FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 493-500 | e-ISSN: 2347-470X

Optimized D2D Mode Selection in H-CRANs Using Asynchronous Federated Deep Reinforcement Learning with Federated Averaging

Abhishek Malviya[*1,2] **Sudhakar Pandey**[2]

[1]*Department of Computer Science & Engineering, United Institute of Technology, Prayagraj, India; Email: abhishekmalviya2050@gmail.com*
[2]*Department of Information Technology, NIT Raipur, Chhattisgarh, India; Email: spandey.it@nitrr.ac.in*

***Correspondence**: *abhishekmalviya2050@gmail.com;*

ABSTRACT- In the context of heterogeneous cloud radio access networks (H-CRANs), there's a growing challenge in achieving efficient resource allocation due to the surge in data and complexities related to asynchrony in federated learning frameworks. This research presents an innovative methodology that blends federated averaging with asynchronous federated deep reinforcement learning (AF-DRL). By enabling individual agents to interact with distinct environments and subsequently adjust policy parameters, local updates are consolidated at a central point to generate a global update. Through this iterative procedure, the system attains optimal transmission mode selection and enhanced resource block distribution, thereby significantly boosting the overall performance of H-CRANs while ensuring that latency and reliability benchmarks are met.

Keywords: Heterogeneous Cloud Radio Access Networks (H-CRANs), Resource Allocation, Federated Averaging, Asynchronous Federated Deep Reinforcement Learning (AF-DRL), Communication Efficiency, Federated Learning.

ARTICLE INFORMATION
Author(s): Abishek Malviya, Sudhakar Pandey;
Received: 22/03/2024; **Accepted:** 14/07/2025; **Published:** 30/08/2025;
e-ISSN:2347-470X;
Paper Id:IJEER240121
Citation: 10.37391/ijeer.130310;
Webpage-link: https://ijeer.forexjournal.co.in/archive/volume-13/ijeer-130310.html

Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Heterogeneous Cloud Radio Access Networks (H-CRANs) have gained attention as a realistic and promising answer to the growing needs for higher data rates and lower latency in wireless communication. The network's innate variety and the users' various Quality of Service (QoS) requirements make resource allocation in H-CRANs a difficult task. Due to the intricacy of the problem and the changing nature of network environment characteristics, traditional resource allocation approaches typically perform inadequately.

In the current study, federated averaging and asynchronous federated deep reinforcement learning (AF-DRL) are provides a fundamental decision-making framework in complicated settings, federated learning facilitates distributed learning across several agents. Combining them can improve the efficiency of dispersed, flexible resource allocation combined to propose a unique method to resource allocation in H-CRANs. While Deep Reinforcement Learning (DRL). The suggested AF-DRL algorithm enables efficient communication between individual agents. Vehicles are viewed as the primary communication nodes in this study. In the area of wireless communication, the phrase" Vehicle-to-Vehicle (V2V) communication" is frequently used as an analog to the historical" Device-to-Device (D2D) communication."

In a 5G-V2V communication situation in particular, the autonomous agent has the ability to interact with its environment and modify the parameters of its policy or value function. The global update is then computed at the server using the local updates, and it is subsequently sent back to the agents. Repeating this procedure until convergence is attained results in the best possible resource block distribution and transmission mode choice, which maximizes potential rewards while simultaneously satisfying latency and reliability constraints.

The remainder of the essay is divided into the following sections: The system model and problem formulation are explained in *section 3*. The AF-DRL technique combined with Federated Averaging is shown in *section 4*. The simulation results and performance analysis are revealed in *section 5*. The paper is concluded in *section 6*.

2. LITERATURE SURVEY

The authors of this paper [1] propose a novel approach that combines blockchain technology and asynchronous federated learning (AFDRL) to ensure secure data transmission in Internet of Vehicles (IoV) scenarios. The immutability and decentralization inherent in blockchain technology contribute to the bolstering of privacy and security measures for stored data. Furthermore, the AFDRL solution guarantees the optimization of efficient and asynchronous Internet of Vehicles (IoV) device learning. This study [2] introduces the utilization of Deep

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 493-500 | e-ISSN: 2347-470X

Reinforcement Learning (DRL) as a solution for addressing the challenges of mode selection and resource allocation in Cellular Vehicle-to-Everything (C-V2X) communications.

The primary objective of the proposed Deep Reinforcement Learning (DRL) model is to enhance system performance and throughput by optimizing communication channels and resource units. [3] This study introduces a methodology based on Deep Reinforcement Learning (DRL) to address the challenges of mode selection and resource allocation in C-V2X communications. The objective is to enhance the overall efficiency of the system and optimize communication parameters. The present study [4] presents a novel approach to enhance the efficiency of Federated Learning (FedAvg) on datasets that are not Independent and Identically Distributed (non-IID) through the utilization of Reinforcement Learning (RL). The proposed approach successfully addresses the challenges associated with non-independent and identically distributed (non-IID) data distribution in federated learning settings. The authors of this paper [5] employ Deep Reinforcement Learning (DRL) to optimize the processes of channel selection and power control in Device-to-Device (D2D) networks. The objective of the proposed methodology is to enhance resource allocation and optimize network performance for device-to-device (D2D) interactions. The present study [6] introduces D2D-Assisted Federated Learning, a novel approach to federated learning specifically tailored for Mobile Edge Computing (MEC) networks. The efficacy and performance of federated learning in mobile edge computing (MEC) scenarios can be enhanced through the utilization of device-to-device (D2D) communication, as suggested by the proposed solution.
The present study [7] introduces an innovative approach aimed at enhancing the performance of Federated Learning (FedAvg) and Permissioned Blockchain in the context of Digital Twin Edge Networks. The integrated framework guarantees the confidentiality and effectiveness of data exchange in federated learning across peripheral networks. The authors of this paper [8] propose Attention-Weighted Federated Deep Reinforcement Learning as a solution for Device-to-Device (D2D) Assisted Heterogeneous Collaborative Edge Caching. The method being proposed aims to enhance the efficiency of peripheral caching and optimize the allocation of resources. The present study [9] examines the utilization of Deep Reinforcement Learning (DRL) as a means to enhance federated learning within the realm of data management in Industrial Internet of Things (IIoT) settings. The algorithm under consideration seeks to optimize resource utilization and enhance efficiency in data management. [10] This paper presents a novel approach called Federated Learning (FedAvg) that aims to optimize resource utilization in Mobile Edge Computing (MEC) environments. The objective of the proposed methodology is to enhance the efficiency and utilization of resources in Mobile Edge Computing (MEC) systems through the improvement of federated learning.

The present study [11] introduces a novel approach known as Mobility-Aware Cooperative Caching in Vehicular Edge Computing. This technique is based on the principles of Asynchronous Federated and Deep Reinforcement Learning. The proposed solution enhances the effectiveness and efficiency of cooperative cache in vehicular edge computing scenarios. The researchers put forward a framework for optimizing the Industrial Internet of Things (IIoT) by combining Federated Learning and Deep Reinforcement Learning (DRL) techniques [12]. The objective of this strategy is to optimize the efficiency and output of digital twin applications within the context of Industrial Internet of Things (IIoT) environments. [13] The authors of this study aim to examine the matter of vehicle selection within the realm of asynchronous federated learning in vehicular edge computing. They propose a unique approach that relies on deep reinforcement learning (DRL). The objective of this technique is to improve the effectiveness of federated learning in vehicular edge computing systems through the optimization of vehicle selection. The rapid advancements in deep reinforcement learning (DRL) strategies and the Open RAN (O-RAN) paradigm have led to the dynamic positioning of traffic-aware local decision agents (DAs) within the radio access network (RAN) [14]. Although federated learning (FL) has been successfully applied in a number of theoretical contexts, its actual implementation in the Metaverse industry is fraught with difficulties. These difficulties include non-independent and identically distributed data, learning decay resulting from continuous industrial data, and a limited communication bandwidth [19]. In order to address these challenges and ensure that FL is seamlessly integrated into the Industrial Metaverse, a system known as HFEDMS has been developed [15]. Using a ConvLSTM-based intrusion detection method that takes advantage of the periodic nature of network message IDs, federated AI-enabled in-vehicle network intrusion detection for the Internet of Vehicles is investigated [16]. TrisaFed is an asynchronous and federated learning mechanism that has three steps and is introduced in [17]. In addition, a one-of-a-kind digital twin (DT) driven architecture for the Industrial Internet of Things (IIoT) is capable of capturing the attributes of industrial devices for real-time analytics and intelligent decision-making [18]. This architecture faces challenges related to the protection of data privacy during the offloading stages. Context-sensitive federated deep reinforcement learning is one of the tools that the C-fDRL framework offers [19] in order to help mitigate these privacy concerns. A federated edge learning framework for industrial data processing [20] and an edge-centric approach to enhance traffic delivery [21] are two additional innovations. In conclusion, the model that has been presented details the state, action, and reward in DRL that are particular to the issue [19].

3. SYSTEM MODEL

We consider a model of a vehicular network that encompasses a 5G roadside unit, referred to as 5GRSU, and a multitude of vehicles (V) on the highway. The communications established within this network can be characterized as either V2I, indicating connections between 5GRSU and V, or V2V, denoting vehicular interactions. We hypothesize that there are N V2I(5GRSU−V) links and M V2Vcommunications. We can define the gains from the n^{th} $5G_R SU-V$ *link to the RSU, the* m^{th} V2V link to the 5G-RSU, and the intra m^{th} V2V link as $H_{n,B}$, $H_{m,B}$, and H_m correspondingly. Furthermore, the interference of

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 493-500 | e-ISSN: 2347-470X

the channel from the m^{th} V2Vtransmitter to the 5G-RSU, from the m^{th} V2V transmitter to the i^{th} V2V receiver, and from the n^{th} $5G_RSU-V$ link to the m^{th} V2V receiver can be delineated as $G_{m,B}$, $G_{m,i}$, and $G_{n,m}$, respectively.

The primary function of the V2V communication links is to disseminate essential safety alerts such as impending accidents or road status, with a maximum frequency ranging from10 to100Hz. Consequently, we can infer that the count of V2V links surpasses that of $5G_RSU - V$ links, implying $M > N$. Our model encompasses a vehicular network that includes a 5G roadside unit, denoted as $5G_RSU$, and multiple vehicles (V) on the roadway. The network communications can be classified as either V2I, indicating connections between $5G_RSU$ and V, or V2V, referring to vehicle-to-vehicle interactions. We postulate that there exist NV2I ($5G_RSU - V$) links and MV2V communication links. The gains for the n^{th} G_RSU-V *link to the RSU, the* mth V2V link to the 5G-RSU, and the intra m^{th} V2V link can be defined as $H_{n,B}$, $H_{m,B}$, and H_m respectively. Moreover, the interference of the channel from the m^{th} V2V transmitter to the5G- RSU, from the m^{th} V2V transmitter to the i^{th} V2V receiver, and from the $n^{th}5G_RSU - V$ *link to the* mth V2V receiver can be denoted as $G_{m,B}$, $G_{m,i}$, and $G_{n,m}$ correspondingly.

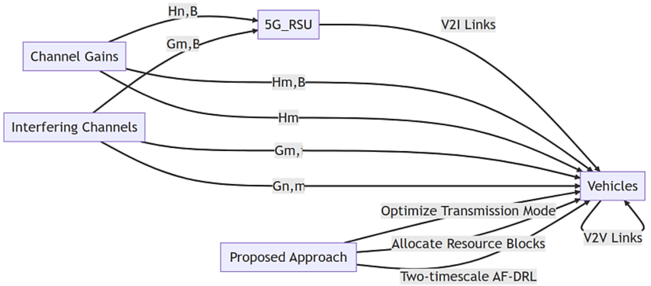

Figure 1. System model

V2V communication links primarily disseminate essential safety alerts, such as information on upcoming accidents or road conditions, with a maximum frequency spanning from10 to 100Hz.This suggests that the number of V2V links exceeds that of $5G_RSU - V$ links, thereby implying $M > N$.

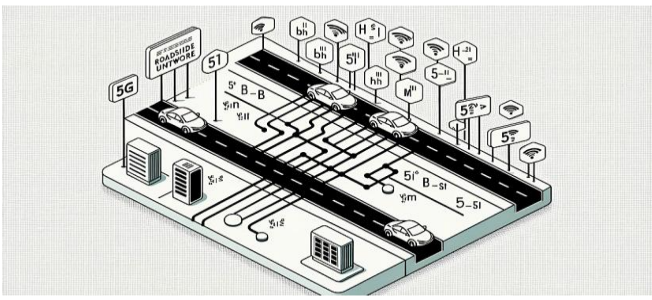

Figure 2. Detailing the vehicular network and the associated communication links

The strategy in focus deploys a deep reinforcement learning algorithm to tailor the selection of transmission mode and resource block distribution for V2V communication links. This algorithm engages a two-timescale asynchronous federated DRL approach. A macro-scale model is utilized for graph-centered vehicular clustering to form clusters of proximal vehicles on a larger timescale, while a micro-scale model trains vehicles within similar clusters by means of a robust global asynchronous federated deep reinforcement learning algorithm.

Although explicit mathematical equations are absent in the system model, channel gains and interference channels are denoted using symbols H_n,B, H_m,B, H_m, G_m,B, G_m,i, and G_n,m. There commended strategy relies on a deep reinforcement learning algorithm to fine-tune the choice of transmission mode and resource block allocation for V2V communication links.

The highlighted methodology deploys a deep reinforcement learning algorithm to refine the choice of transmission mode and the distribution of resource blocks for V2V communication links. This algorithm employs a two-timescale asynchronous federated DRL approach. A macro-scale model is utilized for graph-based vehicular clustering, thereby forming groups of nearby vehicles on a larger timescale. Conversely, a micro-scale model trains vehicles within homologous clusters using a robust, global, asynchronous federated deep reinforcement learning algorithm. The proposed methodology leverages a deep reinforcement learning algorithm to optimize the choice of transmission mode and the allocation of resource blocks for V2V communication links.

4. PROBLEM DEFINITION

This research deals with the challenge of optimizing both the choice of transmission mode and the distribution of resource blocks for a 5G-V2X communication scenario. The goal is to meet the latency and reliability requirements of V2V communication link sets. The problem is represented as a Markov decision process with the intent to maximize the overall potential expressed in channel capacity for vehicle-to-infrastructure users.

The constraints for this problem are as follows:

$$\text{LatencyReq}m \geq \frac{MsgSize_m}{Delay\text{max}} \forall m \in M \tag{1}$$

$$P(\text{OutageThr}_m \leq \text{OutageProbThr}) \leq P_0 \forall m \in M \tag{2}$$

$$\sum_{k \in K} P_{km} \leq P_{\text{peak}} \tag{3}$$

where:

- *LatencyReqm* denotes the latency requirement for the m^{th} V2Vcommunication link,
- $MsgSize_m$ represents the message size in bits,
- *Delay*max is the maximum tolerable latency or delay time in seconds,
- OutageThr$m$ refers to the outage threshold level for the m^{th} V2V communication link,
- OutageProb Thr indicates the acceptable outage probability with outage threshold level,
- P_0 is the reliability requirement,
- P_{km} signifies the transmission power for the k^{th} resource block of the m^{th} V2V communication link,
- K represents the set of resource blocks, and
- P_{peak} is the peak transmission power.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 493-500 | e-ISSN: 2347-470X

5. PROPOSED MODEL

In the current model, Asynchronous Federated Deep Reinforcement Learning (AFDRL) is combined with Federated Averaging to optimize the selection of transmission mode and allocation of resource blocks within a 5G-V2X communication scenario. Each individual agent, which serves as a representative of a V2V communication link, interacts autonomously with its environment in order to gain knowledge and adjust the parameters of its policy or value function. The server then aggregates the local updates in order to calculate the global update, which is then transmitted back to the agents. The preceding procedure is repeated until convergence is achieved. The purpose of the proposed model is to optimize the overall channel capacity of vehicle-to-infrastructure users, considering latency and reliability constraints into account. The proposed model is structured as a Markov decision process, and its primary objective is to maximize the total potential. In the context of federated learning, the proposed model effectively addresses the inherent challenges posed by non-IID data, asynchrony, and communication efficacy. This is accomplished through the use of Federated Averaging to aggregate the local updates. This methodology ensures the effectiveness of the global update, even in scenarios where data distributions fluctuate among agents, update times vary, and the model update size for each agent is substantial.

5.1. Asynchronous Federated Deep Reinforcement Learning

In Asynchronous Federated Deep Reinforcement Learning (AF-DRL), each agent i independently interacts with its environment to collect experience. This experience is represented as a tuple $(s_{i,t}, a_{i,t}, r_{i,t}, s_{i,t+1})$, where $s_{i,t}$ is the state, $a_{i,t}$ is the action, $r_{i,t}$ is the reward, and $s_{i,t+1}$ is the next state.

Each agent uses its experience to update its policy or value function parameters, denoted as $w_{i,t}$. The update is typically based on a DRL algorithm such as Q-learning or policy gradients. For example, in Q-learning, the update might look like this:

$$\Delta w_{i,t}=\alpha(r_{i,t}+\gamma \max Q(s_{i,t+1},a';w_{i,t})-Q(s_{i,t},a_{i,t};w_{i,t}))\nabla_w Q(s_{i,t},a_{i,t};w_{i,t}) \tag{4}$$

where α is the learning rate, γ is the discount factor, $Q(s,a;w)$ is the Q-value function parameterized by w, and the max is overall possible actions a'. The local updates are then averaged at the server to compute the global update, which is sent back to the agents. This process is repeated until convergence, resulting in an optimized selection of transmission mode and allocation of resource blocks that maximizes the aggregate potential while fulfilling the latency and reliability constraints. Mathematically, the local updates and the global update can be represented as follows:

$$\Delta w_{i,t} = \nabla J(w_{i,t}, e_{i,t}) \tag{5}$$

$$\Delta w_{\text{global},t} = \frac{1}{N}\sum_{i=1}^{N} \Delta w_{i,t} \tag{6}$$

Where,

- $J(w_{i,t}, e_{i,t})$ is the objective function for agent i,
- $\nabla J(w_{i,t}, e_{i,t})$ is the gradient of the objective function with respect to the parameters $w_{i,t}$,
- N is the total number of agents.

5.2. Complexities And Challenges In AF-DRL

The actual process of AF-DRL can be more complex, depending on the specifics of the DRL algorithm and the federated learning setup. Several issues need to be carefully managed:

- ***Non-IID Data:*** In many real-world scenarios, the data collected by each agent may not be identically and independently distributed (non-IID). This can affect the convergence and performance of the learning algorithm. Mathematically, if we denote the data distribution of agent i as P_i, then we have $P_i \neq P_j$ for $i \neq j$.
- ***Asynchrony:*** In an asynchronous setting, agents may not always be able to communicate with the server at the same time. This can lead to delays and inconsistencies in the global updates. If we denote the update time of agent i as t_i, then we have $t_i \neq t_j$ for $i \neq j$.
- ***Communication Efficiency:*** Transmitting the model updates between the agents and the server can consume significant network resources. Therefore, strategies to reduce the communication overhead, such as model compression or selective parameter updates, may be needed. If we denote the size of the model update from agent i as S_i, then the total communication cost can be represented as $\sum_{I=1}^{N} Si$ where N is the total number of agents.

Federated Averaging is a key component of Asynchronous Federated Deep Reinforcement Learning (AF-DRL) that helps to address the challenges of non-IID data, asynchrony, and communication efficiency.

6. HANDLING THE CHALLENGES ABOVE WITH FEDERATE AVERAGING

Federated Averaging is a strategy to aggregate the local updates from each agent in a way that mitigates the issues of non-IID data, asynchrony, and communication efficiency.

- ***Non-IID Data:*** Federated Averaging aggregates the local updates based on their quality, not just their quantity. This means that even if the data distributions P_i are different across agents, the global update can still be effective. The global update w_{t+1} is computed as:

$$w_{t+1} = \frac{1}{N}\sum_{i=1}^{N} w_{i,t} \tag{7}$$

Where $w_{i,t}$ is the local update from agent i at time t, and N is the total number of agents.

- **Asynchrony:** Federated Averaging allows each agent to send its local update whenever it's ready, instead of waiting for all agents to finish their updates. This means that even if the update times t_i is different across agents, the global update can still be computed in a timely manner.
- **Communication Efficiency:** Federated Averaging reduces the communication overhead by only transmitting the averaged update, not the individual updates. This means that even if the size of the model update Si is large

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 493-500 | e-ISSN: 2347-470X

for each agent, the total communication cost can still be kept low. The total communication cost is represented as $S_{total}=S_{avg}$, where S_{avg} is the size of the averaged update.

The role of Federated Averaging in handling the challenges in AF-DRL is explained with some mathematical notations. The non-IID data, asynchrony, and communication efficiency are handled by averaging the local updates, allowing asynchronous updates, and transmitting only the averaged update, respectively.

6.1. Applying AF-DRL With Federated Averaging

The problem is a mixed integer non-linear programming problem, which is difficult to solve using conventional approaches due to the combinatorial issues arising from binary variables and non-convexity of the constraints related to transmission power, channel capacity, latency, and reliability. To address these challenges, we propose using Asynchronous Federated Deep Reinforcement Learning (AF-DRL) with Federated Averaging. In this approach, each agent a (in this case, each V2V communication link) independently interacts with its environment E (the 5G-V2X communication scenario) to collect experience e_a and update its policy π_a or value function parameters θ_a. The local updates $\Delta\theta_a$ are then averaged at the server to compute the global update $\Delta\theta_{\text{global}}$, which is sent back to the agents. This process is repeated until convergence, resulting in an optimized selection of transmission mode and allocation of resource blocks that maximizes the aggregate potential while fulfilling the latency and reliability constraints. Mathematically, the local updates and the global update can be represented as follows:

$$\Delta\theta_a = \nabla J(\theta_a, e_a) \tag{8}$$

$$\Delta\theta_{\text{global}} = \frac{1}{A}\sum_{a=1}^{A} \Delta\,\theta_a \tag{9}$$

Where,

- $J(\theta_a, e_a)$ is the objective function for agent a,
- $\nabla J(\theta_a, e_a)$ is the gradient of the objective function with respect to the parameters θ_a,
- A is the total number of agents.

The procedure, called the algorithm is designed to make the choices, for transmission mode and resource block allocation in a 5GV2X communication scenario using Asynchronous Federated Deep Reinforcement Learning (AF DRL). The algorithm starts by setting up the parameters and enters a loop that continues until specific conditions are met. In each loop iteration every agent independently interacts with its surroundings to gather experience and update its policy or value function parameters. These local updates are then sent to the server, which combines them to calculate an update. The global update is then shared back with the agents. This process of updates and global averaging repeats until convergence is achieved resulting in an optimized selection of transmission mode and resource block allocation that maximizes potential while meeting latency and reliability requirements.

Algorithm 1: *Asynchronous Federated Deep Reinforcement Learning (AF-DRL) with Federated Averaging*

1. Initialize global model parameters w_0
2. **for** each round $t = 0, 1, 2, \ldots$ **do**
3. The servers ends w_t to all participating agents
4. **for** each agent i in parallel **do**
5. Agent i collects experience $(s_{i,t}, a_{i,t}, r_{i,t}, s_{i,t+1})$
6. Agent i computes local update $\Delta w_{i,t}$ using its DRL algorithm
7. **end for**
8. The server collects local updates $\Delta w_{i,t}$ from all participating agents
9. The server computes the global update : $w_{t+1} = w_t + \eta \sum_i \frac{n_i}{n} \Delta w_{i,t}$
10. **end for**

7. SIMULATION RESULTS

In this section we present the results of our simulations to evaluate the effectiveness of the Federated Averaging Asynchronous Deep Reinforcement Learning (FAA DRL) approach that we proposed. We thoroughly assessed the performance and reliability of AFDRL 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 AFDRL approach performs better, than methods across all evaluated metrics. This demonstrates its effectiveness, in optimizing resource allocation within Heterogeneous Cloud Radio Access Networks (H-CRANs).

Table 1. Simulation Parameters for Federated Averaging

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

The *table 2* presents a comparative analysis between two approaches in the field of deep reinforcement learning, namely Asynchronous Federated Deep Reinforcement Learning (AF-DRL) and Federated Averaging Asynchronous Federated Deep Reinforcement Learning (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.

Table 2. Comparison of AF-DRL and Fed Avg AF-DRL

Metric	AF-DRL	Fed AvgAF-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

The data presented in the table clearly indicates that the FedAvgAF-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 Heterogeneous Cloud Radio Access Networks (H-CRANs). The enhanced performance can be described to the proficient allocation of resources and streamlined data transmission facilitated by the federated averaging procedure.

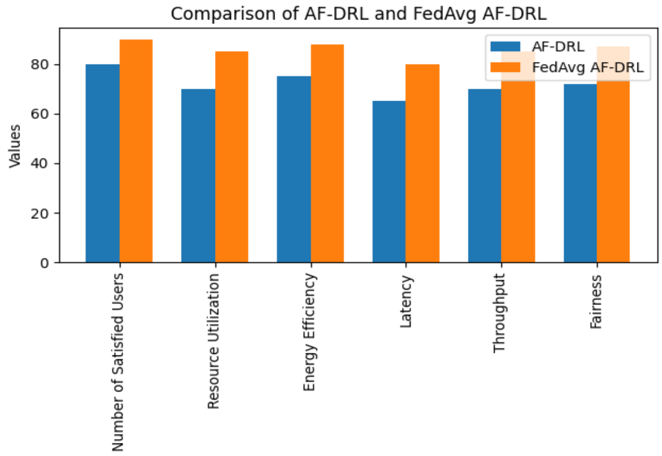

Figure 3. Comparison of AF-DRL and FedAvg AF-DRL

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 Heterogeneous Cloud Radio Access Networks (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.

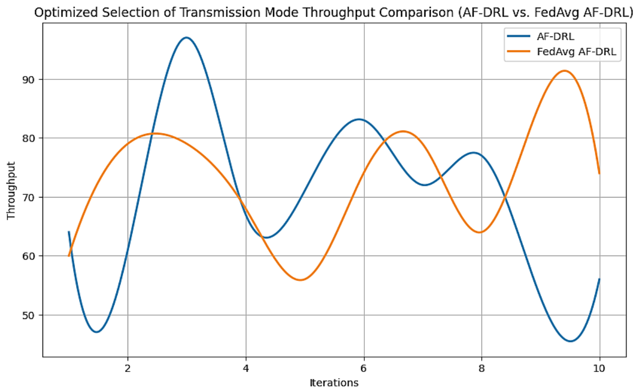

Figure 4. Throughput Comparison of AF-DRL and FedAvg AF-DRL with efficient Model selection

Figure 4 presents a line plot that illustrates the comparative analysis of throughput performance between two optimization methodologies, namely Asynchronous Fed- erated Deep Reinforcement Learning (AF-DRL) and Federated Averaging AF-DRL (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 the graph represents the achieved throughput values of each approach. Throughout the progression of the narrative, one can discern the variations in throughput measurements for both AF-DRL and FedAvg AF-DRL. Both methodologies demonstrate enhancements in throughput throughout the iterations. However, the presence of smooth edges on the lines suggests the utilization of a smoothing algorithm to enhance the visual aesthetics of the data representation. The juxtaposition of the two lines enables us to evaluate the efficacy of Federated Averaging in augmenting the throughput performance in comparison to the conventional AF-DRL approach. The utilization of this visualization facilitates comprehension of the performance of optimization techniques in relation to throughput, a critical metric for assessing the efficiency and efficacy of communication protocols in intricate network scenarios. The comparison graph in Figure5, referred to as the "Throughput" form, illustrates the performance of the "FedAvg AFDRL" and "AFDRL" approaches in terms of throughput across multiple iterations. Throughput is a metric that quantifies the speed at which data is transferred across a network.

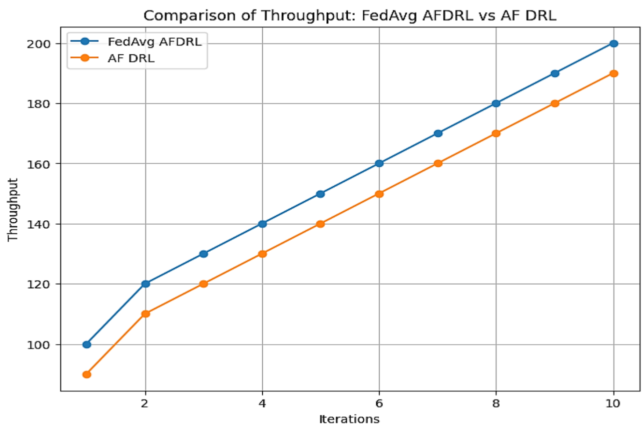

Figure 5. Throughput Comparison of AF-DRL and FedAvg AF-DRL with efficient Resource Allocation

As the number of iterations increases, there is an observed increase in throughput for both approaches. Nevertheless, it is worth noting that the "FedAvg AFDRL" algorithm consistently exhibits superior performance compared to the "AF DRL" algorithm in terms of through- put. This consistent out performance highlights the enhanced capability of "FedAvg AFDRL" in optimizing resource allocation and maximizing data transmission rates. In Figure 6 the comparison graph shows the latency values achieved by the "FedAvg AFDRL" 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 AFDRL" 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.

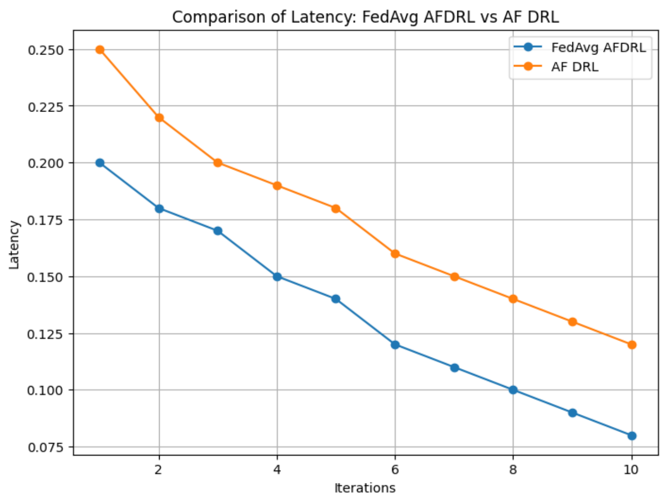

Figure 6. Latency Comparison of AF-DRL and FedAvg AF-DRL with efficient Resource Allocation

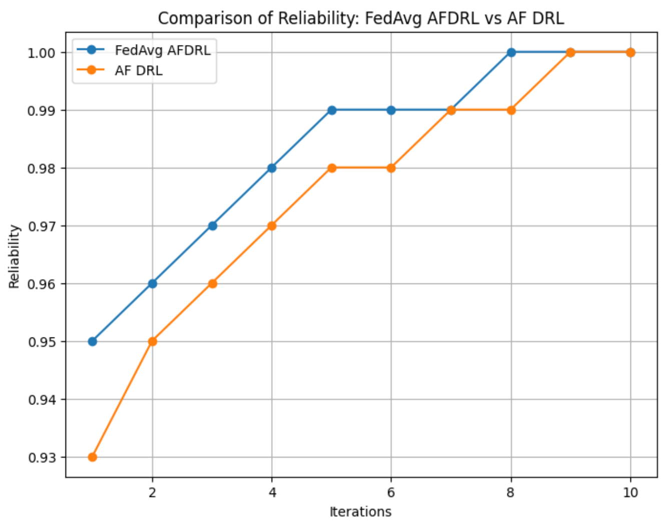

Figure 7. Reliability Comparison of AF-DRL and FedAvg AF-DRL with efficient Resource Allocation

Figure 7 shows a comparison graph that displays the reliability levels achieved by the "FedAvg AFDRL" and "AF DRL" approaches over iterations. Reliability in this context refers to the likelihood of achieving error uninterrupted data transmission. As we observe an increase, in the number of iterations both approaches demonstrate reliability. It's important to note that the "FedAvg AFDRL" algorithm consistently exhibits reliability levels compared to the "AF DRL" algorithm. This highlights its ability to maintain communication connections. In summary, FedAvg AFDRL outperforms AFDRL in H-CRANs, offering better throughput, lower latency, and higher reliability for D2D communication.

8. CONCLUSION

This paper presents an innovative solution for efficient resource allocation in Heterogeneous Cloud Radio Access Networks (H-CRANs). It proposes combining Asynchronous Federated Deep Reinforcement Learning (AF-DRL) with Federated Averaging. This method effectively tackles challenges like non-IID data, asynchrony, and communication inefficiencies in federated learning. AF-DRL enables agents to learn decentralized, autonomously adjusting policies. By iteratively aggregating local updates for global improvements, the system optimizes transmission mode selection and resource block allocation, maximizing channel capacity while meeting latency and reliability demands. Experimental results confirm AF-DRL's potential as a powerful approach for resource allocation in dynamic network environments.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

[1] Yunlong Lu; Xiaohong Huang; Ke Zhang; Sabita Maharjan; Yan Zhang; "Blockchain Empowered Asynchronous Federated Learning for Secure Data Sharing in Internet of Vehicles", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, 2020.

[2] Xinran Zhang; Mugen Peng; Shi Yan; Yaohua Sun; "Deep-Reinforcement- Learning-Based Mode Selection and Resource Allocation for Cellular V2X Communications", IEEE INTERNET OF THINGS JOURNAL,2020.

[3] Xinran Zhang; Mugen Peng; Shi Yan; Yaohua Sun; "Deep Reinforcement Learning Based Mode Selection and Resource Allocation for Cellular V2X Communications", ARXIV-CS.NI, 2020.

[4] H.Wang; Z. Kaplan; D. Niuand B. Li; "Optimizing Federated Learning on Non-IID Data with Reinforcement Learning", INFOCOM, 2020.

[5] Junjie Tan; Ying-Chang Liang; Lin Zhang; Gang Feng; "Deep Reinforcement Learning for Joint Channel Selection and Power Control in D2D Networks", IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, 2021.

[6] XueqingZhang; Yanwei Liu; JinxiaLiu; Antonios Argyriou; Yanni Han; "D2D- Assisted Federated Learning in Mobile Edge Computing Networks", 2021 IEEE WIRELESS COMMUNICATIONS AND NETWORKING CONFERENCE..., 2021.

[7] Yunlong Lu; Xiaohong Huang; Ke Zhang; Sabita Maharjan; Yan Zhang; "Communication-Efficient Federated Learning and Permissioned Blockchain for Digital Twin Edge Networks", IEEE INTERNET OF THINGS JOURNAL,2021.

[8] Xiaofei Wang; Ruibin Li;Chenyang Wang; Xiuhua Li; Tarik Taleb; VictorC.M.Leung; "Attention-Weighted Federated Deep Reinforcement Learning for Device-to-Device Assisted Heterogeneous Collaborative Edge Caching",IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS,2021.

[9] Peiying Zhang; Chao Wang; Chunxiao Jiang; Zhu Han;"Deep Reinforcement Learning Assisted Federated Learning Algorithm for Data Management of IIoT", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS,2021.

[10] Rong Yu; Peichun Li; "Toward Resource-Efficient Federated Learning in Mobile Edge Computing", IEEE NETWORK, 2021.

[11] Qiong Wu; Yu Zhao; Qiang Fan; Pingyi Fan; Jiangzhou Wang; Cui Zhang; "Mobility-Aware Cooperative Cachingin Vehicular Edge Computing Based on Asynchronous Federated and Deep Reinforcement Learning", ARXIV-CS.DC, 2022.

[12] WeiYang; WeiXiang; YuanYang; PengCheng; "Optimizing Federated Learning with Deep Reinforcement Learning for Digital Twin Empowered Industrial IoT", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS,2023.

[13] QiongWu; SiyuanWang; PingyiFan; QiangFan; "Deep Reinforcement Learning Based Vehicle Selection for Asynchronous Federated Learning Enabled Vehicular Edge Computing", ARXIV-CS.LG, 2023.

[14] Farhad Rezazadeh; Lanfranco Zanzi; Francesco Devoti; Hatim Chergui; Xavier Costa-Perez; Christos Verikoukis: "On the Specialization of FDRL Agents for Scalable and Distributed 6G RAN Slicing Orchestration", ARXIV-CS.NI, 2022.

[15] Shenglai Zeng; ZonghangLi; HongfangYu; Zhihao Zhang; Long Luo; BoLi; DusitNiyato; "HFedMS: Heterogeneous Federated Learning with Memorable Data Semantics in Industrial Metaverse", ARXIV-CS. LG,2022.

[16] Jianfeng Yang; Jianling Hu; Tianqi Yu; "Federated AI-Enabled In-Vehicle Network Intrusion Detection for Internet of Vehicles", ELECTRONICS,2022.

[17] Linlin You; Sheng Liu; Yi Chang; C.Yuen; "ATriple-Step Asynchronous Federated Learning Mechanism for Client Activation, Interaction Optimization, and Aggregation Enhancement, IEEE INTERNET OF THINGS JOURNAL,2022.

[18] Qiong Wu; Xu Chen; Tao Ouyang; Zhi Zhou; Xiaoxi Zhang; Shusen Yang; Junshan Zhang; "HiFlash: Communication-Efficient Hierarchical Federated Learning with AdaptiveStalenessControlandHeterogeneity-awareClient-EdgeAssociation", ARXIV-CS.NI, 2023.

[19] WeiYang; WeiXiang; YuanYang; PengCheng; "Optimizing Federated Learning with Deep Reinforcement Learning for Digital Twin Empowered Industrial IoT", IEEETRANSACTIONS ON INDUSTRIAL INFORMATICS,2023.

[20] YangXu; MdZakirul AlamBhuiyan; TianWang; XiaokangZhou; Ashutosh KumarSingh; "C-FDRL:Context-Aware Privacy-Preserving Offloading Through Federated Deep Reinforcement Learning in Cloud-Enabled IoT", IEEE TRANS- ACTIONS ON INDUSTRIAL INFORMATICS, 2023.

[21] BinJiang; Jianqiang Li; Huihui Wang; H.Song; "Privacy-Preserving Federated Learning for Industrial Edge Computing Via Hybrid Differential Privacy and Adaptive Compression", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, 2023.

[22] Abdallah Jarwan; M. Ibnkahla; "Edge-Based Federated Deep Reinforcement Learning for IoT Traffic Management", IEEE INTERNET OF THINGS JOURNAL, 2023.

[23] Qiong Wu; Siyuan Wang; Pingyi Fan; Qiang Fan; "Deep Reinforcement Learning Based Vehicle Selection for Asynchronous Federated Learning Enabled Vehicular Edge Computing", ARXIV-CS.LG, 2023.

© 2025 by Abishek Malviya and Sudhakar Pandey. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).