3
	DL Approach	100.00%	0.00%	10.1.1.3
	SAFE-ADVENT	100.00%	0.00%	10.1.1.3
Benign & Attack TCP	Centralized FL	1.14%	98.86%	No attacker identified*
	SAC-based P2P FL	5.02%	94.98%	No attacker identified*
	DL Approach	0%	100%	No attacker identified*
	SAFE-ADVENT	89.91%	10.09%	10.1.1.4
Benign & Attack UDP	Centralized FL	85.98%	14.02%	10.1.1.5, 10.1.1.9*
	SAC-based P2P FL	88.29%	11.71%	10.1.1.5
	DL Approach	81.30%	18.70%	10.1.1.5, 10.1.1.8*
	SAFE-ADVENT	90.15%	9.85%	10.1.1.5
Benign & Attack TCP & UDP	Centralized FL	29.12%	70.88%	No attacker identified*
	SAC-based P2P FL	33.50%	66.50%	10.1.1.7*, 10.1.1.8*
	DL Approach	24.80%	75.20%	No attacker identified*
	SAFE-ADVENT	95.35%	4.65%	10.1.1.6

*Misclassified as an attacker (false positive) or as a benign user (false negative).

TABLE 10. Impact of adversarial attacks on performances.

Metric	Central. FL	SAC-based P2P FL	DL	SAFE-ADVENT
Accuracy (%)	24.75	29.12	15.82	90.20
Precision (%)	24.10	28.50	15.20	89.70
Recall (%)	24.90	29.40	16.30	90.80
F1-Score (%)	24.40	28.90	15.60	90.10

constraint:

$$C_{\text{comm}}^{(t)} \leq C_{\text{max}}, \quad \forall t \in \{1, \dots, T\}, \quad (6)$$

where $C_{\text{comm}}^{(t)}$ is the total communication overhead incurred during aggregation in round t , and C_{max} is the strict upper bound of allowable bandwidth. Furthermore, the optimization must remain stable even when a fraction of the local training datasets are maliciously poisoned (as detailed in Subsection B).

Consequently, our framework is designed to quantitatively optimize the following metrics:

- **Communication Cost Minimization:** Reducing the total exchanged data (measured in MB) per round below C_{max} by dynamically selecting a reliable subset of agents $K \leq N$.

- **Robust F1-Score and Accuracy Maximization:** Maximizing the harmonic mean of precision and recall (F1-score) on the local validation datasets U_j . These validation metrics serve a dual purpose: they act as the filtering criteria during the dynamic agent selection phase and are aggregated to evaluate the global model's generalization, ensuring high detection rates ($> 90\%$) even under adversarial poisoning.

IV. PROPOSAL: SAFE-ADVENT

A. DESIGN OF SAFE-ADVENT

Our system comprises N agents, each denoted by Agent_j ($j \in \{1, \dots, N\}$). Each agent maintains two datasets: a local training dataset D_j and a local validation dataset U_j , drawn respectively from the global sets $\mathcal{D} = \{D_1, \dots, D_N\}$ and $\mathcal{U} = \{U_1, \dots, U_N\}$. The deployment of the SAFE-ADVENT framework within our 5G-V2X testbed, illustrated in algorithms 1 and 2, is structured into seven sequential steps, detailed as follows.

- 1) **Feature Selection & Engineering:** Each agent extracts Argus flow features (rates, counts, and flags) and computes short-window entropy statistics using a window of $W = 2$ seconds and a stride of

$s = 0.5$ second. Entropy is derived from a packet-size histogram with $B_{\text{size}} = 32$ bins and a destination-port histogram with $B_{\text{port}} = 16$ bins. The resulting inputs are then normalized. This step is executed during initialization and before each training round, when features are recomputed for the current observation window (Algo. 1, lines 1-2).

- 2) **Initial Training:** Each agent $Agent_j$ (i.e., a UE in the 5G-V2X network) begins by training its model locally, based on its training dataset D_j , producing an initial parameter set θ_j . Subsequently, it evaluates θ_j on the validation dataset U_j to obtain performance metrics, e.g., f_j (F1 score) and a_j (accuracy). This initial phase ensures that each agent learns from its respective data before participating in collaborative updates (Algo. 1, lines 1–5).
- 3) **Agent Selection:** After the initial training, agents are evaluated based on their F1 score and accuracy, f_j and a_j , respectively. Then, each agent $Agent_j$ engages in SAC following (1)–(4) to compute the round-specific average performance metrics, μ_f and μ_a . Using these averaged metrics, each agent $Agent_j$ verifies its selection by checking:

$$\mathcal{A}_{\text{sel}} = \{Agent_j \mid f_j \geq \mu_f \text{ and } a_j \geq \mu_a\}. \quad (7)$$

Only those agents meeting or exceeding both thresholds, i.e., those in $\mathcal{A}_{\text{sel}}$, are selected to contribute to the aggregation step, while the remaining agents, $\mathcal{A}_{\text{not_sel}}$, continue in the process via transfer learning. This selection ensures that only the most reliable models contribute to the global model in a given round, thus enhancing overall quality (Algo. 2, lines 1–9).

- 4) **Aggregation and Optimization:** The selected agents $\mathcal{A}_{\text{sel}}$ participate in the SAC approach following (1)–(4) to securely aggregate their local model weights as shown in Fig. 1. This aggregation, derived from the round's top-performing agents, ensures the global model $\Theta_{\mathcal{A}_{\text{sel}}}$ is shaped by the most reliable local updates for that round. As a result, the global model $\Theta_{\mathcal{A}_{\text{sel}}}$ continually incorporates high-quality insights, even though selected agents $\mathcal{A}_{\text{sel}}$ may vary across rounds, enhancing overall performance. Consequently, each $Agent_j \in \mathcal{A}_{\text{sel}}$ obtains the aggregated global model $\Theta_{\mathcal{A}_{\text{sel}}}$ for the next round of updates. This corresponds to Algo. 2, line 12.
- 5) **Transfer Learning to Non-Selected Agents:** Once the global model $\Theta_{\mathcal{A}_{\text{sel}}}$ is created, it is transferred to the agents that were not selected in the current round, gathered in the set $\mathcal{A}_{\text{not_sel}}$, as follows:

$$\theta_j = \Theta_{\mathcal{A}_{\text{sel}}}, \quad \forall Agent_j \in \mathcal{A}_{\text{not_sel}}. \quad (8)$$

This step ensures that all agents, whether selected $\mathcal{A}_{\text{sel}}$ or non-selected $\mathcal{A}_{\text{not_sel}}$, get the updated global model

$\Theta_{\mathcal{A}_{\text{sel}}}$. Therefore, the non-selected agents $\mathcal{A}_{\text{not_sel}}$ can then participate in the next round of training with a strong benchmark. By iteratively applying this transfer learning, even agents in $\mathcal{A}_{\text{not_sel}}$ can contribute to the global model in later rounds if their performance improves (Algo. 2, line 13).

- 6) **Iterative Process until Final Round:** The previous steps are repeated iteratively across multiple rounds. At the start of each new round, agents recompute and update the feature *values* on the current window and normalize the inputs. This keeps input dimensionality consistent across agents and reduces communication overhead. After that, each agent refines its local model by training on its dataset while leveraging the updated global model from the previous round. In each round, agents are re-evaluated and re-selected based on updated F1 scores and accuracy. The continuous process of selection, aggregation, and transfer learning ensures that the final global model is always robust and incorporates the best aspects of all participating agents over time (Algo. 1, lines 8–17).

Algorithm 1 Proposed SAFE-ADVENT

Input: Number of agents N , training datasets $\mathcal{D}$, validation datasets $\mathcal{U}$, number of FL rounds T .

Initialization:

- 1: **for** $j = 1$ to N **do**
 - 2: From D_j , extract Argus flow features and compute short-window entropy-based statistics; Normalize inputs.
 - 3: Get θ_j after local training of D_j .
 - 4: Test the updated model θ_j on validation dataset U_j .
 - 5: Get f_j and a_j
 - 6: **end for**
 - 7: Update $\Theta = \{\theta_1, \dots, \theta_N\}$ using *Algorithm 2*
 - SAC-based P2P FL for subsequent FL rounds:**
 - 8: **while** $t \leq T$ **do**
 - 9: **for** $j = 1$ to N **do**
 - 10: Update Argus-derived features from D_j and entropy statistics for the current window; normalize inputs.
 - 11: Train local dataset D_j using $\Theta_{\mathcal{A}_{\text{sel}}}$ for ε episodes.
 - 12: Get the updated model θ'_j .
 - 13: Test the updated model on validation dataset U_j .
 - 14: Get f_j and a_j .
 - 15: **end for**
 - 16: Update $\Theta = \{\theta_1, \dots, \theta_N\}$ using *Algorithm 2*
 - 17: **end while**
-

- 7) **Final Deployment:** After the final round of training, each agent has the final global model locally. This model is then used for real-time inference and detection of DDoS attacks, ensuring that all agents are equipped with the most effective and up-to-date model.

Algorithm 2 Learning Process

Input: Number of agents N , *performancemetrics* : $[f_j, a_j]_{j=1,\dots,N}$
Agent selection:
1: **function** SELECT($N, [f_j, Acc_i]_{j=1,\dots,N}$)
2: Get $\mu_f \leftarrow \text{SAC}(N, [f_j]_{j=1,\dots,N})$
3: Get $\mu_a \leftarrow \text{SAC}(N, [a_j]_{j=1,\dots,N})$
4: Initialize an empty set $\mathcal{A}_{\text{sel}}$ for selected agents
5: **for** $i = 1$ to N **do**
6: Assign Agent_j to set $\mathcal{A}_{\text{sel}}$ based on (7)
7: **end for**
8: **return** $\mathcal{A}_{\text{sel}}$
9: **end function**
10: Get $\mathcal{A}_{\text{sel}_K} = \{A_{\text{sel}_j}, \dots, A_{\text{sel}_K}\} \leftarrow \mathcal{A}_{\text{sel}}$ % Set of selected agents
11: Get $\mathcal{A}_{\text{not_sel}_R} = \{A_{\text{not_sel}_j}, \dots, A_{\text{not_sel}_R}\} = \text{Agent}_N \setminus \mathcal{A}_{\text{sel}_K}$
 % Set of disregarded agents
Global model design:
12: Get $\Theta_{\mathcal{A}_{\text{sel}}} \leftarrow \text{SAC}(\mathcal{A}_{\text{sel}_K}, K)$
Transfer learning:
13: Update $\theta_j, \forall j \in \mathcal{A}_{\text{not_sel}_R}$ using (8)

B. COMMUNICATION COST AND COMPLEXITY ANALYSIS OF SAFE-ADVENT

In the conventional SAC-based P2P FL framework, each agent transmits the computed partitions and subtotals of the model parameters Θ to the other $(N-1)$ agents during each FL round [19]. Therefore, the communication cost per round is given by $2\Theta N(N-1)$. Extending this over T rounds, the total communication cost is $c^{\text{sac}} = 2\Theta N(N-1)T$.

In contrast, SAFE-ADVENT reduces the communication load through the agent selection mechanism. Specifically, in round t , only $K_t \leq N$ agents participate in SAC, leading to a training-related cost of $c_{t,1} = 2\Theta K_t(K_t - 1), \forall t = 1, \dots, T$. Moreover, the computation of the averaged performance metrics, i.e., μ_f and μ_a , requires an exchange cost of $c_{t,2} = 2QN(N-1)$, where Q denotes the size of the exchanged metrics. Finally, the transfer learning step requires a single broadcast of the global model to the non-selected agents, which adds a cost of $c_{t,3} = \Theta$. The total communication cost of a single round in the proposed scheme is given by $c_{t,1} + c_{t,2} + c_{t,3}$. Hence, the overall cost across all rounds can be expressed as:

$$\begin{aligned} c^{\text{safe-advent}} &= \sum_{t=1}^T (c_{t,1} + c_{t,2} + c_{t,3}) \\ &= \sum_{t=1}^T (2\Theta K_t(K_t - 1) + 2QN(N-1) + \Theta). \end{aligned} \quad (9)$$

For centralized FL [41], the communication cost is given by $c^{\text{centralized-fl}} = \Theta(N+1)T$, as it involves N uploads from the agents to the central server and a single broadcast from the server to all agents for the global model update.

Finally, for the traditional centralized DL benchmark [9], the communication paradigm fundamentally differs. Instead of securely exchanging fixed-size model parameters Θ , centralized DL requires the continuous transmission of raw data to the server for training, followed by the server broadcasting the updated model back to the agents. Let $|D_j^{(t)}|$ denote the number of raw data samples collected by Agent_j during the observation window of round t , and let S_{data} represent the average size of a single data sample in bytes. The communication cost per round for the centralized DL benchmark consists of the massive raw data upload plus the model broadcast. Consequently, the total communication cost over T observation rounds is expressed as:

$$c^{\text{centralized-dl}} = \sum_{t=1}^T \left(\sum_{j=1}^N |D_j^{(t)}| S_{\text{data}} + \Theta \right). \quad (10)$$

Given that the cumulative volume of raw network traffic vastly exceeds the size of the compiled model parameters (i.e., $\sum_{j=1}^N |D_j^{(t)}| S_{\text{data}} \gg \Theta$), the centralized DL benchmark inherently saturates the available network bandwidth, which makes it highly inefficient and unscalable for continuous, real-time anomaly detection in resource-constrained 5G-V2X environments.

Table 3 summarizes the communication complexity of SAFE-ADVENT and all benchmarks. As given in (9), the per-round cost comprises three terms: SAC aggregation among K_t selected agents, metric exchange across all N agents, and a transfer-learning broadcast to non-selected agents. Since $Q \ll \Theta$, the metric-exchange term is negligible, and the dominant complexity reduces to $\mathcal{O}(TK_t^2)$, always lower than $\mathcal{O}(TN^2)$ of conventional SAC-based P2P FL when $K_t < N$. Centralized FL achieves $\mathcal{O}(TN)$ but introduces a single point of failure and removes serverless privacy preservation. The DL benchmark shares the same asymptotic class but incurs a significantly higher practical cost since $S_{\text{data}} \gg \Theta$, as confirmed by Fig. 3. The computational complexity of SAFE-ADVENT follows the same order as its communication complexity per round.

V. IMPLEMENTATION OF SAFE-ADVENT

This section describes the implementation of SAFE-ADVENT, which relies on a dataset generated through our 5G-V2X testbed, and made publicly available in [42]. Subsection A presents the main features inherent to our 5G-V2X testbed, while Subsection B describes in detail the dataset used to train the SAFE-ADVENT models. Finally, Subsection C presents SAFE-ADVENT's model architecture and hyperparameters selection procedure.

A. 5G-V2X TESTBED OVERVIEW

In Fig. 1, we show the main hardware and software modules that compose our 5G-V2X testbed. We also show an example scenario composed by four agents, with Agent 4 as a malicious UE. Here, all agents exchange model

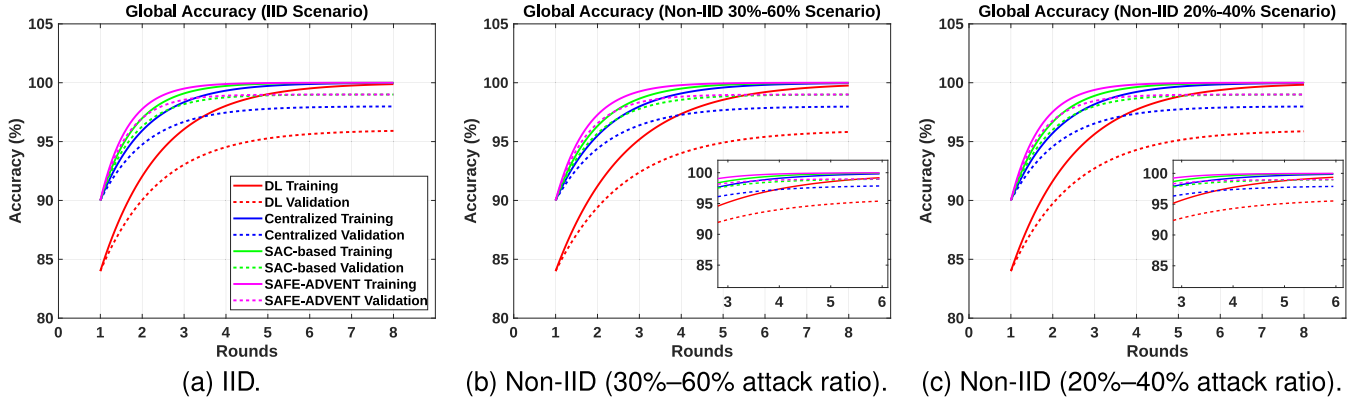


Fig. 2. Global training and validation accuracy over 8 rounds for SAFE-ADVENT and all benchmarks under three data distribution scenarios.

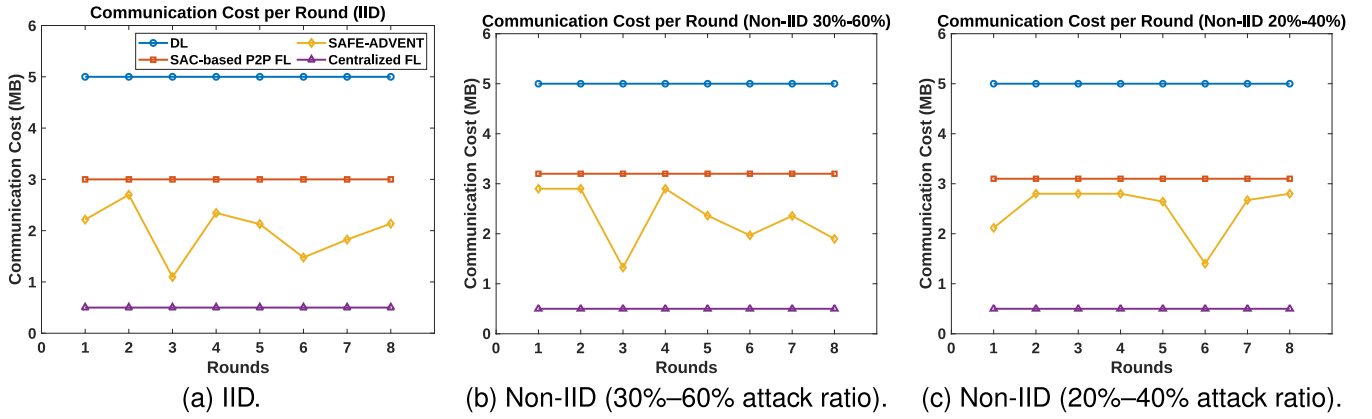


Fig. 3. Communication cost (in megabytes–MB) over 8 rounds for SAFE-ADVENT and all benchmarks under three data distribution scenarios.

updates through SAC-based secure aggregation, as detailed in Subsection III-A.

We configured a 5G SA operator composed of one gNodeB (which in turn represents the RAN), as well as a 5G CN that includes multiple function modules, such as the session management function (SMF), the access and mobility management function (AMF), and the user plane function (UPF). The UPF provides Internet access to all UEs, also uniting them in a single data network (DN). The software used to deploy all RAN/CN modules is OAI, which is open-source and developed by Eurecom to support 4G and 5G NR systems [43].

Two units of Dell Precision 5820 Tower (Intel Xeon W-2265 3.50GHz with 12 cores, 128GB of RAM) are used to deploy the RAN and the CN elements separately. The RAN host also communicates with a USRP X310 card, used to deploy the gNodeB and emulate the Uu interface for communication with all UEs. The UEs, in turn, are deployed with three JetRacer Pro robot cars, each one integrated with a NVIDIA Jetson Nano platform [44], and an emulated UE represented by a second USRP X310 card that is operated through the same machine where the 5G CN is deployed.

The robot cars also include a SIMCom SIM8202G-M2 5G card [45] for 5G communication, while the emulated UE uses OAI. In Fig. 1, we present a scenario where one of the physical agents (i.e., Agent 4) is a malicious UE, while the others are benign.

It is worth highlighting that the number of cars chosen for our testbed aims to mitigate the local RF interference and to align UE traffic generation with the processing and network capacity inherent to an indoor USRP-based setup. Even though a four-UE deployment is enough for an end-to-end validation of the hardware, the SAC protocol execution, and all benchmark non-IID data distributions, validation with larger deployments requires extending beyond these physical hardware constraints.

In our testbed, V2V communication is enabled through an emulated PC5-based *sidelink* interface, mirroring its intended properties as defined in 3GPP Release 16 [46]. To do so, all UEs also include a Wi-Fi ad-hoc adapter to establish direct, infrastructure-free links with neighboring UEs.

The setup above allows UEs to autonomously initiate communication without involving the gNodeB or the CN.

The Wi-Fi ad-hoc interface is used solely for peer discovery, packet exchange and decentralized model updates among vehicles, hereby effectively mimicking the low-latency, localized nature of PC5-based *sidelink* communication.

In parallel with V2V communication, each UE remains connected to the 5G network via the *Uu* interface as mentioned above, thus allowing the aggregation of data used in deep learning (DL)-based approaches, as well as central aggregation inherent to centralized FL approaches. Such a dual connectivity setup allows vehicles to operate in hybrid mode, relying on the 5G CN for global orchestration or fallback connectivity, and data aggregation for DL while performing core FL activities using the emulated *sidelink* interface, such as peer coordination, SAC mechanisms (i.e., secret partitioning and secure multi-party average calculation), transfer learning, and the exchange of information concerning the detected attacker (e.g., its IP address).

B. 5G-V2X TESTBED DATASET SETUP

In our experiments, benign UEs generate TCP and UDP traffic at a constant throughput of 5 MBps, using *iperf3*. On the other side, malicious traffic is generated by malicious UEs through *hping3* and *Mausezahl*. For this type of traffic, TCP packets with the SYN flag set and/or UDP packets are flooded in the 5G network, depending on the scenario, and no payload is set for any of them.

It is worth noting that end-to-end cryptography was intentionally disabled in our experiments. The considered threat model focuses on volumetric DDoS (flooding) attacks targeting *availability*, not confidentiality or integrity. While encryption and message authentication ensure data protection, they cannot prevent a malicious transmitter from sending large volumes of valid packets. In a 5G-V2X environment, such floods saturate radio and core network resources, thus delaying legitimate encrypted traffic. Accordingly, our model relies on flow-level statistics observable without payload access, which remain informative under encryption.

In all our experiments, as also seen in Fig. 1, we consider use cases where one or multiple cars are attackers, and the main targets are neighboring cars in the 5G-V2X network (i.e., using the V2V communication interface), hence threatening all network services, applications and P2P-based approaches that rely on the direct communication between them.

We defined multiple scenarios to create our 5G-V2X testbed dataset, mainly by observing two key attributes: 1) The number of malicious and benign UEs, and 2) the type of malicious traffic (i.e., TCP, UDP or mixed). Table 4 summarizes all parameters set to collect data for our 5G-V2X testbed dataset.

We start setting our data collection scenarios based on all possible combinations of parameters defined in the table. For each scenario, each UE (benign or malicious) performs a network packet capture (PCAP) through *Wireshark*, thus generating PCAP files per scenario.

Furthermore, in order to reflect realistic conditions in 5G-V2X networks, for each parameter combination from Table 4, we finally collect data under two distinct scenarios:

- **IID:** In this scenario, each connected agent (i.e., each robot car and the emulated UE) collects approximately 10,000 packets over a fixed duration of 100 seconds. The packet distribution is balanced, with 50% of all packets corresponding to benign traffic and the remaining 50% corresponding to malicious traffic, this last one generated by each malicious agent at a rate of 1,000 packets per second (cf. Table 4), thus ensuring uniform attack intensity across all malicious agents.
- **Non-IID:** In this scenario, we simulate the heterogeneity inherent to real-world vehicular environments and evaluate how variations in local data distributions affect SAFE-ADVENT's detection accuracy and robustness. While the overall dataset size per agent remains the same as for the IID scenario (i.e., around 10,000 packets in a 100-second data collection round), the attack traffic ratio varies across agents. Specifically, we defined two distinct data collection sets. In the first one, the attack ratios are fixed in 30%, 40%, 50%, and 60% for Car 1, Car 2, Car 3, and Car 4, respectively. On the other hand, in the second set, a milder imbalance is introduced with attack ratios of 20%, 25%, 30%, and 40% for Car 1, Car 2, Car 3, and Car 4, respectively. Both sets introduce significant variability in attack intensity, simulating situations where different vehicles experience different levels of adversarial traffic. For example, one vehicle might capture only 30% of all attack packets generated, while another may record up to 60%. Such variability better emulates the dynamic nature of network attacks in different regions or under different network conditions.

C. SAFE-ADVENT'S MODEL ARCHITECTURE AND HYPERPARAMETERS SELECTION

The model architecture and hyperparameters inherent to SAFE-ADVENT were selected dynamically to optimize performance on the 5G-V2X testbed. The model was designed to handle the specific challenges inherent to a DDoS attack detection procedure, leveraging features generated using the Argus tool [47].

In this step, we keep only flow-level features that are useful for flooding-style DDoS detection in 5G-V2X and that can be computed without payload access. It is worth noting that feature selection and engineering are critical, as Argus exports many redundant and/or irrelevant fields. Selecting a compact set of features reduces model size, inference latency, and communication overhead, while ensuring applicability across 5G-V2X scenarios.

From Argus, we keep the following base features: (i) packet count; (ii) byte count; (iii) flow duration; (iv) protocol and TCP flag counters; and (v) timing summaries. On top of these, we derive short-window features using a *sliding window* of $W = 2$ seconds with *stride* $s = 0.5$ seconds,

including: (i) packet and byte rates; (ii) average packet size and its short-window variance; (iii) mean and variance of inter-arrival times; (iv) TCP flag ratios (e.g., SYN/total, ACK/total) and protocol mix ratios; and (v) entropy terms over binned packet sizes and destination ports. Entropy is computed as:

$$H(\mathcal{X}) = - \sum_{x \in \mathcal{X}} p(x) \log p(x),$$

where $\mathcal{X}$ represents a histogram over packet sizes with $B_{\text{size}} = 32$ bins spanning $[0, 1500]$ Bytes, and over destination ports with $B_{\text{port}} = 16$ bins (port bucket index $\lfloor \text{port}/4096 \rfloor$). These statistics are observable under encryption and remain stable under mobility [48].

On the training split, we apply the following preprocessing steps: (i) discard low-variance features, (ii) remove highly correlated pairs (e.g., $|\rho| > 0.9$), (iii) rank the remaining features by label relevance (e.g., using mutual information) and (iv) retain the top set compatible with the UE latency/memory constraints. Inputs are then z-score normalized using training statistics. This approach works without payload access, tolerates dynamic peers, and is lightweight, maintaining effective detection under non-IID traffic while controlling on-device and on-air cost.

Following the feature selection and preprocessing phase, the precise input to the neural networks of SAFE-ADVENT and all evaluated benchmarks (i.e., Centralized FL, SAC-based P2P FL and Centralized DL) is a continuous, z-score normalized feature vector $\mathbf{x} \in \mathbb{R}^{42}$. Specifically, this 42-dimensional input space is constructed from five distinct feature categories to capture both volumetric and targeted evasion attacks:

- 1) *Basic Flow Statistics* (five features), including total packet counts, byte counts, and flow duration;
- 2) *Protocol and Flag Counters* (12 features), capturing specific TCP flag occurrences (e.g., SYN, ACK, RST, FIN) and UDP/ICMP flow indicators;
- 3) *Timing and Rate Summaries* (15 features), encompassing packet and byte transmission rates, the mean and short-window variance of packet sizes, and inter-arrival time statistics;
- 4) *Ratio Metrics* (eight features), such as TCP flag ratios (e.g., SYN/Total) and protocol mix proportions;
- 5) *Entropy Metrics* (two features), specifically the short-window entropy computed over packet size and destination port histograms.

This comprehensive vector $\mathbf{x}$ is fed directly into the model's input layer. To ensure a fair comparative analysis, SAFE-ADVENT and all benchmarks share the identical core neural network architecture. This architecture comprises two dense hidden layers with 30 and 10 neurons respectively, utilizing the *ReLU* activation function, and concludes with a two-neuron output layer employing a *softmax* function for binary classification (DDoS attack or no attack). All models apply the *ADAM* optimizer and *L1* regularization.

As detailed in Subsection B, each agent captures approximately 10,000 local network packets. At each agent, this local dataset is strictly partitioned into an 80%–20% training/validation split. The validation set is used locally to compute the F1-score and accuracy for the dynamic agent selection process. Because this is a decentralized 5G-V2X deployment without a central testing repository, the global model's offline generalization is evaluated by aggregating the performances across these local validation sets. Local models are trained using TensorFlow 2.x and deployed via *LiteRT* to accommodate the resource constraints of the UEs.

Finally, to ensure rigorous reproducibility, we consolidate our complete hardware, software, and ML experimental settings in Table 5. Specifically for the FL-based approaches (SAFE-ADVENT, SAC-based P2P FL and Centralized FL), the learning process spans $T = 8$ communication rounds, with 10 local training episodes per round, a batch size of 100, and a learning rate of 10^{-4} . To maintain computational equivalence with the FL models, which process their data 80 times (eight rounds $\times$ 10 local episodes), the Centralized DL benchmark is trained over 80 global epochs across the entire aggregated dataset.

VI. EXPERIMENTAL RESULTS

This section focuses on SAFE-ADVENT's performance in terms of DDoS attacks detection in the network traffic. To gain comprehensive insights, experiments were conducted under offline and online (live traffic) scenarios, also following an established approach as described in [22]. First, Subsection A describes the performance metrics and benchmarks adopted to assess the effectiveness of our proposal, while subsections B and C present the obtained results.

A. PERFORMANCE METRICS AND BENCHMARKS DESCRIPTION

To evaluate the performances of SAFE-ADVENT, we adopt the following metrics: 1) *Accuracy*: The ratio of correctly predicted instances (both true positives and true negatives) to the total instances (i.e., the sum of true/false positives/negatives); 2) *Precision*: The ratio of true positive predictions to the total predicted positives (i.e., the sum of true positives/negatives); 3) *Recall*: The ratio of true positive predictions to the total actual occurrences (i.e., the sum of true positives and false negatives); 4) *F1-Score*: The harmonic mean of precision and recall; and 5) *Communication Cost*: The total amount of data transferred during training, measured in megabytes (MB).

SAFE-ADVENT is compared to three benchmarks: 1) A traditional DL approach, where model training is performed in a centralized manner by aggregating raw data from all agents [9], 2) a centralized FL approach, where agents train their models locally and only exchange model updates with a central server [49], and 3) a conventional SAC-based P2P FL approach, where agent selection and transfer learning are not done, and all agents are included in each round in SAC [16]. In our 5G-V2X testbed, DL and centralized FL

approaches rely on a dedicated central server hosted within the 5G RAN, which performs data aggregation and model orchestration.

It is worth highlighting that this work extends our prior study [21] by validating the framework in a real 5G-V2X testbed. To keep the evaluation controlled and directly comparable, we retain the same benchmarks used in [21]. Specifically, we adopt *centralized FL* as a representative *server-based FL* setting. We also adopt the conventional *SAC-based P2P FL* benchmark as the closest *serverless* collaborative training framework to our proposal, since it relies on secure aggregation among peers but does not incorporate SAFE-ADVENT's dynamic agent selection or transfer learning [16]. We additionally include a centralized DL benchmark as a reference point when raw data aggregation is feasible, to quantify the performance/overhead trade-off against collaborative learning [9]. Recent studies further report that federated and centralized learning can achieve comparable accuracy under many settings, while differing in privacy, resource, and communication trade-offs [50], [51], [52].

B. RESULTS OF OFFLINE EXPERIMENTS

In this subsection, we present an in-depth analysis of the overall performance of SAFE-ADVENT and all benchmarks in an offline training environment, using the dataset collected from our 5G-V2X testbed.

Fig. 2 shows the global accuracy obtained for training and validation datasets over eight rounds, under three data distribution scenarios, namely IID (Fig. 2a), non-IID with attack ratios ranging from 30 to 60% (Fig. 2b), and non-IID with attack ratios ranging from 20 to 40% (Fig. 2c).

As can be seen in Fig. 3, in the IID scenario, all FL-based approaches start on the first round with an accuracy of ~90% for both training and validation datasets, while the DL benchmark begins at ~84%. The DL benchmark converges slowly and reaches 100% (resp. 96%) training accuracy (resp. validation accuracy) at the 8th round. In contrast, centralized FL converges faster, achieving 100% and 98% in training and validation accuracies, respectively, around the 6th round. Meanwhile, the SAC-based P2P FL benchmark converges to 100% training accuracy and 99% validation accuracy during the 5th round. SAFE-ADVENT is the fastest converging approach, reaching 100% in training accuracy and ~100% validation accuracy at the 4th round. In the non-IID (30%–60% attack ratio) scenario, a higher data heterogeneity postpones the convergence for all benchmarks, compared to the IID scenario. Nevertheless, SAFE-ADVENT still has the fastest convergence at the 5th round, while the reduced gap between training and validation accuracy curves indicates a good generalization and low overfitting risk. Finally, in the non-IID (20%–40% attack ratio) scenario, similar observations to the previous case can be made, where SAFE-ADVENT is still the fastest approach to converge to the best accuracy performance in both training and validation.

Fig. 3 presents the communication cost for SAFE-ADVENT and all benchmarks under the defined data distribution scenarios. In the IID scenario (Fig. 3a), DL benchmark consistently communicated 5 MB of data per round since it aggregates raw data and uses the full available throughput. In contrast, the SAC-based P2P FL benchmark remains static at 3 MB of exchanged data in all rounds, hence reflecting a fixed communication cost that is equivalent to 60% of total throughput. Meanwhile, the centralized FL operates constantly at 0.5 MB per round, thus using only 10% of the available throughput (without accounting for any data securing overhead). The fixed communication costs of the benchmarks result from the fact that all agents participate in every training round. In contrast, SAFE-ADVENT's communication costs fluctuate due to its dynamic agent selection and transfer learning process. Its communication cost varies by up to 20% around an average of 2 MB per round, and is maintained at worst 0.3 MB lower than that of the SAC-based P2P FL benchmark, resulting in an average decrease of 33.3%. Hence, SAFE-ADVENT achieves significant communication cost savings compared to both DL and SAC-based P2P FL benchmarks, while being more accurate and resilient to the single point of failure, compared to centralized FL. In the non-IID (30%–60% attack ratio) scenario (Fig. 3b), the DL and centralized FL benchmarks maintain the same communication costs, while the one for SAC-based P2P FL slightly increases to 3.2 MB due to the needed extra overhead to manage non-IIDness. However, SAFE-ADVENT fluctuates without exceeding the cost of SAC-based P2P FL. In the non-IID (20%–40% attack ratio) scenario (Fig. 3c), similar observations are made, where our proposed method's communication cost is still between those of centralized FL and SAC-based P2P FL.

It is worth noting that, in order to ensure a rigorous and computationally fair comparison in figs. 2 and 3, the x-axis metric of a "Round" for the centralized DL benchmark is defined as a synchronized time window of data collection and model updating. In a continuous 5G-V2X deployment, agents continuously generate data. During each round, the DL benchmark requires all agents to upload their newly collected raw data to the central server, which then updates the global model and broadcasts it back. To maintain strict computational equivalence with the FL models which execute 10 local training episodes per communication round, a single round for the centralized DL model is mapped to 10 global training epochs over the newly aggregated dataset.

Notably, the FL-based approaches demonstrate a more rapid initial convergence trajectory compared to the centralized DL benchmark. From an optimization perspective, this phenomenon occurs because the centralized dataset contains a massive, highly heterogeneous mixture of traffic flows, which induces a highly non-convex global loss landscape with significant gradient noise during early epochs. In contrast, the FL agents initially optimize their models over smaller, localized data distributions. By executing multiple local gradient steps ($E = 10$) on these less noisy subsets, the

local models quickly descend into stable local minima. The subsequent SAC acts as an ensemble averaging mechanism, smoothing the aggregated weight updates and propelling the global model to a high-accuracy state faster than the centralized stochastic gradient descent (SGD) processing the raw global dataset.

C. RESULTS OF ONLINE LIVE EXPERIMENTS

1) TESTING ROUND

In this subsection, we assess SAFE-ADVENT's performance by deploying it on our 5G-V2X testbed and running a live traffic test round with one attacking car. Specifically, in the live traffic test round, all four agents (Car 1, Car 2, Car 3, and Emulated UE) process incoming live traffic with their local FL models. The live traffic was generated by simulating benign traffic using `iperf3` at a fixed throughput of 5 Mbps to emulate typical vehicular communication patterns. Concurrently, malicious traffic was produced by one of the agents (in our experiments, we selected Car 3 as the attacker) employing two instances of `hping3` and three instances of `Mausezahn` to flood the network with TCP SYN and UDP packets, hence replicating realistic DDoS attack scenarios. This experimental setup ensures that all vehicles will receive a mixture of benign and attack traffic, consistent with the conditions defined during our dataset collection phase.

To facilitate supervised local updates and accurately compute the performance metrics (e.g., F1-score and accuracy) during these live traffic experiments, ground-truth labels must be provided in real-time. It is important to note that while RF transmissions are physically live, the testing environment is strictly orchestrated. As is standard practice in vehicular network security evaluations [22], where roles are pre-assigned to network entities, our testbed orchestrator deterministically assigns specific IP addresses and ports to the attack generation scripts and benign applications.

Consequently, a localized programmatic labeling module running alongside the Argus extractor on each UE can automatically classify the traffic. Specifically, this module matches the incoming flow 5-tuples (Source IP, Destination IP, Source Port, Destination Port, and Protocol) against the active orchestrated attack profiles. Flows originating from the designated attacker IPs are deterministically tagged as malicious (1), while `iperf3` flows are tagged as benign (0). These ground-truth labels are dynamically appended to the 42-dimensional feature vector $\mathbf{x}$ before it is fed into the local neural network, thereby enabling the real-time computation of the loss function and the execution of supervised local model updates without requiring manual intervention. While this orchestrated deterministic labeling is necessary for establishing rigorous ground-truth within the experimental testbed, a real-world 5G-V2X production deployment would instead align with O-RAN standards [53], utilizing an offline-trained model to infer and provide pseudo-labels for continuous online AI/ML training.

Each agent updates its model based on real-time data, and these updates are securely aggregated to produce a global prediction. The results for each agent are shown in Table 6.

The average accuracy for the testing round is 95.04%, while the average F1-score is 95.08%. Once the FL models run on each agent and the aggregated performance indicates high accuracy under live traffic conditions, an attack is flagged. Following this detection, we further investigate the source of the anomaly through registered users' data provided by the 5G-V2X network. This additional monitoring confirms the attacker's identity by correlating the registered IP address with the device's international mobile subscriber identity (IMSI). This robust identification method ensures that even if the attacker spoofs its IP address, the underlying IMSI remains unchanged and is used for reliable attribution. In our experiment, Car 3 is correctly identified as the malicious agent with IP address 10.1.1.2. As a result, Car 3 would be excluded from the secure aggregation process in subsequent rounds. If the attacker was not detected in this initial live testing round, additional rounds would be conducted to further refine the model. However, since the attacker was identified with high accuracy in the first round of our experiment, then no additional FL rounds were required.

2) IMPACT OF DATA DISTRIBUTIONS

To further assess the robustness of our SAFE-ADVENT models, we compared its performance to all benchmarks using live traffic under three data distribution scenarios: 1) IID, 2) non-IID with attack ratios ranging from 30 to 60%, and 3) non-IID with attack ratios ranging from 20 to 40%. Here, we evaluate the detection rate for DDoS attacks and classification accuracy for benign traffic, and identify the IP address of the attacker. For each data distribution condition, we consider different scenarios that include pure traffic types, i.e., "Benign Only," "Attack TCP Only," "Attack UDP Only" and "Attack TCP & UDP", as well as mixed traffic types, i.e., "Benign & Attack TCP," "Benign & Attack UDP," and "Benign & Attack TCP & UDP". Finally, the attacker's IP address was systematically spoofed, and respectively assigned from 10.1.1.1 to 10.1.1.6 to the scenarios involving DDoS attacks. Concurrently, the benign agents used IP addresses 10.1.1.7, 10.1.1.8, and 10.1.1.9 to diversify the profile of legitimate traffic. This setup ensures that detection models do not overfit to a static IP range, which reflects real-world conditions.

Table 7 shows the obtained results for the IID scenarios. In scenarios "Benign Only", "Attack TCP Only", "Attack UDP Only", and "Attack TCP & UDP", all approaches detect and classify all respective flows with near-100% detection rates, indicating that single-protocol attacks or exclusively benign flows are simple to distinguish. Nevertheless, note that DL benchmark misclassified a legitimate UE as an attacker when mixing two types of malicious traffic.

In the "Benign & Attack TCP" scenario, centralized FL achieved an 87.24% DDoS detection rate without any

misclassification. The SAC-based P2P FL benchmark reached a DDoS detection rate of 86.37%, but misclassified IP 10.1.1.7 as an attacker. DL approach had the lowest DDoS detection rate (84.79%) with two UE misclassifications, suggesting that mixed traffic leads to an underperforming model. In contrast, SAFE-ADVENT reported the highest DDoS detection rate (93.06%) without any misclassification. Similar trends are also observed for the “Benign & Attack UDP” and “Benign & Attack TCP & UDP” scenarios, with DL benchmark realizing the lowest detection rate. Both centralized FL and SAC-based P2P FL maintained high detection rates. However, except from centralized FL, the other benchmarks misclassified legitimate UEs as attackers. In the latter scenario, SAFE-ADVENT outperforms all benchmarks by at least 4.2% and up to 7.1% in terms of DDoS detection rate.

Table 8 presents the results for the non-IID scenarios with attack ratios between 30% and 60%. Similarly to the IID scenarios, near-100% detection rates and no misclassifications are achieved by all solutions when only benign or only attack traffic is flowing in the network. In the “Benign & Attack TCP” scenario, centralized FL attained a DDoS detection rate of 89.45% and no misclassifications. Meanwhile, SAC-based P2P FL achieved DDoS detection rate of 90.32%, although it misclassified IP 10.1.1.7 as an attacker. In contrast, DL benchmark’s DDoS detection rate dropped to 80.24% and misclassified two IPs (10.1.1.8 and 10.1.1.9), thus indicating a significant difficulty in managing non-IIDness. Meanwhile, SAFE-ADVENT remained the best approach with a DDoS detection rate of 94.58% and no misclassifications. A similar trend is noticed in the “Benign & Attack UDP” scenario, with the DL approach degrading further in performances, while SAFE-ADVENT maintaining a DDoS detection rate of 95.16%. In the meantime, SAC-based P2P FL and centralized FL kept similar performances as to the previous scenario. Finally, in the “Benign & Attack TCP & UDP” scenario, centralized FL detected 90.02% of DDoS traffic but misclassified one UE (10.1.1.8), while SAC-based P2P FL realized a DDoS detection rate of 91.64% without any misclassifications. The DL benchmark is further degraded to only 71.53% DDoS detection rate and misclassified three UEs as attackers (10.1.1.7, 10.1.1.8 and 10.1.1.9). SAFE-ADVENT remained the best approach, with a DDoS detection rate of 94.92% and no misclassifications.

As shown in Table 9, all methods are able to handle pure benign/attack traffic in the non-IID (20–40% attack rate) scenarios, demonstrating their effectiveness in simple network conditions. However, in the “Benign & Attack TCP” scenario, both centralized FL and SAC-based P2P FL identified only 1.14% and 5.02% of all DDoS flows, respectively, and failed to detect the attacker. At the same time, the DL approach performed worse, failing to detect any attacker or DDoS attack flows. In contrast, SAFE-ADVENT was able to detect 89.91% of DDoS traffic, and

accurately pinpointed IP 10.1.1.4 as the attacker. This result highlights how agent selection and transfer learning enabled SAFE-ADVENT to adapt successfully to intense data non-IIDness. In the “Benign & Attack UDP” scenario, the benchmarks were able to achieve above 80% DDoS detection rates, highlighting a higher ease in detecting DDoS UDP-based attacks than TCP-based ones. Preference is still for SAC-based P2P FL compared to centralized FL and DL benchmarks. In any case, SAFE-ADVENT is the best performing method with a DDoS detection rate of 90.15% and no attacker misclassifications. In the “Benign & Attack TCP & UDP” scenario, all benchmarks realize DDoS detection rates below 34% and fail to identify any attacker correctly. This is mainly due to their inability to accurately detect TCP attacks. Conversely, SAFE-ADVENT detected 95.35% of DDoS traffic and accurately identified the attacker with IP 10.1.1.6.

3) IMPACT OF ADVERSARIAL ATTACKS

Considering the most challenging non-IID scenario with attack rates ranging from 20% to 40%, we present in Table 10 SAFE-ADVENT’s performance under network-level adversarial attacks in comparison with all benchmarks.

As shown, all benchmarks achieved a poor performance, with the DL method as the most vulnerable one, achieving an accuracy of 15.82% and a F1-score of 15.60%, which reflects its high sensitivity to corrupted data samples. Centralized FL is slightly better with performances ranging between 24% and 25%, while SAC-based P2P FL realizes higher performances in the range of 28% and 29.4%. This result emphasizes a better resilience of this benchmark, however, it is still significantly affected by the contaminated local updates, preventing the global model from converging. In contrast, SAFE-ADVENT preserves its performances around 90%, thus outperforming all benchmarks. In comparison with its average accuracy and F1-score obtained with live traffic and non-adversarial attacks (i.e., Table 6), we note a slight decrease in SAFE-ADVENT’s performance (around 5%) when adversarial attacks are involved. The obtained results demonstrate that SAFE-ADVENT’s key agent selection and transfer learning mechanisms ensure that only relevant and lowly-corrupted model updates from high-performing agents are involved in SAC P2P FL, thus effectively mitigating the impact of adversarial noise.

not only aims to mitigate the local RF interference, but also aims to align UE traffic generation with the processing and network capacity inherent to an indoor, USRP-based setup.

4) IMPACT OF SCALABILITY

Table 11 shows the robustness of SAFE-ADVENT and all benchmarks under an increased network scale. To rigorously evaluate the algorithmic robustness and the communication efficiency as the network density increases, we expanded our 5G-V2X testbed with four extra virtual agents, deployed as containers, and tested such a deployment under the

TABLE 11. Impact of scalability on performance.

Metric	Central. FL	SAC-based P2P FL	DL	SAFE-ADVENT
Accuracy (%)	89.45	88.97	81.23	95.83
Precision (%)	89.28	88.82	80.95	95.72
Recall (%)	89.60	89.10	81.38	96.01
F1-Score (%)	89.53	89.04	81.31	95.97

Non-IID (20–40% attack rate) scenario. By doing so, we effectively double the number of connected agents from four to eight, thus bypassing the limitations inherent to the primary testbed, while keeping a mixed traffic profile (i.e., combining benign and TCP and UDP attack traffic).

As illustrated, SAFE-ADVENT stands out as the best approach, achieving performances between 95% and 96%. In contrast, centralized FL and SAC-based P2P FL benchmarks showed moderate performances around 89%, while the DL approach performed the worst with accuracy and F1-score being approximately 81.3%. These results indicate the robustness of SAFE-ADVENT in handling a higher number of agents compared to the benchmarks, and in the worst data distribution scenario.

D. DESIGN RECOMMENDATIONS FOR 5G-V2X DEPLOYMENTS

Based on our empirical evaluations using our testbed, we distill the following design guidelines for deploying FL-based anomaly detection in practical 5G-V2X networks:

- **Implement Threshold-Based Client Isolation:** Our adversarial experiments (Table 10) demonstrate that standard aggregation protocols fail under data poisoning conditions. Network operators should never aggregate all available local models. Instead, systems must implement strict, dynamic local validation thresholds (e.g., rejecting models performing below the current round's average F1-score) to naturally filter out compromised updates and mitigate the impact of severe non-IID data distributions.
- **Prioritize Payload-Agnostic Feature Engineering:** Since 5G-V2X communications rely heavily on end-to-end encryption, deep packet inspection is largely unfeasible. Our results confirm that highly accurate detection (> 94%) can be achieved purely through flow-level statistics. Detection systems should prioritize lightweight, payload-agnostic features, specifically short-window packet/byte rates and entropy metrics derived from packet sizes and destination ports which remain robust indicators of volumetric DDoS attacks even when payloads are encrypted.
- **Adopt Hybrid PC5/Uu Communication Architectures:** To resolve the $\mathcal{O}(N^2)$ communication bottleneck inherent to secure P2P FL, deployments should utilize hybrid communication strategies. Heavy, peer-to-peer model exchanges should be aggressively offloaded to the PC5 sidelink (e.g., C-V2X or Wi-Fi ad-hoc) to preserve cellular bandwidth. However, the system must

maintain the 5G Uu interface as a reliable fallback for global model broadcasts or critical alerts if *sidelink* packet reception rates drop below acceptable thresholds due to high vehicular mobility.

VII. CONCLUSION

In this paper, we proposed SAFE-ADVENT, a novel SAC-based P2P FL approach adapted for use within the 5G-V2X network to detect DDoS attacks. We aimed to reduce the communication cost and enhance the robustness against data non-IIDness and adversarial attacks by leveraging agent selection and transfer learning. First, we trained and tested our proposed approach under offline scenarios, using IID and non-IID data collected from a realistic 5G-V2X testbed. Results showed that SAFE-ADVENT converged faster than DL, centralized FL and SAC-based P2P FL benchmarks. It also proved to be adaptable to different network conditions, thus reducing the communication cost over time. Through live traffic testing under IID and non-IID conditions, we demonstrated that SAFE-ADVENT can detect more than 90% of all DDoS attacks in most scenarios, thus outperforming all benchmarks. Under adversarial attacks, our method achieved the best accuracy and F1-score performances (above 90%), while all benchmarks realized degraded performances below 30%. Finally, doubling the number of FL agents did not prevent SAFE-ADVENT from preserving performances above 95%, thus proving its high scalability, compared to the benchmarks' performance remaining under 90%.

Synthesizing our empirical findings, we strongly recommend that future 5G-V2X anomaly detection architectures implement strict threshold-based client isolation to withstand poisoning attacks, leverage payload-agnostic flow heuristics (e.g., short-window entropy) to inspect encrypted traffic, and utilize hybrid PC5/Uu orchestration to offload peer-to-peer exchanges from the CN. Future work will focus on refining SAFE-ADVENT's resilience and adaptability under extreme and unpredictable network loads, managing the heterogeneity of computational resources and communication bandwidths, ensuring attacks detection performance under extreme and unpredictable network loads, and optimizing real-time inference to meet the strict latency requirements of real-world, heterogeneous 5G-V2X applications.

ACKNOWLEDGMENT

(Badre Bousalem and Fahdah Alalyan contributed equally to this work.)

REFERENCES

- [1] K. Ramezanpour and J. Jagannath, "Intelligent zero trust architecture for 5G/6G networks: Principles, challenges, and the role of machine learning in the context of O-RAN," *Comput. Netw.*, vol. 217, Nov. 2022, Art. no. 109358.
- [2] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 957–975, 2020.

- [3] S. K. Singh, R. Singh, and B. Kumbhani, "The evolution of radio access network towards open-RAN: Challenges and opportunities," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2020, pp. 1–6.
- [4] U. Ghafoor, M. Ali, H. Z. Khan, A. M. Siddiqui, and M. Naeem, "NOMA and future 5G & B5G wireless networks: A paradigm," *J. Netw. Comput. Appl.*, vol. 204, Aug. 2022, Art. no. 103413.
- [5] A. Afaq, N. Haider, M. Z. Baig, K. S. Khan, M. Imran, and I. Razzak, "Machine learning for 5G security: Architecture, recent advances, and challenges," *Ad Hoc Netw.*, vol. 123, Dec. 2021, Art. no. 102667.
- [6] A. Alalewi, I. Dayoub, and S. Cherkaoui, "On 5G-V2X use cases and enabling technologies: A comprehensive survey," *IEEE Access*, vol. 9, pp. 107710–107737, 2021.
- [7] A. Boualouache and T. Engel, "A survey on machine learning-based misbehavior detection systems for 5G and beyond vehicular networks," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 1128–1172, 2nd Quart., 2023.
- [8] B. Bousalem, V. F. Silva, R. Langar, and S. Cherrier, "Deep learning-based approach for DDoS attacks detection and mitigation in 5G and beyond mobile networks," in *Proc. IEEE 8th Int. Conf. Netw. Softwarization (NetSoft)*, Jun. 2022, pp. 228–230.
- [9] B. Bousalem, V. F. Silva, R. Langar, and S. Cherrier, "DDoS attacks detection and mitigation in 5G and beyond networks: A deep learning-based approach," in *Proc. GLOBECOM - IEEE Global Commun. Conf.*, Dec. 2022, pp. 1259–1264.
- [10] B. Bousalem, M. A. Sakka, V. F. Silva, W. Jaafar, A. Ben Letaifa, and R. Langar, "DDoS attacks mitigation in 5G-V2X networks: A reinforcement learning-based approach," in *Proc. 19th Int. Conf. Netw. Service Manage. (CNSM)*, Oct. 2023, pp. 1–5.
- [11] D. Garcia-Roger, E. E. González, D. Martín-Sacristán, and J. F. Monserrat, "V2X support in 3GPP specifications: From 4G to 5G and beyond," *IEEE Access*, vol. 8, pp. 190946–190963, 2020.
- [12] A. Karukutla et al., "A modular and lightweight SDN-enabled 5G network slicing framework for additive manufacturing," in *Proc. IEEE 16th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2026, pp. 662–668.
- [13] A. M. Alwakeel, "Network slicing and stand alone 5G architectures: Paving the way for smart cities and IoT," in *Proc. IEEE 15th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jun. 2025, pp. 79–85.
- [14] V. Sritapan et al., "5G security evaluation process investigation," Dept. Defense (DoD), DHS Sci. Technol. Directorate, Cybersecurity Infrastruct. Secur. Agency (CISA), Tech. Rep. Version 1, May 2022.
- [15] M. Liyanage, A. Braeken, S. Shahabuddin, and P. Ranaweera, "Open RAN security: Challenges and opportunities," *J. Netw. Comput. Appl.*, vol. 214, May 2023, Art. no. 103621.
- [16] D. Attanayaka, P. Porambage, M. Liyanage, and M. Ylianttila, "Peer-to-peer federated learning based anomaly detection for open radio access networks," in *Proc. ICC - IEEE Int. Conf. Commun.*, May 2023, pp. 5464–5470.
- [17] 3GPP, *Security Aspects of 3GPP Support for Advanced Vehicle-to-Everything (V2X) Services*, document ETSI TS 133 536, 2022. [Online]. Available: <https://www.3gpp.org/dynareport/33536.htm>
- [18] A. Abouamar, A. Taik, A. Filali, and S. Cherkaoui, "Federated deep reinforcement learning for open RAN slicing in 6G networks," *IEEE Commun. Mag.*, vol. 61, no. 2, pp. 126–132, Feb. 2023.
- [19] T. Wink and Z. Nocht, "An approach for peer-to-peer federated learning," in *Proc. 51st Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. Workshops (DSN-W)*, Jun. 2021, pp. 150–157.
- [20] J. Le, D. Zhang, X. Lei, L. Jiao, K. Zeng, and X. Liao, "Privacy-preserving federated learning with malicious clients and honest-but-curious servers," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 4329–4344, 2023.
- [21] F. Alalyan, B. Bousalem, W. Jaafar, and R. Langar, "Secure peer-to-peer federated learning for efficient cyberattacks detection in 5G and beyond networks," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2024, pp. 1752–1757.
- [22] H. Yakan, I. Fajjari, N. Aitsaadi, and C. Adjih, "Federated learning for V2X misbehavior detection system in 5G edge networks," in *Proc. Int. ACM Conf. Model. Anal. Simul. Wireless Mobile Syst.*, Oct. 2023, pp. 155–163.
- [23] H. Wang et al., "Federated learning for V2X communications: Opportunities and challenges," in *Proc. IEEE Int. Symp. Broadband Multimed. Syst. Broadcast.*, Jun. 2024.
- [24] R. Song et al., "V2X-boosted federated learning for cooperative intelligent transportation systems with contextual client selection," in *Proc. ICRA Wrkshp. Collabo. Percept. Learn.*, 2023, pp. 1–5.
- [25] B. Hussain, Q. Du, B. Sun, and Z. Han, "Deep learning-based DDoS-attack detection for cyber-physical system over 5G network," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 860–870, Feb. 2021.
- [26] R. Asensio-Garriga et al., "ZSM-based E2E security slice management for DDoS attack protection in MEC-enabled V2X environments," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 485–495, 2024.
- [27] D. Han, H. Li, X. Fu, and S. Zhou, "Traffic feature selection and distributed denial of service attack detection in software-defined networks based on machine learning," *Sensors*, vol. 24, no. 13, p. 4344, Jul. 2024.
- [28] S. Anbalagan, G. Raja, S. Gurumoorthy, R. D. Suresh, and K. Dev, "IIDS: Intelligent intrusion detection system for sustainable development in autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 12, pp. 15866–15875, Dec. 2023.
- [29] S. Wijethilaka and M. Liyanage, "A federated learning approach for improving security in network slicing," in *Proc. GLOBECOM - IEEE Global Commun. Conf.*, Dec. 2022, pp. 915–920.
- [30] V. Turina, Z. Zhang, F. Esposito, and I. Matta, "Combining split and federated architectures for efficiency and privacy in deep learning," in *Proc. 16th Int. Conf. Emerg. Netw. Experiments Technol.*, Nov. 2020, pp. 562–563.
- [31] A.-A. Maiga, E. Ataro, and S. Githinji, "XGBoost and deep learning based-federated learning for DDoS attack detection in 5G core network VNFs," in *Proc. 6th Int. Conf. Comput. Commun. Internet (ICCCI)*, Jun. 2024, pp. 128–133.
- [32] S. Sheikhi and P. Kostakos, "DDoS attack detection using unsupervised federated learning for 5G networks and beyond," in *Proc. IEEE J. Europ. Conf. Netw. Commun. 6G Summit*, Jun. 2023.
- [33] Y. Liu, J. Peng, J. Kang, A. M. Ilyasu, D. Niyato, and A. A. A. El-Latif, "A secure federated learning framework for 5G networks," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 24–31, Aug. 2020.
- [34] F. Alalyan, M. Awad, W. Jaafar, and R. Langar, "Secure distributed federated learning for cyberattacks detection in B5G open radio access networks," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 3067–3081, 2025.
- [35] Google.LiteRT. Accessed: May 27, 2026. [Online]. Available: <https://ai.google.dev/edge/litert>
- [36] S. Ennaji, F. de Gaspari, D. Hitaj, A. Kbid, and L. Vincenzo Mancini, "Adversarial challenges in network intrusion detection systems: Research insights and future prospects," *IEEE Access*, vol. 13, pp. 148613–148645, 2025.
- [37] V. Tolpegin, S. Truex, M. E. Gürsoy, and L. Liu, "Data poisoning attacks against federated learning systems," in *Proc. Eur. Symp. Res. Comput. Secur.*, 2020, pp. 480–501.
- [38] ETSI TS 123 287 V16.3.0 (2020-07): 5G; Architecture Enhancements for 5G System (5GS) to Support Vehicle-to-Everything (V2X) Services, document ETSI TS 123 287 V16.3.0 (2020-07), 2020.
- [39] IEEE Std 802.11-2020: Wireless LAN MAC PHY Specifications, Standard IEEE Std 802.11-2020, 2020. [Online]. Available: <https://www.normsplash.com/Samples/IEEE/128485443/IEEE-802.11-2020-en.pdf>
- [40] K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning," in *Proc. ACM CCS*, 2017, pp. 1175–1191.
- [41] X. Li et al., "On the convergence of FedAvg on non-IID data," in *Proc. Int. Conf. Learn. Represent.*, 2020, pp. 1–27.
- [42] B. Bousalem, "DDoS Attacks in 5G-V2X Networks," IEEE Dataport, Apr. 2025, doi: [10.21227/vz7v-7951](https://doi.org/10.21227/vz7v-7951).
- [43] EurekaOpenAirInterface. Accessed: May 27, 2026. [Online]. Available: <https://openairinterface.org/>
- [44] Waveshare.JetRacer Pro AI Kit. Accessed: May 27, 2026. [Online]. Available: <https://www.waveshare.com/jetracer-pro-ai-kit.htm>
- [45] Waveshare.SIMCom SIM8202G-M2 5G Module Designed for Jetson Nano. Accessed: May 27, 2026. [Online]. Available: <https://www.waveshare.com/sim8202g-m2-5g-for-jetson-nano.htm>
- [46] M. Harounabadi, D. M. Soleymani, S. Bhadauria, M. Leyh, and E. Roth-Mandutz, "V2X in 3GPP standardization: NR sidelink in release-16 and beyond," *IEEE Commun. Standards Mag.*, vol. 5, no. 1, pp. 12–21, Mar. 2021.
- [47] QoSient.Argus. Accessed: May 27, 2026. [Online]. Available: <https://openargus.org/>
- [48] R. Neres Carvalho, J. Luiz Bordim, and E. Adilio Pelinson Alchieri, "Entropy-based DoS attack identification in SDN," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshops (IPDPSW)*, May 2019, pp. 627–634.

- [49] B. Mao, J. Liu, Y. Wu, and N. Kato, "Security and privacy on 6G network edge: A survey," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 1095–1127, 2nd Quart., 2023.
- [50] S. Kamei and S. Taghipour, "A comparison study of centralized and decentralized federated learning approaches utilizing the transformer architecture for estimating remaining useful life," *Rel. Eng. Syst. Saf.*, vol. 233, May 2023, Art. no. 109130.
- [51] S. Garst et al., "A comprehensive experimental comparison between federated and centralized learning," *Database*, vol. 2025, 2025, Art. no. baaf016.
- [52] G. Drainakis, P. Pantazopoulos, K. V. Katsaros, V. Sourlas, A. Amditis, and D. I. Kaklamani, "From centralized to federated learning: Exploring performance and end-to-end resource consumption," *Comput. Netw.*, vol. 225, Apr. 2023, Art. no. 109657.
- [53] O. Alliance, "O-RAN AI/ML workflow description and requirements 1.03," in *Proc. ML Workflow Description Requirements*, vol. 1, 2021, pp. 1–49.

BADRE BOUSALEM received the M.Sc. degree in computer science from ENSIAS, Morocco, in 2021. He is currently pursuing the Ph.D. degree with University Gustave Eiffel (UGE), France. He is the Research and Development Project Manager and the Information Security Officer at Metanext, France. His main research interests include machine learning, federated learning, reinforcement learning, cybersecurity, cloud computing, and next-generation vehicular networks. His current research focuses on cyberattacks response and mitigation using machine learning and software-defined networking in 5G and vehicle-to-everything (5G-V2X) environments.

FAHDAH ALALYAN (Graduate Student Member, IEEE) received the M.Sc. degree in information systems security from Concordia University, Canada, in 2021. She is currently pursuing the Ph.D. degree in computer engineering with the École de Technologie Supérieure (ÉTS), University of Quebec, Montreal, Canada. She has also worked on cyberattack detection in distributed systems. Her research interests include artificial intelligence, privacy preservation, quantum machine learning, anomaly detection, and next-generation wireless networks. Her current research focuses on machine learning, federated learning, quantum machine learning, and cybersecurity in 5G and open RAN environments.

VINICIUS F. SILVA received the M.Sc. and Ph.D. degrees in computer science from the Universidade Federal de Minas Gerais (UFMG), Brazil, in 2014 and 2021, respectively. He is currently a Post-Doctoral Researcher with UFMG. Previously, he was a Post-Doctoral Researcher at University Gustave Eiffel (UGE), France. His main research interests also include wireless sensor networks, vehicular networks, opportunistic networks, software-defined networks, device-to-device networks, and the Internet of Things. His current research focuses on 5G and beyond cellular networks, as well as 5G vehicle-to-everything (5G-V2X) networks.

SYLVAIN CHERRIER received the Ph.D. degree in computer science from University Paris-Est in 2013. He has been an Associate Professor at University Gustave Eiffel (UGE), France, since 2016. His main research interests include services choreography for the Internet of Things, as well as wireless sensor networks and vehicular networks. His current research focuses on 5G and vehicular networks.

WAEEL JAAFAR (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees from Polytechnique Montreal, Canada. From 2019 to 2022, he was with the Systems and Computer Engineering Department, Carleton University, as an NSERC Post-Doctoral Fellow. From 2014 to 2018, he has pursued a career in the telecommunications industry, where he has been involved in designing telecommunication solutions for projects across Canada and abroad. During his career, he was a Visiting Researcher at Khalifa University, Abu Dhabi, United Arab Emirates, in 2019; Keio University, Tokyo, Japan, in 2013; and UQAM, Montreal, Canada, in 2007. He has been an Associate Professor at the Software and IT Engineering Department, École de Technologie Supérieure (ÉTS), University of Quebec, Montreal, since September 2022. His research interests include wireless communications, integrated terrestrial and non-terrestrial networks, resource allocation, edge caching and computing, and machine learning for communication and networks. He received several prestigious grants, including the National Sciences and Engineering Research Council of Canada (NSERC) Alexander-Graham Bell Scholarship, the Fonds de Recherche du Québec-Nature et Technologies (FRQNT) Scholarship, and the best paper awards at IEEE ICC 2021, ISCC 2023, and CIoT 2024.

RAMI LANGAR (Member, IEEE) received the M.Sc. degree in network and computer science from University Pierre and Marie Curie (now Sorbonne University) in 2002 and the Ph.D. degree in network and computer science from Telecom Paris Tech, Paris, France, in 2006. He was a Post-Doctoral Research Fellow with the School of Computer Science, University of Waterloo, ON, Canada, from 2006 to 2008, and an Associate Professor with LIP6, University Pierre and Marie Curie, from 2008 to 2016. He has been a Full Professor of computer science at the École de Technologie Supérieure (ÉTS), Montreal, Canada, since 2021, and has been affiliated with University Gustave Eiffel (UGE), France, since 2016. His research interests include resource management in 5G/6G and O-RAN networks, green networking, machine learning for communications and networks, cybersecurity, digital twin, and quantum AI. He was a co-recipient of several best paper awards, including the IEEE/IFIP CNSM 2014, IFIP PEMWN 2016, IEEE NetSoft 2022, and CIoT 2024. He was the Chair of the IEEE ComSoc Technical Committee on Information Infrastructure and Networking (TCIIN) from January 2018 to December 2019.