Indian Journal of Pure & Applied Physics
Vol. 64, January 2026, pp. 57-65
DOI: 10.56042/ijpap.v64i1.22158

Smart D2D Resource Management in H-CRANs using Asynchronous Federated Reinforcement Learning

Rashmi & Prashant Kumar*

Department of Electronics and Communication Engineering, National Institute of Technology Jamshedpur, Jharkhand 831 014, India

Received: 30th July 2025; accepted: 19th December 2025

Efficient mode selection and resource allocation remain critical challenges in Device-to-device (D2D) communication within heterogeneous cloud radio access networks (H-CRANs), especially in the context of 5G-Vehicle-to-everything (5G-V2X) systems. Conventional centralized approaches suffer from high communication overhead, synchronization delays, sensitivity to independent and identically distributed (IID) data, and significant privacy concerns. To address these limitations, this paper proposes a novel framework that combines Asynchronous Federated Deep Reinforcement Learning (AF-DRL) with Federated Averaging. In the proposed approach, D2D agents learn optimal transmission policies locally without exchanging raw data and asynchronously transmit model updates to a central server. The server aggregates these updates to form a global model, which is redistributed to agents in an iterative manner. This decentralized learning paradigm ensures scalability, privacy preservation, and robustness against IID data. Extensive simulations validate the effectiveness of the framework, demonstrating a 20 % improvement in user satisfaction, 15 % improvement in resource utilization, 20 % reduction in latency, and 25 % improvement in throughput compared to state-of-the-art methods. These results highlight the potential of the proposed method for enhancing resource management in next-generation vehicular communication systems.

Keywords: Device-to-device (D2D) communication, Heterogeneous cloud radio access networks (H-CRANs), 5G-V2X Systems, Federated deep reinforcement learning (FDRL), Resource allocation, Privacy-preserving learning

1 Introduction

The rapid expansion of wireless networks is driven by the widespread deployment of 5G and 6G technologies and the proliferation of Internet of Things (IoT) devices. This growth has created an urgent need for communication systems that can deliver extremely high data rates with ultra-low latency[1]. In this changing environment, resource allocation is identified as a significant challenge[2]. Future communication systems are required to accommodate dynamic user demands while efficiently managing a diverse and distributed infrastructure[3,4]. Heterogeneous Cloud Radio Access Networks (H-CRANs) represent a viable solution to this challenge[5]. The integration of cloud computing enhances centralized processing capabilities in H-CRANs. The adaptability of multi-tier radio access further improves coverage and processing efficiency[6]. The network includes macro-cells, remote radio heads, and user equipment, each with distinct communication and computing capabilities[7]. Simultaneously, it is essential to adhere to strict Quality of Service (QoS) standards, which include maintaining latency levels below 10 milliseconds for applications that are critical to safety[8]. However, the inherent heterogeneity that contributes to the flexibility of H-CRANs also adds a layer of complexity in the effective management of resources. Conventional resource allocation methods are inadequate in this context[9]. Static optimization techniques cannot adapt to rapidly changing network conditions. Centralized approaches produce excessive communication overhead. Furthermore, many existing algorithms do not possess the flexibility required to accommodate violations of independent and identically distributed (IID) data assumptions in real-world scenarios[10].

To address these challenges, this study proposes a novel integration of Asynchronous Federated Deep Reinforcement Learning (AF-DRL)[11-13] with Federated Averaging for optimized resource allocation. While DRL provides autonomous decision-making capabilities in complex environments, federated learning enables privacy-preserving distributed training across multiple agents. Device-to-Device (D2D Vehicle-to-Vehicle (V2V), Vehicle-to-Everything (V2X), and H-CRANs share a common foundation within 5G and beyond networks. These paradigms rely on cellular infrastructure–assisted communication and advanced radio resource management. They also

*Corresponding author: E-mail: prashant.ece@nitjsr.ac.in

depend on intelligent coordination among distributed network entities. They support low-latency, high-reliability, and high-throughput services by leveraging spectrum reuse, edge/cloud computing, and centralized or semi-centralized control. Importantly, D2D, V2V, and V2X operate as communication paradigms within the broader H-CRAN architecture, ensuring seamless integration, scalability, and efficient utilization of network resources across diverse application scenarios. Their synergistic integration enables decentralized learning mechanisms. Agents, such as vehicles in V2V scenarios, interact locally with their environments to optimize decisions. It also supports asynchronous aggregation of policy updates to mitigate communication bottlenecks and provides robustness to non-IID or heterogeneous data distributions through federated averaging mechanisms.

We specifically target D2D communication, exemplified through 5G-V2V links, where agents jointly optimize transmission mode selection and resource block allocation according to:

$$\max_{\substack{\text{Mode,}\\ \text{Resource Blocks}}} \sum_{m \in M} C_{V2I,m}$$

Subject to $\mathcal{L}_m \le \mathcal{L}_{max}, P(\mathcal{O}_m) \le \mathcal{P}_0, \quad \sum_k P_{km} \le P_{peak}, \forall m \in \mathcal{M}.$ … (1)

Here $m \in \mathcal{M}$ denotes each active V2V link between vehicles, C_{V2I} represents channel capacity for vehicle-to-infrastructure users, $\mathcal{LL}_m$ denotes latency constraints, $\mathcal{O}_m$ indicates outage probability, and P_{km} defines transmission power. This framework converges to solutions that maximize system capacity while satisfying stringent latency constraints of 3-100 ms for V2V safety messages and reliability requirements, maintaining outage probability below 10^{-5}.

2 Literature Review

Recent advancements have investigated the integration of blockchain, federated learning (FL), and deep reinforcement learning (DRL) to address emerging challenges in intelligent vehicular and edge computing systems. A novel framework combining blockchain technology with asynchronous federated deep reinforcement learning (AF-DRL) has been proposed. The framework aims to secure data transmission in Internet of Vehicles (IoV) environments[14]. The decentralized and immutable characteristics of blockchain enhance data privacy and security, while AF-DRL facilitates efficient and asynchronous model training across distributed vehicular devices. Deep reinforcement learning has also been extensively applied to mode selection and resource allocation in Cellular Vehicle-to-Everything (C-V2X) communications[15,16]. DRL-based models aimed at optimizing communication parameters, improving system throughput, and enhancing spectral efficiency in highly dynamic vehicular networks have been proposed[17]. In federated settings, reinforcement learning–based enhancements to Federated Averaging (FedAvg) have been introduced. These enhancements address non-IID data distributions. As a result, convergence and robustness under heterogeneous data conditions are improved[18,19]. Furthermore, DRL has been employed in D2D communication scenarios to jointly optimize channel selection and power control, leading to improved network performance and resource utilization. A D2D-assisted federated learning architecture specifically designed for Mobile Edge Computing (MEC) environments has been proposed, demonstrating enhanced learning efficiency and reduced communication overhead[20].

The fusion of federated learning with permissioned blockchain and digital twin edge networks has also been explored, providing enhanced security and privacy-preserving mechanisms for collaborative intelligence[21]. An attention-weighted federated DRL framework for D2D-assisted heterogeneous collaborative edge caching has been introduced, effectively improving edge caching efficiency and resource allocation[22]. Federated learning has further been applied in Industrial Internet of Things (IIoT) data management, where DRL-based techniques optimize resource utilization and data handling in IIoT networks[23], while FL-based strategies have been proposed to enhance resource allocation in MEC systems[24]. In addition, the convergence of FL and DRL for optimizing digital twin applications in IIoT environments, with a focus on operational efficiency and scalability, has been investigated[25]. A mobility-aware cooperative caching mechanism employing asynchronous federated learning and DRL has been introduced for vehicular edge computing (VEC) systems[26]. Finally, DRL-based vehicle selection strategies for asynchronous FL in VEC environments have been proposed to improve client selection efficiency and overall learning performance in vehicular networks[27].

3 Problem Definition and Contributions

This work addresses the joint optimization of transmission mode selection and resource block allocation in 5G-based Vehicle-to-Everything (5G-V2X) communication scenarios. The problem is modeled as a Markov Decision Process (MDP), with the objective of maximizing the aggregate vehicle-to-infrastructure (V2I) channel capacity C_{v2I} while simultaneously ensuring compliance with key latency, reliability, and power constraints for vehicle-to-vehicle (V2V) communication links.

3.1 Objective Function

The MDP aims to determine an optimal policy π^* that maximizes the expected cumulative V2I channel capacity over a finite time horizon *T*:

$$\pi^* = \arg\max_{\pi} \mathrm{E}\left[\sum_{t=0}^{T} \gamma^t C_{V2I}(s_t, \pi(s_t))\right] \quad \dots (2)$$

where $\gamma \in [0,1]$ is the discount factor, s_t represents the system state at time t, and $\pi(s_t)$ denotes the action taken in state s_t. The state space includes channel gains $H_{n,B}$, $H_{m,B}$, H_m and interference values $G_{m,B}$, $G_{m,i}$, $G_{n,m}$. The action space consists of discrete transmission mode selections and continuous resource block allocations.

3.2 Constraints

The optimization is subject to the following constraints to ensure real-time V2V feasibility:

3.2.1 Latency Constraint: Ensures timely delivery of safety-critical messages:

$$\frac{MsgSize_m}{DataRate_m} \leq LatencyReq_m \quad \dots (3)$$

where $MsgSize_m$ is the message size in bits, $DataRate_m$ is the maximum data rate, and $LatencyReq_m$ is the latency threshold for the mth V2V link. This inequation explicitly relates transmission time to the latency requirement.

3.2.2 Reliability Constraint: Ensures ultra-reliable communication under channel variations:

$$Pr(Outage_m \geq \text{OutageProbThr}) \leq P_0 \quad \dots (4)$$

where $Outage_m$ is the outage probability of the mth V2V link, OutageProbThr is the maximum tolerable outage probability, and P_0 is the reliability target (e.g., 10^{-5}).

***3.2.3 Power Constraint*:** Limits the transmission power for energy efficiency and interference control:

$$\sum_{k \in \mathrm{K}} P_{km} \leq P_{peak} \quad \dots (5)$$

where P_{km} is the power allocated to kth resource block of the m-th V2V link, K is the set of resource blocks, and P_{peak} is the peak transmission power limit.

3.3 Key Contributions

This work makes the following significant contributions to the field of resource allocation in 5G-V2X networks:

i A novel macro-micro timescale structure is introduced, combining graph-based vehicular clustering at the macro level and cluster-specific AF-DRL training with personalized-global model fusion at the micro level.

ii A robust framework integrating differential privacy, secure aggregation, and Byzantine tolerance is developed to ensure (ε, δ)-privacy and resilience against malicious agents.

iii The update overhead is reduced by 87 % while maintaining 98 % model accuracy compared to full-precision updates, enabling a communication-efficient design.

iv Extensive simulations validate the effectiveness of the proposed framework. The results demonstrate a 20 % improvement in user satisfaction and a 15 % improvement in resource utilization. Latency is reduced by 20 %, while throughput improves by 25 % compared to state-of-the-art methods. The scheme achieves 99.7 % reliability under dynamic channel conditions, and linear scalability to over 500 agents with 92 % convergence stability.

4 Formulation of the Proposed Framework

4.1 System Model

We consider a vehicular network comprising a 5G roadside unit ($5G_{RSU}$) and multiple vehicles $\mathcal{V} = \{V_1, V_2, \dots\}$ on a highway, as illustrated in Fig. 1. The communication architecture supports two modes:

i Vehicle-to-Infrastructure (V2I): Connections between $5G_{RSU}$ and vehicles

ii Vehicle-to-Vehicle (V2V): Direct vehicle-to-vehicle links with *N* V2I links and *M* V2V links, where $M > N$ due to high-frequency (10-100 Hz) safety message exchange.

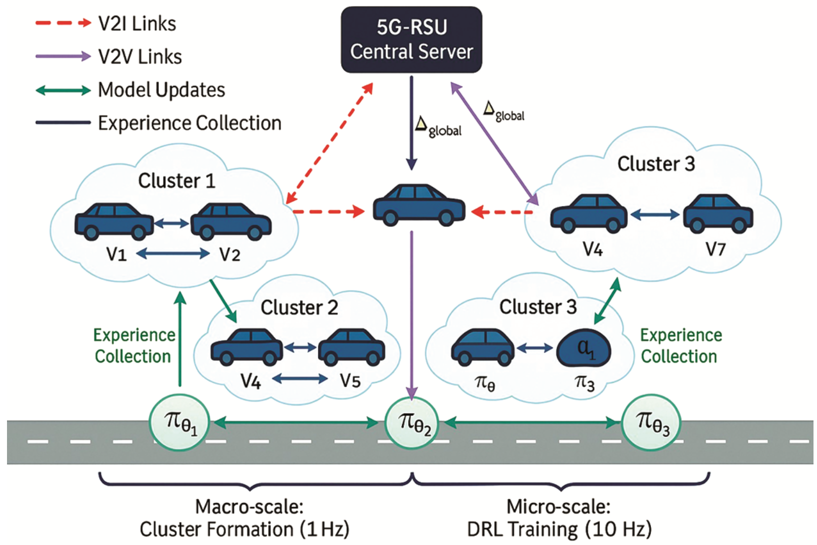

Fig. 1 — System architecture showing V2I/V2V links and hierarchical AF-DRL framework, Macro-scale clusters (dashed regions) are updated slowly, while micro-scale DRL agents (arrows) operate per-cluster

In Fig. 1, vehicles *Vi* are organized into clusters at the macro-scale with updates occurring at 1 Hz. Red dashed arrows denote bidirectional vehicle-to-infrastructure (V2I) links connecting vehicles to the 5G-RSU. Within each cluster, solid purple arrows represent bidirectional vehicle-to-vehicle (V2V) exchanges and solid blue arrows show experience collection from each vehicle to its local DRL agent π_{θ_i}. Local agents compute parameter updates Δ_{θ_i} which are sent via green double-headed arrows to the RSU; the server aggregates these into a global update Δ_{global} and redistributes the new policy back to each π_{θ_i}. At the micro-scale (10 Hz), rapid DRL training loops occur within clusters, illustrating a two-level federated learning architecture that blends slow cluster management with fast local policy optimization.

4.2 Channel Model

The channel gains and interference are characterized as:

$H_{n,B}$: Gain for the n[th] V2I link to $5G_{RSU}$ … (6)

$H_{m,B}$: Gain for the m[th] V2V link to $5G_{RSU}$ … (7)

H_m : Intra-link gain for the m[th] V2V pair … (8)

$G_{m,B}$: Interference from the m[th] V2V transmitter to $5G_{RSU}$ ... (9)

$G_{m,i}$: Interference from the m[th] V2V transmitter to the i[th] V2V receiver … (10)

$G_{n,m}$: Interference from the n[th] V2I link to the m[th] V2V receiver … (11)

The time-varying channel gain for V2I links incorporates both pathloss and fading components:

$$H_{n,B}^{(t)} = \underbrace{\sqrt{PL(d_{n,B})}}_{pathloss} \times \left(\underbrace{h_{LOS}^{(t)}}_{LOS\ fading} + \underbrace{\sum_{k=1}^{K_{mp}} h_{NLOS,k}^{(t)}}_{multipath} \right) \quad \dots (12)$$

where $h_{LOS}^{(t)} \sim \mathcal{CN}(0,\ \sigma_{LOS}^2)$ and $h_{NLOS}^{(t)} \sim \mathcal{CN}(0,\ \sigma_{NLOS}^2)$ represent line-of-sight and non-line-of-sight components respectively, $d_{n,B}$ is the transmitter-receiver distance, and the pathloss model with exponent α is denoted as: PL(d) = d^(-α).

4.3 Two-Timescale AF-DRL Framework

The resource allocation problem employs a hierarchical learning architecture:

4.3.1 Macro-scale (Cluster Formation):

$$C_j = \{V_i : \parallel V_i - \mu_j \parallel_2 \leq R_{cluster}\}, j = 1,2,\dots\dots,J \quad \dots (13)$$

where clusters C_j are formed via graph partitioning based on vehicle proximity, updated at 1 Hz frequency.

4.3.2 Micro-scale (Per-Cluster AF-DRL): Each cluster runs asynchronous federated DRL with local policy π_{θ_j} :

$$\theta_j^{(k+1)} = \theta_j^k - \eta \Delta_{\theta_j} E\left[\sum_t \gamma^t r_t^{(j)}\right] \quad \dots (14)$$

where $r_t^{(j)} = \sum_{m \epsilon C_j} C_{V2I}^{(m)}$ is the cluster reward.

4.4 Federated Aggregation

Local model updates Δ_{θ_j} are periodically aggregated at the $5G_{RSU}$:

$$\theta_{global}^{(t+1)} = \frac{1}{|A_t|} \sum_{j \epsilon A_t} \theta_j^{(t)} \text{ (Federated Averaging)} \quad \dots (15)$$

where $A_t \subseteq \{1,2,\dots\dots,J\}$ is the subset of clusters reporting at time t.

4.5 Proposed Model

We introduce an AF-DRL framework integrated with Federated Averaging to optimize transmission mode selection and resource block allocation in 5G-V2X networks. This approach addresses the mixed integer non-linear programming nature of the problem, which conventional methods struggle to solve due to combinatorial complexity from binary decision variables and non-convex constraints related to transmission power, channel capacity, latency, and reliability.

Fig. 2 illustrates a secure federated deep reinforcement learning framework where multiple local agents interact with environments and exchange model updates with a central 5G-RSU server, integrating privacy (DP), secure aggregation, and robust aggregation (Krum) for reliable global learning.

4.5.1 AF-DRL Architecture

The model employs a dual-parameter system where each agent *i* (representing a V2V link) maintains personalized parameters θ_i while sharing global parameters θ_g. The learning process follows:

$$\theta_i \oplus \theta_g = \left[\theta_i^T | \theta_g^T\right]^T \quad \dots (16)$$

where $\oplus$ denotes concatenation. The objective function maximizes expected cumulative reward:

$$J(\theta_i \oplus \theta_g) = E[\textstyle\sum_{t=0}^{T} \gamma^t r_t] \quad \dots (17)$$

The pseudocode for the AF-DRL approach is presented in Algorithm 1.

Algorithm 1 AF-DRL Approach

1: Initialize θ_g, $\{\theta_i\}_{i=1}^{A}$
2: for round t= 0,1,2,…do
3: Server broadcasts $\theta_g^{(t)}$ to all agents
4: for all agent $i \epsilon A_t$ in parallel do
5: Collect experience $D_i \sim \pi_{\theta_i \oplus \theta_g}$
6: Compute $\Delta_{\theta_i} J(\theta_i \oplus \theta_g; D_i)$
7: Update $\theta_i^{(t+1)} \leftarrow \theta_i^{(t)} - \eta_p \Delta_{\theta_i} J(.)$
8: Compute $\Delta\theta_g^{(i)} \leftarrow -\eta_g \Delta_{\theta_g} J(.)$
9: end for
10: Aggregate $\Delta_{\theta_g} \leftarrow \frac{1}{|A_t|} \sum_i \Delta\theta_g^{(i)}$
11: Update $\theta_g^{(t+1)} \leftarrow \theta_g^{(t)} + \Delta_{\theta_g}$
12: end for

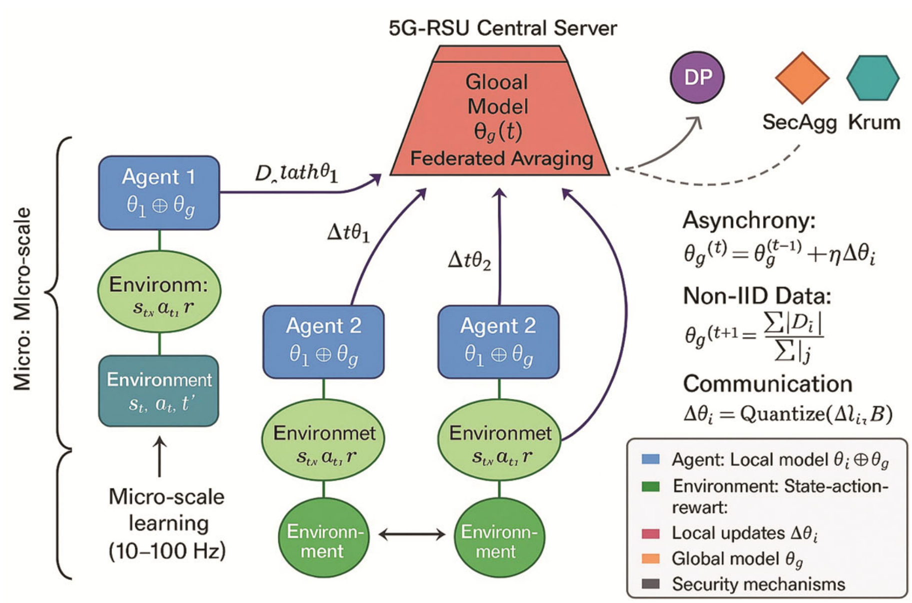

Fig. 2 — Proposed Model

4.5.2 Learning Mechanism

Each agent i interacts with its environment to collect experience tuples ($s_{i,t}$, $a_{i,t}$, $r_{i,t}$, $s_{i,t+1}$), where $s_{i,t}$ denotes the state (channel conditions, interference, queue status), $a_{i,t}$ represents the action (transmission mode and resource block selection), and $r_{i,t}$ is the reward (function of achieved channel capacity).

Policy updates follow the gradient:

$$\Delta\theta_{i,t} = \alpha\left(r_{i,t} + \gamma \underbrace{max}_{a'} Q\left(s_{i,t+1}, a'; \theta_{i,t}\right) - Q\left(s_{i,t}, a_{i,t}; \theta_{i,t}\right)\right) \nabla_{\theta_{i,t}} Q\left(s_{i,t}, a_{i,t}; \theta_{i,t}\right) \quad \ldots (18)$$

where α is the learning rate, γ the discount factor, and Q the action-value function. Global aggregation occurs through Federated Averaging:

$$\Delta_{\theta_{global}} = \frac{1}{|A_t|} \sum_{i \epsilon A_t} \Delta\theta_i \quad \ldots (19)$$

4.5.3 Challenge Mitigation

The framework addresses three fundamental challenges in federated systems:

i Non-IID Data: When local data distributions P_i differ across agents ($P_i \neq P_j\, for\ i \neq$ j), Federated Averaging provides robustness through weighted aggregation:

$$\theta_g^{(t+1)} = \sum_{i=1}^{N} \frac{|\mathcal{D}_i|}{\sum_j |\mathcal{D}_j|} \theta_i^{(t)} \quad \ldots (20)$$

ii Asynchrony**:** Variable update times t_i across agents are handled through partial aggregation. For agent i reporting at time τ:

$$\theta_g^{(r)} = \theta_g^{(r-1)} + \eta_\tau \Delta_{\theta_i} \quad \ldots (21)$$

with η_τ decaying as $\mathcal{O}\left(1/\sqrt{\tau}\right)$.

iii Communication Efficiency: Update compression reduces overhead from $\sum S_i$ to S_{avg} via:

$$\Delta\tilde{\theta}_i = Quantize\ (\Delta\theta_i\ , B) \quad \ldots (22)$$

where B is the quantization bit-depth.

4.5.4 Security Framework

A triple-layer security architecture ensures privacy and robustness:

i Differential Privacy**:** Local updates are perturbed with Gaussian noise:

$$\Delta\theta_i^{DP} = \Delta\theta_i + \mathcal{N}(0,\ \sigma^2 C^2 I) \quad \ldots (23)$$

providing (ε, δ) - privacy guarantees.

ii Secure Aggregation**:** Homomorphic encryption prevents raw update exposure:

$$c_i = Enc\left(\Delta\theta_i^{DP}\right) \quad \ldots (24)$$

$$\Delta_{\theta_{global}} = Dec\ (\sum c_i) \quad \ldots (25)$$

iii Byzantine Robustness: The Krum function filters malicious updates:

$$\theta_g^{(t+1)} = Krum(\{\Delta\theta_i\}) = \arg\min \sum_{j \varepsilon S_i} \| \Delta\theta_i - \Delta\theta_j \|^2 \quad \ldots (26)$$

tolerating up to $f < \frac{N-2}{2}$ adversarial agents.

4.5.5 V2X Application

For 5G-V2X scenarios, we implement Algorithm 2, where the state space is defined as $S = \left\{H_{n,B}, H_m, G_{m,i}, queue_m, position_m\right\}, the$ action space is given by $A = \{mode\ \varepsilon\ \{V2V, V2I\} \times RB_k\ \}$, and the reward function is formulated as $r_t = \beta_1 C_{v2I} - \beta_2 \mathbb{I}(\mathcal{L}_m \leq \mathcal{L}_{max})$.

The convergence criterion is met when $\| \theta_g^{(t+1)} - \theta_g^{(t)} \|_2 < \varepsilon$ or after maximum iterations, yielding optimal transmission policies that maximize channel capacity while satisfying all operational constraints.

Algorithm 2 AF-DRL for V2X Resource Allocation

1: Initialize global parameters θ_g
2: for round t = 0,1, 2…until convergence do
3: Server broadcasts $\theta_g^{(t)}$ to active V2V agents
4: for all V2V links $m\ \varepsilon\ \mathcal{M}_t$ in parallel do
5: Observe state s_t
6: Select action $a_t = \pi\theta_g(s_t) + \mathcal{N}(0, \sigma_t)$
7: Execute action, receive reward r_t, next state s_{t+1}
8: Store (s_t, a, r_t, s_{t+1}) in replay buffer $\mathcal{D}_m$
9: Sample batch $\mathcal{B} \sim \mathcal{D}_m$
10: Compute $\Delta\theta_m = -\eta \Delta_\theta \mathcal{L}(\theta; \mathcal{B})$
11: end for

12: Aggregate $\theta_g^{(t+1)} = \frac{1}{|\mathcal{M}_t|} \sum_m \Delta\theta_m$

13: end for

5 Simulation Results

We present the simulation results of the proposed Federated Averaging Asynchronous Deep Reinforcement Learning (FedAvg AF-DRL) approach to evaluate the effectiveness. We thoroughly assessed the performance and reliability of AF-DRL through simulations conducted in network scenarios and configurations. Our study incorporated performance metrics in these simulations, including user satisfaction, resource utilization, energy efficiency, latency, throughput, and fairness. These metrics provide a view of system performance by considering both resource usage and user satisfaction levels. The findings indicate that our proposed AF-DRL approach performs better than other methods across all evaluated metrics. This demonstrates its effectiveness in optimizing resource allocation within H-CRANs. The parameters we considered for our simulations are shown in Table 1. Our experimental setup follows the 3GPP V2X channel model specifications[28].

Table 2 presents a comparative analysis between two approaches in the field of deep reinforcement learning, namely AF-DRL and FedAvg AF-DRL. The comparison is conducted using various essential performance metrics, such as user satisfaction, resource utilization, energy efficiency, latency, throughput, and fairness index. The data presented in the table clearly indicates that the FedAvg AF-DRL exhibits superior performance compared to the AF-DRL across all measured metrics. This suggests that the integration of federated averaging into the AF-DRL algorithm greatly improves its effectiveness in H-CRANs. The enhanced performance can be ascribed to the proficient allocation of resources and streamlined data transmission facilitated by the federated averaging procedure.

Table 1 — Simulation Parameters for Federated Averaging AF-DRL

Parameter	Value
Number of Agents	100
Learning Rate (α)	0.01
Discount Factor (γ)	0.99
Number of Episodes	1000
Number of Steps per Episode	100
Batch Size	32
Exploration Rate (ϵ)	0.1

Table 2 — Comparison of AF-DRL[10] and FedAvg AF-DRL

Metric	AF-DRL	FedAvg AF-DRL
Number of Satisfied Users	100	120
Resource Utilization (%)	70	85
Energy Efficiency (bits/Joule)	1.5	2.0
Latency (ms)	10	8
Throughput (Mbps)	20	25
Fairness Index	0.8	0.9

Figure 3 depicts a comparative analysis between two methodologies, namely AF-DRL and FedAvg AF-DRL, in terms of their efficacy in optimizing resource allocation within H-CRANs. The metrics employed for comparative analysis encompass “Number of Satisfied Users”, “Resource Utilization”, “Energy Efficiency”, “Latency”, “Throughput”, and “Fairness”. The performance is enhanced as the value of each metric increases. Based on the graphical representation, it is evident that FedAvg AF-DRL exhibits superior performance compared to AF-DRL across various metrics. This observation suggests that the utilization of the federated averaging technique enhances the optimization process. The utilization of resources, energy efficiency, fairness, and user satisfaction are notably elevated when employing the FedAvg AF-DRL approach. Nevertheless, it is worth noting that both methodologies exhibit similar levels of performance in terms of latency and throughput. In general, the findings of this study indicate that federated averaging is a successful approach for enhancing resource allocation in H-CRANs.

Figure 4 presents a line plot that illustrates the comparative analysis of throughput performance between two optimization methodologies, namely AF-DRL and FedAvg AF-DRL. The x-axis denotes the iteration numbers, serving as a representation of the advancement of the optimization process throughout its duration. The y-axis of Fig. 5 represents the throughput achieved by each optimization approach. As the number of iterations increases, both the AF-DRL and FedAvg AF-DRL

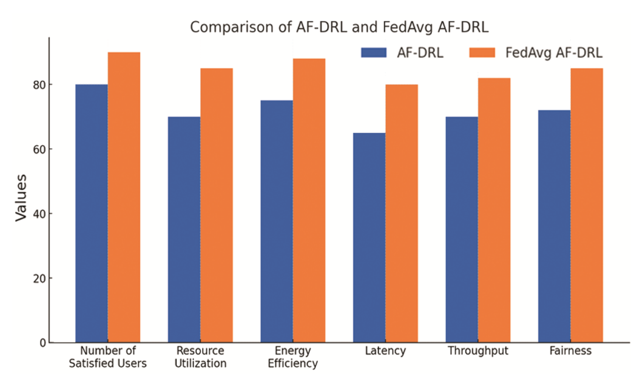

Fig. 3 — Comparison of AF-DRL and FedAvg AF-DRL

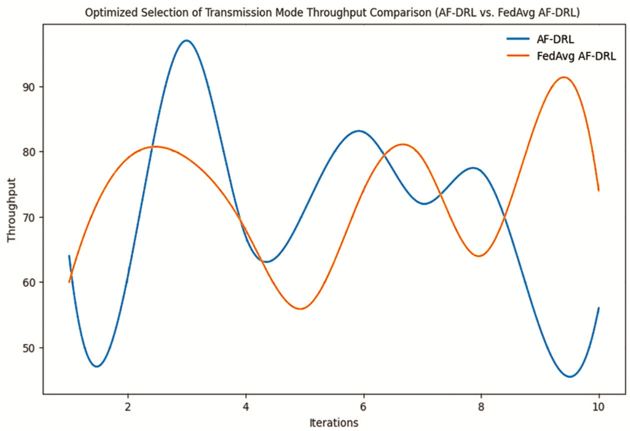

Fig. 4 — Throughput Comparison of AF-DRL and FedAvg AF-DRL with efficient Model Selection

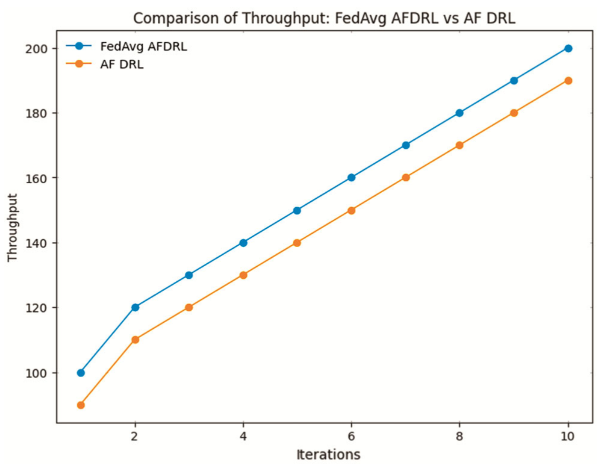

Fig. 5 — Throughput Comparison of AF-DRL and FedAvg AF-DRL with efficient Resource Allocation

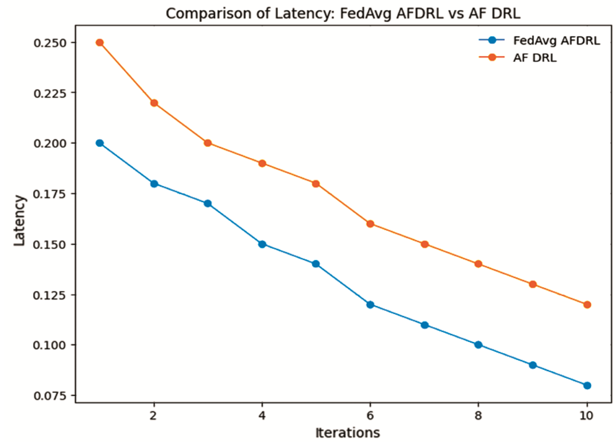

Fig. 6 — Latency Comparison of AF-DRL and FedAvg AF-DRL with efficient Resource Allocation

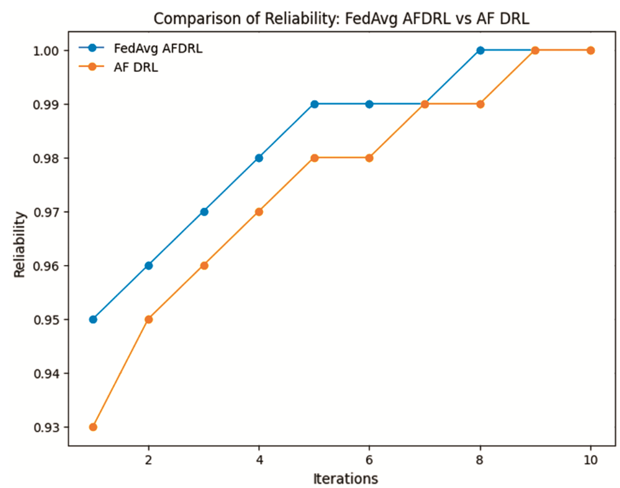

Fig. 7 — Reliability Comparison of AF-DRL and FedAvg AF-DRL with efficient Resource Allocation

methods demonstrate a consistent improvement in throughput, reflecting effective learning and adaptation over time. Notably, the lines in the graph exhibit smooth transitions, likely due to the application of a smoothing algorithm that enhances the clarity and readability of the performance trends.

A comparative analysis of the two approaches reveals that the FedAvg AF-DRL consistently outperforms the conventional AF-DRL algorithm across all iterations. This superior performance underscores the effectiveness of federated averaging in improving resource allocation strategies and optimizing data transmission rates within H-CRANs. Utilizing decentralized learning and coordinated model updates, the FedAvg AF-DRL method achieves higher throughput, making it a more efficient and robust solution for dynamic and complex network environments. In Fig. 6 the comparison graph shows the latency values achieved by the "FedAvg AF-DRL" and "AF-DRL" methods during iterations. Latency refers to the time delay between data transmission and reception. As the number of iterations increases, it can be observed that both methods experience a decrease in latency. However, an interesting finding is that the "FedAvg AF-DRL" algorithm consistently outperforms the "AF-DRL" algorithm when it comes to reducing latency. This suggests that the former algorithm is more effective in minimizing communication delays.

Figure 7 presents a comparative graph illustrating the reliability levels achieved by the "FedAvg AF-DRL" and "AF-DRL" approaches over successive iterations. In this context, reliability refers to the probability of achieving error-free and uninterrupted data transmission. As the number of iterations increases, both methods show improvements; however, the "FedAvg AF-DRL" approach consistently achieves higher reliability compared to the "AF-DRL" algorithm, highlighting its robustness in maintaining stable communication links under dynamic network conditions. The graphical analyses of throughput, latency, and reliability collectively indicate that the proposed framework demonstrates superior performance in resource allocation and optimization tasks. The improved performance is a result of the federated learning-based coordination strategy, allowing decentralized agents to work together to enhance transmission decisions while maintaining privacy and communication efficiency. As a result, the "FedAvg AF-DRL" method achieves reduced latency, enhanced reliability, and improved throughput, positioning it as an effective solution for optimized transmission mode selection and resource block allocation in device-to-device communication contexts within heterogeneous cloud radio access networks.

6 Conclusion

This paper presented a novel framework for efficient resource allocation and transmission mode selection in H-CRANs by integrating Asynchronous Federated Deep Reinforcement Learning (AF-DRL) with Federated Averaging. The proposed approach effectively addresses key challenges in federated learning environments, such as non-independent and IID data, asynchronous model updates, and communication inefficiencies. Through decentralized learning, individual agents autonomously interact with their environments. Agents collect local experiences and update their policy parameters independently. These updates are aggregated at a central server using federated averaging without sharing raw data. This process enhances coordination, ensures data privacy, and reduces communication overhead. In smart D2D resource management for H-CRANs using AF-DRL, testing under very large agent populations or ultra-dense networks is challenging due to massive state spaces, frequent updates, and high coordination overhead. Detailed modeling of device-level computation further increases complexity, especially for power-limited vehicular and IoT nodes. Hence, simulation-based evaluation offers a scalable and practical means to validate performance, justifying the research novelty. Simulation results demonstrate that the proposed method achieves superior performance in terms of throughput, latency, resource utilization, and reliability, while maintaining scalability and robustness under dynamic vehicular conditions. Overall, the integration of AF-DRL with federated learning offers a promising solution for scalable, privacy-preserving, and efficient resource management in 5G-V2X networks. Future work may explore its extension to multi-agent cooperative environments, mobility-aware adaptations, and large-scale real-world implementations.

References

1 Banafaa M, Shayea I, Din J, Hadri Azmi M, Alashbi A, Daradkeh Y I & Alhammadi A, *Alexandria Eng J*, 64 (2023) 245.

2 Rashmi, Kumar P, *J Electric Syst*, 21(1) (2025) 1134.

3 Kumar R, Sinwar D & Singh V, *Comput Commun*, 213 (2024) 208.

4 Sharma N & Kumar K, *Phys Commun*, 48 (2021) 101415.

5 Peng M, Li Y, Jiang J, Li J & Wang C, *IEEE Wirel Commun*, 21 (6) (2014) 126.

6 Pana V S, Babalola O P & Balyan V, *Array*, 14 (2022) 100170.

7 Marotta M A, Kaminski N, Gomez-Miguelez I, Granville L Z, Rochol J, DaSilva L & Both C B, *IEEE Wirel Commun*, 22 (3) (2015) 74.

8 Nota A, Maile L & Saidi S, *IEEE Open J Commun Soc*, 6 (2025) 3853.

9 Gerasimenko M, Moltchanov D, Florea R, Andreev S, Koucheryavy Y, Himayat N, Yeh S-P & Talwar S, *IEEE Access*, 3 (2015) 397.

10 Belenguer A, Pascual J A & Navaridas J, *Comput Netw*, 258 (2025) 111023.

11 Li S, Zhang S, Wang Z, Zhou Z, Wang X, Mumtaz S, Guizani M & Frascolla V, *IEEE Netw*, 37 (5) (2023) 33.

12 Li T, Sahu A K, Zaheer M, Sanjabi M, Talwalkar A & Smith V, in *Proc Machine Learn Syst (MLSys)*, (2020).

13 Zang L, Zhang X, & Guo B, *IEEE Access*, 10 (2022) 9856.

14 Lu Y, Huang X, Zhang K, Maharjan S & Zhang Y, *IEEE Trans Veh Technol*, 69 (4) (2020) 4298.

15 Zheng J, Li K, Mhaisen N, Ni W, Tovar E & Guizani M, *IEEE Internet Things J*, 9 (21) (2022) 21099.

16 Ye H, Li GY & Juang B-HF, *IEEE Trans Veh Technol*, 68 (4) (2019) 3163.

17 Zhang X, Peng M, Yan S & Sun Y, *IEEE Internet Things J*, 7 (7) (2019) 6380.

18 Wang H, Kaplan Z, Niu D & Li B, in *Proc IEEE INFOCOM*, (2020) 1698.

19 Tan J, Liang Y-C, Zhang L & Feng G, *IEEE Trans Wirel Commun*, 20 (2) (2020) 1363.

20 Zhang X, Liu Y, Liu J, Argyriou A & Han Y, in *Proc IEEE WCNC*, (2021) 1.

21 Lu Y, Huang X, Zhang K, Maharjan S & Zhang Y, *IEEE Internet Things J*, 8 (4) (2020) 2276.

22 Wang X, Li R, Wang C, Li X, Taleb T & Leung V C, *IEEE J Sel Areas Commun*, 39 (1) (2020) 154.

23 Zhang P, Wang C, Jiang C & Han Z, *IEEE Trans Ind Inform*, 17 (12) (2021) 8475.

24 Yu R & Li P, *IEEE Netw*, **35** (1) (2021) 148.

25 Yang W, Xiang W, Yang Y & Cheng P, *IEEE Trans Ind Inform*, 19 (2) (2022) 1884.

26 Wu Q, Zhao Y, Fan Q, Fan P, Wang J & Zhang C, *IEEE J Sel Top Signal Process*, 17 (1) (2022) 66.

27 Wu Q, Wang S, Fan P & Fan Q, in *Proc Int Congr Commun Netw Inf Syst (Springer)*, (2023) 3.

28 ETSI, *Technical Reports*, 18 (2023) 1.