# Federated Reinforcement Learning for Load Balancing in O-RAN 5G and Beyond Networks

Mohammed Imad Aal-Nouman, Sarmad M. Hadi and Haider A. H. Alobaidy

***Abstract*—The intelligent distributed control is a critical factor in next-generation mobile networks for connecting a massive number of devices and handling diverse services. The Open Radio Access Network (O-RAN) has recently gained significant traction as a flexible architecture based on standards that enables disaggregated, programmable network components. However, because O-RAN is inherently distributed, it creates an imbalance in traffic load when user mobility and traffic are dynamic. This paper proposes a Federated Reinforcement Learning (FedRL) framework to help load balance O-RAN units (O-RUs). The framework enables each O-RU to train a local reinforcement learning agent using real-time information on signal quality, mobility dynamics, and load conditions. The trained agents periodically share their learned information via federated averaging of their Q-tables at the Near-Real-Time RAN Intelligent Controller (Near-RT RIC), enabling collaborative learning without requiring centralized data collection. Using a synthetic dataset obtained from MATLAB-generated simulations of a mobile network of 500 m x 500 m, the proposed system is trained and tested. It simulates realistic user movement and radio conditions between multiple cells. In the experiments, FedRL reduces load imbalance by over 50% compared to rule-based schemes and by around 25% compared to centralized RL. It also achieves fairness values exceeding 0.95, increases throughput by more than 15%, increases handover success rates by around 13%, and reduces ping-pong events by 30-40 % compared to the baselines.**

***Index Terms*—O-RAN, Load balance, B5G, Federated Reinforcement Learning, Handover.**

## I. INTRODUCTION

Mobile communication has advanced from the early analog first-generation (1G) systems to the advanced capabilities of fifth-generation (5G), with sixth-generation (6G) on the horizon. These changes have led to significant shifts in how networks are designed and operated [1]. Over time, traditional Radio Access Network (RAN) designs have struggled to keep up with rising demands for flexibility, scalability, and cost efficiency. To overcome these challenges, the Open-RAN (O-RAN) was introduced as a new architectural model. O-RAN focuses on openness, intelligent control, and interoperability [2].

Formed in 2018, the O-RAN Alliance launched the Open RAN initiative as a new approach to rethink traditional RAN designs. The purpose of this is to disassemble the previous 'system structure', which is still held intact by standardized interfaces and open communication protocols [3]. O-RAN, through the separation of hardware and software as well as interoperability of equipment from different vendors, creates an environment for innovation that reduces dependence on a single supplier and enables cheaper network deployment [4]. The Near-Real-Time RAN Intelligent Controller (RT RIC) and Service Management and Orchestration (SMO) platform are key components introduced by this architecture that can improve how the RAN is programmed, managed, and optimized [5]. The O-RAN architecture consists of a radio domain and a management domain architecture. The radio architecture contains four significant elements. The O-RU is responsible for low-level physical (PHY) layer functions and radio frequency (RF) signals. The Open Distributed Unit (O-DU) handles MAC, RLC and High-PHY layer-related tasks. The O-CU-CP handles control signaling, while the O-CU-UP is for user data traffic [6]. In operational terms, the Service Management and Orchestration platform is responsible for controlling and optimizing the network. The SMO's Non-Real-Time RIC (Non-RT RIC) leverages the A1 interface to facilitate long-term policy enforcement and the training of Artificial Intelligence (AI) and Machine Learning (ML) models. Simultaneously, the Near-Real-Time RIC (Near-RT RIC) facilitates faster decision-making by hosting dynamic xApps that interact directly with the O-DU and O-CU for optimization purposes [4].

The O-RAN is driven by several intersecting demands. The growth of mobile data is accelerating rapidly, driven by new applications such as augmented reality, ultra-high-definition video streaming, and the IIoT network. Thus, demand for low latency and intelligent resource allocation is growing [7]. Operators are under mounting pressure to achieve lower operational costs while retaining the same level of quality of service (QoS). Also, the proprietary, closed designs of conventional RAN platforms have historically hampered innovation and limited flexible large-scale deployments [8]. O-RAN tackles these issues by enabling intelligence based on AI and ML through open interfaces while supporting adaptation and optimization at both the local and network-wide level [9]. Despite the architecture, O-RAN's failure to load balance is an issue. When the O-RUs traffic is imbalanced, it can lead to poor user experience, increased latency, and inefficient spectrum use. Most load-balancing techniques leverage either a handover policy based on signal strength or a fixed threshold. These do not respond to any real-time traffic or mobility changes [10-13]. Centralized control approaches are more coordinated, but they are prone to latency and scalability issues in large-scale heterogeneous networks [14]. While recent works explore the use of

Mohammed Imad Aal-Nouman and Haider A. H. Alobaidy are faculty members at the Department of Information and Communications Engineering, College of Information Engineering, AL-Nahrain University, Baghdad, Iraq (e-mail: m.aalnouman@nahrainuniv.edu.iq, and Haideralobaidy@nahrainuniv.edu.iq )

Sarmad M. Hadi is a faculty member at the Department of Computer Networks Engineering, College of Information Engineering, AL-Nahrain University, Baghdad, Iraq (e-mail: sarmad@nahrainuniv.edu.iq)

DOI: 10.36244/ICJ.2026.2.6

reinforcement learning and federated learning for resource optimization in O-RAN, most such approaches remain limited. Centralized RL techniques face limitations in latency and scalability in large deployments [15], while multi-agent RL solutions allow local adaptation but typically lack coordination, leading to inconsistent results [16]. Existing federated learning approaches have focused on slice-level orchestration or static allocation [17-19], but not on those that account for mobility-driven handover events. More recent papers also point out that the lack of mobility adaptation ultimately limits the efficiency of FL-based schemes under realistic O-RAN conditions [20, 21]. The identified gaps demonstrate the need for a mobility-aware federated load-balancing mechanism that guarantees scalability, privacy and responsiveness in a real-world O-RAN scenario. This paper analyzes the dynamic and intelligent load balancing challenge in an O-RAN-based 5G Network and beyond.

Federated Reinforcement Learning (FedRL) framework offers a formulation in which distributed agents associated with an O-RU make local decisions on handovers and load distribution. These decisions depend on the current states within the network. Federated learning has agents synchronizing their models periodically using the Near-RT RIC, allowing learning without exchanging data. The purpose of this research is to enhance load balancing performance while ensuring scalability, adaptability, and user privacy. We use a synthetic mobile network scenario to generate a detailed dataset. Each record in the dataset contains information on user mobility, radio signal quality, and local cell load. FedRL is assessed using basic metrics such as Max-to-Min Load Difference, Handover Failure Rate, and Ping-Pong Occurrences, compared with centralized and multi-agent RL. Most centralized RL approaches assume global state observability (i.e., all agents are fully observable) and that centralized control involves only a single-point control decision. Also, classic multi-agent RL schemes operate independently without coordination. However, the proposed FedRL-O-RAN framework proposes coordinated yet decentralized learning for mobility-aware handover optimization. The main contribution of this work is threefold. First, rather than directly selecting the target cell to handover, the approach formulates the handover control as A3-offset learning. Second, the basic building block of this approach is tabular Q-learning implemented at each O-RU to keep computation simple and easy to understand. Finally, the learning mechanism employs federated Q-table aggregation at the Near-RT RIC to achieve consistency in their handover policies without disclosing any user data. This combination enables scalable, privacy-preserving, mobility-aware load balancing within the operational constraints of O-RAN. The principal contributions of this paper are summarized as follows:

- Development of distributed Q-learning agents, one per O-RU, synchronized by means of federated averaging through the Near-RT RIC.
- Inclusion of signal quality, cell load, and user mobility context in the state inputs that make the decision.
- The development of the MATLAB multi-cell simulation framework, which includes interference, mobility, and dynamic traffic and is evaluated against rule-based, centralized RL, and MARL baselines.

The rest of the paper is organized as follows. Related work is discussed in Section II; Section III presents the system model architecture and methodology; Section IV presents all the simulation parameters and metrics; Section V highlights the main findings and results; and finally, conclusions are drawn in Section VI.

## II. Related Works

Intelligent load balancing and handover optimization are playing a central role in O-RAN and C-RAN performance, especially with the integration of reinforcement learning and federated learning. As network densification increases and mobility becomes more dynamic, intelligent decision-making mechanisms are increasingly required to maintain fairness, throughput stability, and handover reliability.

From a centralized optimization perspective, Pamuklu et al. (2021) present an RL-based strategy for adaptively managing function splitting between CU and DU components in disaggregated O-RAN settings, aiming to minimize operational costs in energy-constrained scenarios [23]. Ali (2025) showed that when using centralized learning strategies in C RAN environments, it improves energy efficiency and spectral efficiency and reduces BBU centralized deployments' handover failure rate [24]. Researchers have studied DDPG-based reinforcement learning techniques for dynamic optimization of handover parameters in dense 5G scenarios, which can minimize ping-pong effect and handover failure rate [25]. Gupta and others in (2021) implemented joint load balancing and handover optimization with deep RL in a multi-band LTE/5G network. They showed improved stability and throughput for the network [26]. Centralized methods make use of global observability and coordinated decision making, which allows optimization objectives for the whole system to be included in learning directly. Nevertheless, difficulties when it comes to scaling and signaling could arise for these techniques in large O-RANs, especially when rapid mobility dynamics require low-latency local adaptation.

Orhan et al. (2021) presented a Graph Neural Network-based RL xApp for intelligent user association of the O RAN RIC to move towards distributed and multi-agent learning approaches. The RL xApp from Orhan et al. manages to achieve 20 to 45% gains in load balancing in comparison to greedy baselines [15]. In the context of O-RAN, Lai et al. (2023) designed mmLBRA, a multi-agent, multi-armed bandit-based scheme which effectively balances load among O-RUs [16]. Several additional works [16, 18, 26, 27] propose multi-agent systems or hierarchical FL structures for distributing decision-making in O-RAN. When each entity learns from its local environment, then the overall entity can respond to heterogeneous traffic and centralized bottlenecks can be reduced. However, many do not perform a structured policy synchronization across their O-RUs, or rely on centralized inference components. This may detract from achieving consistent global fairness, and lead to unstable performance as traffic patterns change rapidly across neighboring cells.

In O-RAN, collaborative optimization that protects privacy may be possible with federated learning. Zhang et al. (2022) designed algorithms for slice-aware xApps coordination based on federated deep reinforcement learning (FedDRL) [17]. A federated deep Q-learning xApp for 5G load balancing was developed by Lin et al (2024), which improves throughput and protects privacy [28]. The authors of reference [29], Cao et al., designed a federated DQN access control mechanism in O RAN by using UCB-based UE selection to maximize long-term throughput and minimize unnecessary handovers. Ahmed and colleagues developed a framework that utilizes federated DRL for distributed slice management to optimize O-RAN environments. Zhao et al. also proposed a multi-layer collaborative federated learning (MLCFL) scenario framework for next-generation (6G) O-RAN coordination. Extensions of this line of research include Elastic Federated Learning (EFL) which emphasizes federated RL for spectrum sharing in NTN deployments. The these work show that a federated coordination can trade-off between distributing the intelligence and increasing the performance of the system globally without interfering with the data locality. Nonetheless, several federated methods are primarily tailored for slice management, access control, or spectrum allocation, as opposed to fine-grained, mobility-aware, user-level handover optimization under dynamic load fluctuations.

Frameworks that are hybrid and predictive have also been studied to improve adaptability. In 2023, Kavehmadavani et al. proposed a traffic steering solution using LSTM for edge-centric O-RAN deployments [33] that will help in load steering based on traffic prediction.

Singh et al. (2023) introduces Cascade RL (CaRL) that leverages state-space decomposition across O DU layers to balance QoS trade-offs [34]. Semiari et al. (2025) proposed a graph-RL load balancing mechanism for QoS-guaranteed O RAN slices, allowing achieving consistent resource fairness under diverse traffic mixes [35]. These approaches may enhance both orchestration and predictive capabilities, but they often abstract away from direct control of mobility-driven handover parameters at the O-RU level. Instead, they focus on other coordination objectives at higher layers.

While RL and FL have been used in the past for load balancing (LB) and handover (HO) optimization in O-RAN and C-RAN, a number of papers focus on global coordination, slice level allocation, static functional distribution etc. A constant limitation across these works is the limited introduction of real-time user-level mobility adaptation to federated or distributed decision loops. A deployment with dynamic conditions in the speed of users, signal value, localized congestion and the like can lead to unfairness or ping pong occurrences if a mobility-aware handover control is not used.

Pamuklu et al. [23], for instance, work on dynamic function splitting without user mobility pattern or federated coordination. The authors Orhan et al [15] improve load fairness through RL-driven user association but do not consider explicit model-specific handover design nor distributed model synchronization. Lin and coworkers [28] propose a federated DQN framework. Their experiments, however, do not include dynamic handover logic nor do they use aggressive load conditions. Likewise, Zhang et al. [17] allocate resources across network slices, but not with any granular mobility-aware handover decision-making, while Cao et al. [29] access control, not adaptive handover parameter optimization.

FedRL-O-RAN is a decentralized and user-centric reinforcement learning framework for dynamic handover optimization unlike the previous work. Every O-RU integrates a distributed Q-learning agent that leverages radio signal readings, user mobility patterns, and local cell load to formulate localized decisions. Through the near-RT RIC, periodic model synchronization occurs through federated averaging, which aligns all policies in the system without raw user data sharing. The proposed framework tackles a key limitation of previous works [15, 17, 28] by addressing the challenges of centralized global optimization and fully independent multi-agent control. It explicitly integrates mobility context into distributed learning with federated synchronization.

## III. METHODOLOGY

Overall, although prior works have made contributions through architectural novelties or specialized learning schemes, FedRL-O-RAN provides a more holistic, scalable and privacy-aware solution. This study directly integrates user mobility into federated coordination in a realistic O-RAN deployment which bridges important research gaps and offers a practical approach for intelligent handover management and load balancing in 5G and beyond the generations of networks.

The issue of load balancing and handover control in O-RAN was formulated as a Markov Decision Process (MDP), with each Open Radio Unit (O-RU) acting as an independent reinforcement learning agent. Rather than directly choosing a target cell, each agent learns to modify the A3 handover offset. In this way, it is able to control handover aggressiveness, affecting load distribution.

In each of O-RU i's decision steps, O-RU i observes a locally constructed state that incorporates its own current load, the average SINR of its connected users and their average mobility speed. Consequently, the state of O-RU i is defined as:

$$s_i(t) = \left[\rho_i(t), \overline{\mathrm{SINR}}_i(t), \bar{v}_i(t)\right] \tag{1}$$

Where $\rho_i(t)$ denotes the current load of O-RU $i, \overline{\mathrm{SINR}}_i(t)$ represents the average SINR of the UEs served by O-RU $i$, and $\bar{v}_i(t)$ denotes the average mobility speed of its connected users. Together, these variables describe the load and mobility conditions within the cell. Since all state information is derived from locally available measurements, the representation remains compatible with the distributed nature of O-RAN operation.

The continuous state variables were discretized before learning because tabular Q-learning requires a finite state space. The local load, average SINR, and average mobility speed were mapped into discrete state categories, enabling the agent to maintain a Q-table while preserving the information required for handover decision-making. This representation is compatible with the distributed operation of O-RAN.

From the available action set, the agent selects an action according to the observed state.

$$a_i(t) \in \{-2,\ 0,\ 2\} \tag{2}$$

The value of the A3 handover offset controls the trigger margin corresponding to these actions, thereby influencing handover decisions. A handover occurs from a serving O-RU to a candidate O-RU only if the RSRP difference between the best candidate and the serving cell exceeds the configured A3 offset. When the offset is adjusted, the learning agent indirectly controls the aggressiveness of handover and load distribution.

After taking an action, the agent receives a reward that reflects the impact of the selected A3 offset on local load conditions and handover stability. To ensure that decisions rely only on local information, the reward is computed at each O-RU using only locally available information rather than network-wide load-balancing metrics. The reward of O-RU i at time t is defined as:

$$r_i(t) = \Delta L_i^{\text{local}}(t) + \beta_1 I_{\text{success}} - \beta_2 I_{\text{failure}} - \beta_3 I_{\text{ping-pong}} \tag{3}$$

where $r_i(t)$ denotes the reward received by O-RU $i$ at time step $t$, and $\Delta L_i^{\text{local}}(t)$ represents the improvement in the local load condition of O-RU i resulting from the selected action. The binary indicators $I_{\text{success}}$, $I_{\text{failure}}$, and $I_{\text{ping-pong}}$ correspond to successful handovers, failed handovers, and ping-pong handover events, respectively. The weighting coefficients $\beta_1, \beta_2$, and $\beta_3$ control the contribution of each component to the overall reward.

This reward formulation guides the agent toward actions that improve local load distribution while maintaining handover stability. The reward increases following successful handovers, while failed handovers and ping-pong events reduce it.

Each O-RU employs a distributed reinforcement learning agent that uses tabular Q-learning to learn state-action optimal mapping. The Q-values are updated at each time step according to the standard temporal-difference update rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \tag{4}$$

Where $\alpha$ is the learning rate, $\gamma$ is the discount factor, $r$ is the observed reward, and $\max_{a'} Q(s',a')$ represents the maximum estimated future value over all possible actions. This update gradually refines the Q-table maintained locally at each O-RU, enabling the agent to learn an effective handover control policy over time.

To enable collaboration without exposing raw user data, the agents engage in a Federated Averaging (FedAvg) process. After a set number of local training episodes, each O-RU sends its local Q-table $Q_i$ to the Near-Real-Time RIC, which computes the global Q-table:

$$Q_{\text{global}} = \frac{1}{N}\sum_{i=1}^{N} Q_i \tag{5}$$

The collected Q-table is then sent to all O-RUs, reinitializing their local agents, so that all O-RUs benefit from each other's experiences. This speeds up learning network. It also aligns your policies and protects users. Thus, FedRL-O-RAN will function as a coordinated/context-aware decision system, which can enable load balancing and stable handovers in dynamic environments. This framework facilitates independent O-RUs' decision-making by leveraging federated learning to jointly enhance the optimization of policies without requiring central data collection. The overall system model is based on the O-RAN architecture, an integration of FedRL to enable intelligence load balancing for 5G and beyond. The scheme comprises multiple O-RUs, a Near-Real-Time RIC, and centralized orchestration by the Service Management and Orchestration layer, as presented in Algorithm 1.

**Algorithm 1: Proposed FedRL-O-RAN**

```
Initialize local Q-table Qi for each O-RU_i
FOR round = 1 to R:
  FOR each O-RU_i in parallel:
    FOR episode = 1 to E:
      FOR each time t:
        Observe state s_i
        Select action a_t using ε-greedy policy
        Apply a_i(t) as the A3 offset for handover triggering
        Execute association update and observe next state s_i(t+1)
    Compute local reward r_i(t) using local load improvement,
handover outcome, and ping-pong indicator
        Update Q_i using Q-learning Eq(4)
        END FOR
     END FOR
    Upload Q_i to Near-RT RIC
  END FOR
  Q_global = (1/N) Σ_i Q_i       # FedAvg aggregation
  Broadcast Q_global and update Q_i = Q_global for all O-RUs
END FOR
  Return {Q_i}, Q_global
```

In this approach, every O-RU creates a local Q-table capable of selecting an A3 offset. The Near-RT RIC uses FedAvg to aggregate Q-tables and shares global table to synchronize policy without UE data sharing. The simulation unfolds in discrete time steps of 1 second. At every interval, the mobility positions of users are updated first. The local state, si(t), for each O-RU, i, is computed from the aggregated statistics of the UEs served by that O-RU. The A3 offset that is chosen will be activated, while handover decisions will be executed based on the trigger condition. When updates for all associations are completed for the current time step, the loads of all O-RUs are updated, and the reward ri(t) is computed. Subsequently, a local update of the Q-table is performed at each O-RU. It ensures every agent's state changes at the same time.

## IV. SIMULATION SETUP

This part discusses the scenario, settings, and reference approaches used to assess FedRL-O-RAN. This paper outlines the simulation environment, provides details of data generation, the performance metrics utilized, and the baseline methods against which the proposed approach is compared.

### *A. Simulation Environment*

The implementation of the suggested FedRL-O-RAN framework and its performance were evaluated in a MATLAB-based simulation environment that emulates the operation of an

O-RAN-enabled mobile network. The simulated network consists of seven O-RUs uniformly and randomly distributed throughout the 500 m × 500 m service area. All O-RUs use the same settings for hardware and power transmission to achieve uniform coverage characteristics. Because this work primarily focuses on modeling load balancing and mobility-aware handover optimization, dynamic power consumption modeling is outside its scope.

The simulation takes place in discrete time with a fixed sampling period of 1 second. At every time step, all O-RUs observe their local state information and simultaneously execute their decision updates according to the synchronization scheme.

Every O-RU acts as an independent reinforcement learning agent utilizing tabular Q-learning. The learning configuration uses a discount factor of $\gamma = 0.95$, a learning rate of $\alpha = 0.001$ and an ε-greedy exploration policy that decays from 1.0 to 0.1.

Through 50 communication rounds, federated learning is executed. After 10 local training episodes, each O-RU sends its local Q-table to the Near-RT RIC each round. To produce a global Q-table, the RIC combines the received Q-tables by using Federated Averaging (FedAvg), which is then sent back to all O-RUs for the next round. Each O-RU will use local experience and knowledge along with the network-wide experience.

All dataset used during training/testing is generated in the simulation environment itself. Every record represents the overall state of an O-RU at a specific timestep, including its load, the average SINR of served users, and the average user mobility speed. The label reflects the chosen A3 offset action. The reward is computed based on load balance improvement and handover outcome. Table 1 presents the features of the state; the definition of actions and the components of the reward.

TABLE 1.
STATE FEATURES, ACTION DEFINITION, AND REWARD COMPONENTS

| Category | Name | Description |
|---|---|---|
| State Features | Local load $\rho_i(t)$ | Number of UEs currently connected to O-RU $i$ |
| | Average SINR $\overline{\mathrm{SINR}}_i(t)$ | Mean SINR of UEs served by O-RU $i$ |
| | Average speed $\bar{v}_i(t)$ | Mean mobility speed of UEs connected to O-RU $i$ |
| Action (Label) | A3 offset $a_i(t) \in \{-2,0,2\}$ | Selected handover trigger margin controlling handover aggressiveness |
| Reward Components | $\Delta L_i^{\mathrm{local}}(t)$ | Local load-improvement term after the selected action |
| | *I_success, I_failure, I_ping-pong* | Indicators reflecting successful handover, failed handover, and ping-pong occurrence |

The simulation considers a dense small-cell O-RAN deployment covering an area of 500 m × 500 m with seven O-RUs. To evaluate scalability, additional deployment scenarios were examined by varying the coverage area and the number of O-RUs. The objective was to determine whether the proposed FedRL framework can maintain its load-balancing and mobility-management performance as the network size and density increase. The scalability scenarios considered in this study are summarized in Table 2.

Scenario S2 examines the impact of increasing O-RU density within the original deployment area. Scenario S3 evaluates the effect of expanding the coverage area while keeping the number of O-RUs unchanged. Scenario S4 combines both factors.

TABLE 2
SCALABILITY SCENARIOS USED IN THE EVALUATION

| Scenario | Area (m × m) | O-RUs |
|---|---|---|
| S1 | 500 × 500 | 7 |
| S2 | 500 × 500 | 14 |
| S3 | 1000 × 1000 | 7 |
| S4 | 1000 × 1000 | 14 |

## B. Learning Algorithms

In order to assess the performance of the proposed FedRL-O-RAN framework, performance comparison is conducted against three representative baseline methods, which correspond to a different philosophy for load balancing and handover decision of O-RAN networks. These baselines incorporate both learning-free heuristics and learning-based techniques for a fair and comprehensive evaluation.

Rule-Based Load Balancing with RSRP (rLBRA): A heuristic approach whereby each UE connects to the O-RU with the highest Reference Signal Received Power (RSRP). Despite being easy and inexpensive, load imbalance often occurs due to some signal strong cells becoming congested while others remain under-loaded.

Centralized Reinforcement Learning (Centralized RL): A unified global entity situated at the Near-Real-Time RIC is responsible for making handover and load balancing decisions for all O-RUs. It relies on aggregated network state information to facilitate these decisions. This centralization allows for consideration of global conditions by the agent but limits its response to rapid local changes and increases signaling overhead.

Multi-Agent Reinforcement Learning (MARL): Every O-RU is an RL agent acting independently based on local observations without coordination. Such design is effective at local phenomena but does not result in an overall balanced loading on the network as a whole.

The choice of these baselines is based on popular strategies from the literature, which allow direct performance comparisons of different decision-making paradigms, including static heuristics, coordinated, and uncoordinated agent learning-based models.

## C. Performance Metrics

Following the description of the simulation setup and the evaluated algorithms, the next step is to define the performance metrics that capture multiple aspects of network performance.

The first metric, Max-to-Min Load Difference, is used to quantify the imbalance in UE distribution across O-RUs, Let:

$\mathcal{B} = \{b_1, b_2, \dots, b_N\}$ : the set of O-RUs (base stations), $\mathcal{U}(t) = \{u_1(t), u_2(t), \dots, u_M(t)\}$ : the set of active UEs (users) at time $t$, $A_j(t) \in \mathcal{B}$: the serving O-RU for the user $u_j$ at time $t$, $L_i(t)$ : the load (number of connected users) on O-RU $b_i$ at time $t$.

We define load per O-RU as:

$$L_i(t) = \sum_{j=1}^{M} 1_{\{A_j(t)=b_i\}} \quad (6)$$

Where $1_{\{A_j(t)=b_i\}}$ is the indicator function:

Equals 1 if user $j$ is connected to $\mathrm{O-RU}b_i$,

Equals 0 otherwise.

Form the Load Vector. Construct the load vector at time $t$ :

$$\mathbf{L}(t)=[L_1(t),L_2(t),\ldots,L_N(t)] \tag{7}$$

This vector represents the instantaneous load distribution across all O-RUs. Then, the maximum and the minimum load for t time will be:

$$L_{\max}(t)=\max_{i\in\{1,\ldots,N\}} L_i(t) \tag{8}$$

$$L_{\min}(t)=\min_{i\in\{1,\ldots,N\}} L_i(t) \tag{9}$$

The instantaneous Max-to-Min Load Difference is given by:

$$\Delta_L(t)=L_{\max}(t)-L_{\min}(t) \tag{10}$$

This value captures the degree of imbalance in the user distribution at a specific moment. To evaluate performance over the entire simulation period, the average load difference is calculated as

$$\bar{\Delta}_L=\frac{1}{T}\sum_{t=1}^{T}\Delta_L(t) \tag{11}$$

While the Max-to-Min Load Difference focuses on extreme values, the Load Standard Deviation captures the overall variability in UE distribution across O-RUs. Let $L_i(t)$ denote the number of UEs connected to O-RU $i$ at time step $t$, and let $\bar{L}(t)$ be the mean load across all $N$ O-RUs at that time step, computed as

$$\bar{L}(t)=\frac{1}{N}\sum_{i=1}^{N}L_i(t) \tag{12}$$

The load variance at time $t$ is then obtained by

$$\sigma_L^2(t)=\frac{1}{N}\sum_{i=1}^{N}(L_i(t)-\bar{L}(t))^2 \tag{13}$$

and the load standard deviation is given by

$$\sigma_L(t)=\sqrt{\sigma_L^2(t)} \tag{14}$$

To capture the overall behavior over the simulation duration, the average standard deviation is calculated as

$$\overline{\sigma_L}=\frac{1}{T}\sum_{t=1}^{T}\sigma_L(t) \tag{15}$$

In which T is the Total Number of time step. The lower values of $\overline{\sigma_L}$ mean that the loads are better balanced among the O-RUs as opposed to higher values indicating greater imbalance. This metric is a statistical view of variability and therefore provides additional information to the Max-to-Min Load Difference that uses extremes of a distribution.

Using the first two load distribution measures, Jain proposes a Fairness Index (JFI) which is a normalized measure that reflects fairness with which the total available network resources are shared by all O-RUs. Contrary to Max-to-Min Load Difference and Load Standard Deviation that use absolute values, the value produced by JFI is bounded between 0 and 1. The values closer to 1 put depicts nearly perfect fairness. Let $L_i$ (t) represent the load of O-RU i at time step t, and N be the total number of O-RUs. At time t, Jain's Fairness Index of Jain is defined as.

$$\mathrm{JFI}(t)=\frac{\left(\sum_{i=1}^{N}L_i(t)\right)^2}{N\cdot\sum_{i=1}^{N}L_i^2(t)} \tag{16}$$

The average fairness across the simulation is then:

$$\overline{\mathrm{JFI}}=\frac{1}{T}\sum_{t=1}^{T}\mathrm{JFI}(t) \tag{17}$$

where $T$ denotes the total number of time steps. Higher $\overline{\mathrm{JFI}}$ values reflect a more equitable distribution of UEs among O-RUs, while lower values indicate uneven resource allocation.

Overall network efficiency is evaluated using throughput, which reflects the aggregate data rate achievable under the prevailing load and channel conditions [36]. Let $r_u(t)$ denote the instantaneous throughput of user $u$ at time step $t$. The network throughput at time $t$ is:

$$R_{\text{net}}(t)=\sum_{u=1}^{U}r_u(t) \tag{18}$$

where $U$ is the number of active users. Under equal resource sharing, per-user throughput is modeled as:

$$r_u(t)=\eta B\log_2(1+\mathrm{SINR}_u(t))\cdot\frac{1}{n_{i(u,t)}(t)} \tag{19}$$

where $\eta$ is the efficiency factor, $B$ is the available bandwidth, and $n_{i(u,t)}(t)$ is the number of users connected to the serving $\mathrm{O-RU}i(u,t)$.

The average network throughput over $T$ time steps is:

$$\bar{R}_{\text{net}}=\frac{1}{T}\sum_{t=1}^{T}R_{\text{net}}(t) \tag{20}$$

Higher $\bar{R}_{\text{net}}$ indicates better spectral efficiency under realistic sharing; Usually, the improvements occur when the SINR is enhanced. In addition, bottleneck or congestion is avoided via load balancing.

Another metric to calculate is the success rate (HSR) of handovers, which is the percentage of such attempts completed successfully without service interruption. We define $N_{\text{succ}}$ as the number of successful handovers during the simulation and $N_{\text{total}}$ as the total number of handovers attempted. Method of Calculating HSR.

$$\mathrm{HSR}=\frac{N_{\text{succ}}}{N_{\text{total}}} \tag{21}$$

The range of this metric is 0 to 1. Higher values are more reliable for maintaining a connection while mobile. A high HSR indicates that the handoff decision process, signaling, and resource allocation all function effectively in practice.

The repeated hand-offs of users between cells may lead to inefficient mobility, even when the cell is quite successful. This term refers to the Ping-Pong Rate (PPR). To allow for its calculation, let $N_{\text{pp}}$ denote the total number of detected ping-pong events and $N_{\text{succ}}$ the total number of successful handovers. The measurement has the following:

$$\mathrm{PPR}=\frac{N_{\text{pp}}}{N_{\text{succ}}} \tag{22}$$

PPR low values denote more stable handover decisions as users often do not oscillate between cells unnecessarily. In balanced networks, the ping-pong rate is likely to be low when the success rate is high for proper mobility management.

## V. RESULTS AND DISCUSSION

The proposed FedRL-O-RAN framework evaluation is structured on 2 categories of performance indicators. The first classification aims to assess the efficiency of load balancing using parameters such as Max-to-Min Load Difference, Load Standard Deviation, and Jain's Fairness Index (JFI), which measure the uniformity and temporal stability of user distributions across O-RUs. The 2nd category is user-experience quality. This is measured using Throughput, Handover Success Rate (HRS), and Ping-Pong Rate (PPR). The first metric we consider is the maximum-to-minimum load difference, which measures the difference in UE load distribution across O-RUs at each time step. A low value indicates a more balanced load across the network, whereas a high value indicates severe congestion in some cells and underutilization in others. In an O-RAN environment, when load distribution is uneven, spectral efficiency decreases. Moreover, Quality of Service (QoS) also suffers due to uneven loads. Figure 1 shows the allocation of the distributed resource for single- and mixed-load types.

The figure shows a small average difference and a tight variability range, indicating a consistently balanced distribution of the UE. On the other hand, the static, signal-strength based association policy of rLBRA produces the widest spread. Although centralized RL and MARL yield moderate gains, they still occasionally experience imbalances. The FedRL remains stable because it builds on locally adaptable decisions, followed by a global model sync to handle congestion.

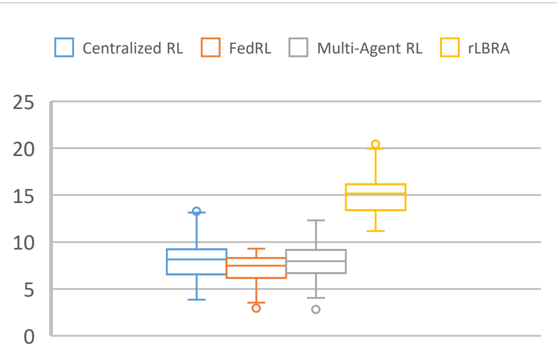


FIGURE 1 the boxplot for the Max-to-Min Load Difference across all evaluated approaches

We also analyze the Load Standard Deviation, which reflects the overall variability in load allocation across the O-RUs over time in addition to extreme-value differences. Lower values indicate enduring fairness and minor fluctuations, which are vital for sustaining network performance. Figure 2 displays comparative boxplots illustrating the Load standard deviation. Our framework has the lowest average standard deviation, coupled with a very compact interquartile range, indicating a stable, properly distributed load. The rLBRA methodology exhibits the greatest variability due to its frequent overload of strong signal cells.

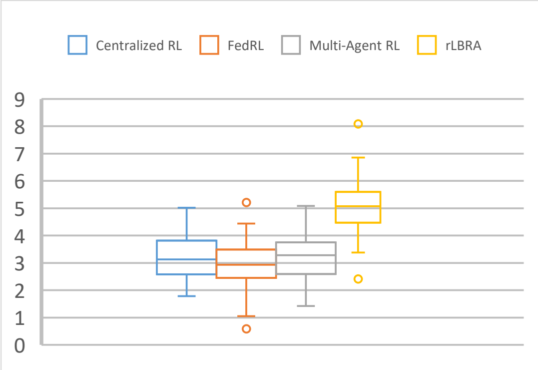


FIGURE 2 the comparative boxplot for the Load Standard Deviation among the evaluated methods

Models for single-agent forecasting of future service supply and demand have been developed. The special characteristic of FedRL is that it can respond quickly to local model change while synchronously and coherently updating the global policy. It preserves transient stability when a delay in the onset of load increase occurs. To measure the fairness in the allocation of resources Jain's Fairness Index (JFI), which is a normalized measure and varies from 0 to 1, with closer to 1 being more fair. With high fairness, no particular subset of users has consistently worse performance (Figure 3).

The fairness values of the proposed FedRL algorithm exceed 0.95 for all simulation sets. Thus, FedRL can produce more reliable fairness values than the baseline methods. The rLBRA method results in the lowest fairness values. This is because the static, signal-strength-driven association overutilizes some O-RUs and underutilizes others. While centralized reinforcement learning (RL) enhances rLBRA, its response to localized variations is sluggish.

In contrast, multi-agent RL (MARL) allows local adaptation but lacks coordination between agents, leading to unequal outcomes. FedRL is superior and fairer than prior work due to its two-tier architecture. In particular, it achieves local learning sensitive to the cell's conditions, while the global model aligns resource-sharing strategies across the network.

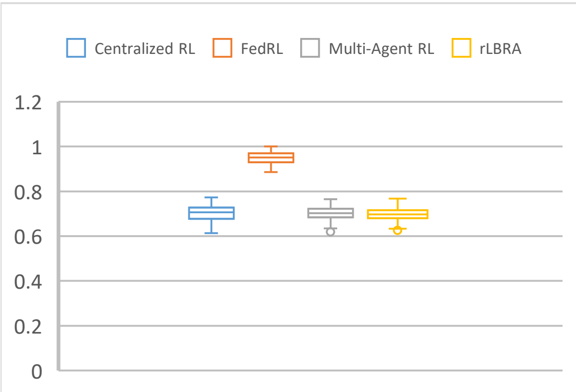


FIGURE 3 JFI distribution for all compared methods

While these performance results reinforce FedRL's ability to achieve even, stable load distribution, it is noteworthy that load balancing alone cannot guarantee a good end-user experience. Consequently, the evaluation of user-experience metrics, Throughput, Handover Success Rate (HSR) and Ping-Pong Rate (PPR) will be done to find out the practical advantage of balanced resources.

Throughput is defined as the total data rate of all active users under prevailing load and channel conditions. Higher throughput means better spectrum utilization and better quality of service (QoS), while stability over time indicates robustness against mobility and traffic.

As shown in figure 4, the practical impact of load balancing on a given metric can be easily gauged. Throughput assessment results indicate that the median values for FedRL were the highest while exhibiting the lowest variability, showcasing its broad adaptability. The rLBRA method has the least performance and shows a lower median and higher variance efficiency due to congestion in strong signal cells. Centralized RL and MARL achieve moderate performance improvements, with MARL slightly better due to faster adaptation to local changes. FedRL's better performance is due to its balanced load distribution, reduced unnecessary handovers, and consistent maintenance of good SINR values.

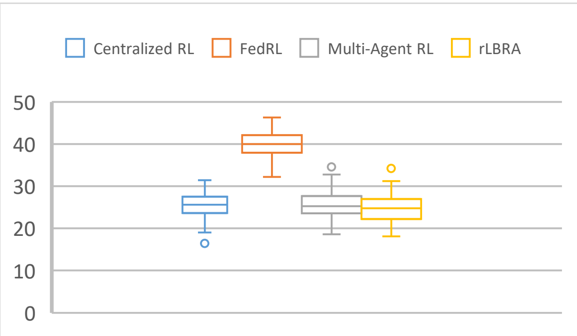


FIGURE 4 comparative throughput distributions for the evaluated methods.

Handover Success Rate is the ratio of handover attempts completed successfully without service interruption. The higher the value, the more reliable the mobility management; the lower it is, the more likely it may signal problems with decision-making, signalling, or resource assignment during handovers. In dynamic O-RAN deployments, consistently high HSR is essential to maintain user connectivity without interruption. According to Figure 5, this metric allows direct comparison of the reliability of different decision-making strategies.

According to HSR results, FedRL maintained values of 85% to 88%, the highest among the compared schemes, while rLBRA, Centralized RL, and MARL cluster at 75% to 80%. Thus, it can be inferred that rLBRA's limited performance stems from its static association policy, which fails to account for real-time mobility and load variations. Although MARL and Centralized RL improved on reactive agents, neither used advanced hyperparameter tuning (Centralized RL) or coordination among agents (MARL). FedRL combines a context-aware local decision-making process with an infrequent global policy update mechanism. Thus, timely and stable handover execution becomes possible with lower failure rates.

The Ping-Pong Rate (PPR) quantifies the proportion of handovers that quickly reverse, returning a UE to its previous serving cell within a short interval. High PPR values indicate instability in mobility management, leading to unnecessary signaling, increased latency, and degraded quality of experience. Reducing this metric is essential to ensuring smooth, efficient handover operations. As illustrated in Figure 6, PPR provides a measure of handover stability across the evaluated methods.

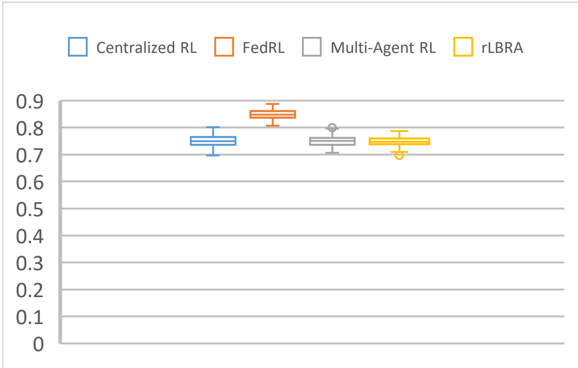


FIGURE 5 comparative HSR performance for all evaluated methods

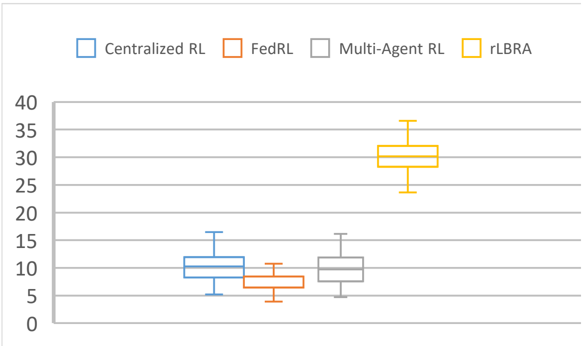


FIGURE 6 The comparative PPR results for all evaluated methods

According to the PPR analysis, we achieve the lowest median value with the most compact distribution. This indicates that our handover decisions are stable and well-judged. The rLBRA method shows the most ping-pong effect, caused by a static signal-threshold trigger that reacts to transient fluctuations and does not take the user trajectory into account. MARL outperforms Centralized RL on this parameter, as it effectively adapts to local mobility patterns. However, it lacks coordination, making it less consistent than FedRL. The edge of the FedRL over other methods is its own reward function, which discourages instability in Handover. Other is the O-RU of neighboring it, aligned in policy, so it helps avoid discordant action. The recommended FedRL-O-RAN framework naturally aligns with the Near-RT RIC architecture in the O-RAN standards from a deployment viewpoint. The signaling overhead for federated updates is still relatively modest, as each O-RU only needs to send its compact Q-table rather than raw user data, as in centralized state collection. In addition, the aggregation function at the Near-RT RIC consists of simple averaging operations, which scale linearly with the number of O-RUs that join the process. As a result, it leads to a low computation load that can be adopted in real-time RIC. As more devices are connected to the network, hierarchical or clustered aggregation approaches can further enhance scalability.

To evaluate the scalability of the proposed framework, the deployment scenarios summarized in Table 2 were evaluated. By varying the number of O-RUs and the deployment area, the analysis examines whether FedRL can maintain its load-balancing and mobility-management performance as the network size increases.

The scalability results presented in Table 3 show that the proposed framework remains effective across all evaluated scenarios. While larger deployment areas increase mobility complexity and user traversal distances, a higher number of O-RUs creates a denser handover environment. Despite these challenges, the federated learning mechanism maintains policy consistency across distributed agents, resulting in stable fairness, handover reliability, and load-balancing performance.

Table 3 summarizes the scalability performance of the proposed FedRL framework under different deployment areas and O-RU densities. The results show that FedRL maintains a high level of fairness across all evaluated scenarios, with Jain's Fairness Index remaining above 0.95. Likewise, the handover success rate remains close to 88%, indicating that the learned policy continues to make effective mobility-management decisions as the network scales.

TABLE 3
SCALABILITY PERFORMANCE SUMMARY

| Scenario | O-RUs | JFI ↑ | HSR ↑ | PPR ↓ | Load Difference ↓ |
|---|---|---|---|---|---|
| S1 | 7 | 0.968 | 0.879 | 0.032 | 8.04 |
| S2 | 14 | 0.964 | 0.881 | 0.031 | 8.71 |
| S3 | 7 | 0.962 | 0.876 | 0.034 | 9.12 |
| S4 | 14 | 0.959 | 0.878 | 0.029 | 9.93 |

The load difference shows only a modest increase as the deployment area and O-RU density increase, indicating that the proposed federated learning framework continues to balance traffic effectively across neighboring O-RUs. Similarly, the ping-pong rate remains consistently low across all scenarios. Overall, the results demonstrate that the proposed framework scales effectively while preserving the fairness, mobility-management, and load-balancing benefits achieved in the baseline deployment.

## VI. CONCLUSIONS

The paper presents FedRL-O-RAN, a decentralized load-balancing and handover-optimization framework for O-RAN-enabled 5G and beyond networks. The use of distributed reinforcement learning via Q-learning agents at every O-RU which enables the agents to make context-aware decisions locally. Furthermore, periodic federated averaging enables the agents to coordinate their decision-making. Thus, this scheme facilitates local context-aware decisions while maintaining global policy alignment and data privacy. In a modelled multi-cell O-RAN topology, FedRL achieved a reduction of more than 50% in load imbalance compared to rule-based schemes and 25% compared to centralized RL. Jain's Index consistently indicates fairness above 0.95. The framework's use further improved throughput by more than 15%, increased handover success rates by nearly 13% compared to the rule-based method, and reduced ping-pong events by 30 to 40%, thereby improving stability and signaling overhead. These benefits indicate that the system can provide balanced load distribution, fairness, and quality of service despite varying mobility and traffic. The proposed solution incorporates proven architectures and principles to build a scalable, agile foundation for intelligent RAN automation, while leveraging the flexibility and openness of next-generation architectures. In addition to showing improved performance, this project shows a practical path to integrate federated reinforcement learning into control loops using Near-RT RIC. The suggested architecture can be used for more detailed coordination tasks, such as multi-objective optimization, energy-aware control, and slice-level orchestration, in B5G and 6G systems. Additional scalability experiments showed that the proposed FedRL framework continues to deliver effective load balancing and reliable mobility management as the deployment area and O-RU density increase. In the future, research can explore hierarchical federated aggregation (HFA), non-IID traffic scenarios, and integration with advanced antenna and beamforming models.

## REFERENCES

[1] N. Q. A. AlShaikhli, M. A. Neamah, and M. I. Aal-Nouman, "Design of triband antenna for telecommunications and network applications," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 1084–1090, 2022. DOI: 10.11591/eei.v11i2.3233

[2] C. Adamczyk and A. Kliks, "Conflict mitigation framework and conflict detection in O-RAN near-RT RIC," *IEEE Communications Magazine*, vol. 61, no. 12, pp. 199–205, 2023. DOI: 10.1109/MCOM.018.2200752

[3] O-RAN Alliance, "O-RAN: Towards an Open and Smart RAN," White paper, Aug. 2018.

[4] M. Polese, L. Bonati, S. D'oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376–1411, 2023. DOI: 10.1109/COMST.2023.3239220

[5] W. Azariah, F. A. Bimo, C.-W. Lin, R.-G. Cheng, N. Nikaein, and R. Jana, "A Survey on Open Radio Access Networks: Challenges, Research Directions, and Open Source Approaches," *Sensors*, vol. 24, no. 3, 2024, DOI: 10.3390/s24031038.

[6] K. Alam et al., "A Comprehensive Tutorial and Survey of O-RAN: Exploring Slicing-Aware Architecture, Deployment Options, Use Cases, and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 28, pp. 1637–1678, 2025, DOI: 10.1109/COMST.2025.3598406

[7] Q. Sun, N. Li, C. I. -L., J. Huang, X. Xu, and Y. Xie, "Intelligent RAN Automation for 5G and Beyond," *IEEE Wirel Commun*, vol. 31, no. 1, pp. 94–102, 2024, DOI: 10.1109/MWC.014.2200271.

[8] R. Ntassah, G. M. Dell'Aera, and F. Granelli, "PPO-EPO: Energy and Performance Optimization for O-RAN Using Reinforcement Learning," arXiv preprint. DOI: 10.48550/arXiv.2504.14749

[9] A. Almeida, P. Rito, S. Brás, F. C. Pinto, and S. Sargento, "A machine learning approach to forecast 5G metrics in a commercial and operational 5G platform: 5G and mobility," *Comput Commun*, vol. 228, p. 107 974, 2024. DOI: 10.1016/j.comcom.2024.107974

[10] E. Gures, I. Shayea, M. Ergen, M. H. Azmi, and A. A. El-Saleh, “Machine learning-based load balancing algorithms in future heterogeneous networks: A survey,” I*EEE Access*, vol. 10, pp. 37 689– 37717, 2022. **DOI**: 10.1109/ACCESS.2022.3161511

[11] J. Ochoa-Aldeán, C. Silva-Cárdenas, R. Torres, J. I. Gonzalez, and S. Fortes, “Algorithms for Load Balancing in Next-Generation Mobile Networks: A Systematic Literature Review,” *Future Internet*, vol. 17, no. 7, p. 290, 2025. **DOI**: 10.3390/fi17070290

[12] R. K. Mohammed and N. N. Khamiss, “Hybrid Multiple Access Techniques Performance Analysis Of Dynamic Resource Allocation,” *Iraqi Journal of Information and Communication Technology*, vol. 7, no. 1, pp. 23–34, 2024. **DOI**: 10.31987/ijict.7.1.243

[13] Kiran, Kotaru, and D. Rajeswara Rao. "Analytical review and study on various vertical handover management technologies in 5G heterogeneous network." *Infocommunications Journal 14*, no. 2 (2022): 28-38. **DOI**: 10.36244/ICJ.2022.2.3

[14] N. Saqib, N. F. Abdullah, A. Abu-Samah, H. A. H. Alobaidy, and R. Nordin, “Adaptive VNF Placement Considering Overall Latency and 5G Wireless Channel Reliability in Industry 4.0: A Reinforcement Learning Based Approach,” *IEEE Access*, vol. 12, pp. 88883–88896, 2024, **DOI**: 10.1109/ACCESS.2024.3419065.

[15] O. Orhan, V. N. Swamy, T. Tetzlaff, M. Nassar, H. Nikopour, and S. Talwar, “Connection management xAPP for O-RAN RIC: A graph neural network and reinforcement learning approach,” in *2021 20th IEEE international conference on machine learning and applications (ICMLA)*, IEEE, 2021, pp. 936–941. **DOI**: 10.1109/ICMLA52953.2021.00154.

[16] C.-H. Lai, L.-H. Shen, and K.-T. Feng, “Intelligent load balancing and resource allocation in O-RAN: A multi-agent multi-armed bandit approach,” in *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, IEEE, 2023, pp. 1–6. **DOI**: 10.1109/PIMRC56721.2023.10293795

[17] H. Zhang, H. Zhou, and M. Erol-Kantarci, “Federated deep reinforcement learning for resource allocation in O-RAN slicing,” in *GLOBECOM 2022-2022 IEEE Global Communications Conference*, IEEE, 2022, pp. 958–963. **DOI**: 10.1109/GLOBECOM48099.2022.10001658

[18] B. Zhao, Q. Cui, W. Ni, X. Li, and S. Liang, “Multi-Layer Collaborative Federated Learning architecture for 6G Open RAN,” *Wireless Networks*, vol. 31, no. 2, pp. 1377–1390, 2025. **DOI**: 10.1007/s11276-024-03823-0

[19] M. Behjati, H. A. H. Alobaidy, R. Nordin, and N. F. Abdullah, “UAV-assisted federated learning with hybrid LoRa P2P/LoRaWAN for sustainable biosphere,” *Frontiers in Communications and Networks*, vol. Volume 6-2025, 2025, **DOI**: 10.3389/frcmn.2025.1529453.

[20] M. Q. Hamdan et al., “Recent Advances in Machine Learning for Network Automation in the O-RAN,” *Sensors*, vol. 23, no. 21, 2023, **DOI**: 10.3390/s23218792.

[21] Y. Azimi, S. Yousefi, H. Kalbkhani, and T. Kunz, “Mobility aware and energy-efficient federated deep reinforcement learning assisted resource allocation for 5G-RAN slicing,” *Comput Commun*, vol. 217, pp. 166– 182, 2024, **DOI**: 10.1016/j.comcom.2024.01.028.

[22] S. Gu, C. You, B. Ren, and D. Guo, “Communication and Computation Efficient Split Federated Learning in O-RAN,” arXiv preprint 2025. **DOI**: 10.48550/arXiv.2508.02534

[23] T. Pamuklu, M. Erol-Kantarci, and C. Ersoy, “Reinforcement learning based dynamic function splitting in disaggregated green open RANs,” in *ICC 2021-IEEE International Conference on Communications*, IEEE, 2021, pp. 1–6. **DOI**: 10.1109/ICC42927.2021.9500721

[24] Q. Ali, “Performance Optimization and Architectural Advancements in Cloud Radio Access Networks (C-RAN) for 5G and Beyond,” *Journal of Electronic & Information Systems| Volume*, vol. 7, no. 01, 2025. **DOI**: 10.30564/jeis.v7i1.9960

[25] T. S. Kumar et al., “Advanced Handover Optimization (AHO) using deep reinforcement learning in 5G Networks,” *Journal of King Saud University Computer and Information Sciences*, vol. 37, no. 6, pp. 1–14, 2025. **DOI**: 10.1007/s44443-025-00124-0

[26] M. Gupta, R. M. Dreifuerst, A. Yazdan, P.-H. Huang, S. Kasturia, and J. G. Andrews, “Load balancing and handover optimization in multi-band networks using deep reinforcement learning,” in *2021 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2021, pp. 1–6. **DOI**: 10.1109/GLOBECOM46510.2021.9685781

[27] P. Sroka, Ł. Kułacz, S. Janji, M. Dryjański, and A. Kliks, “Policy-based traffic steering and load balancing in o-ran-based vehicle-to-network communications,” *IEEE Trans Veh Technol*, vol. 73, no. 7, pp. 9356–9369, 2024.

[28] H. Lin, Y.-K. Su, H.-Q. Chen, and L.-F. Ko, “Federated Deep Q-Learning and 5G load balancing,” *arXiv preprint arXiv:2403.08813*, 2024. **DOI**: 10.48550/arXiv.2403.08813

[29] Y. Cao, S.-Y. Lien, Y.-C. Liang, K.-C. Chen, and X. Shen, “User access control in open radio access networks: A federated deep reinforcement learning approach,” *IEEE Trans Wirel Commun*, vol. 21, no. 6, pp. 3721–3736, 2021. **DOI**: 10.1109/TWC.2021.3123500

[30] F. Ahmed et al., “Federated Deep Reinforcement Learning-Driven O-RAN for Automatic Multirobot Reconfiguration” *NOMS 2025 IEEE Network Operations and Management Symposium*, Honolulu, HI, USA, 2025, pp. 1–7, **DOI**: 10.1109/NOMS57970.2025.11073682.

[31] P. Abdisarabshali, N. Accurso, F. Malandra, W. Su, and S. Hosseinalipour, “Synergies between federated learning and o-ran: Towards an elastic virtualized architecture for multiple distributed machine learning services,” arXiv preprint 2023. **DOI**: 10.48550/arXiv.2305.02109.

[32] V.-D. Nguyen et al., “Network-aided intelligent traffic steering in 6G O-RAN: A multi-layer optimization framework,” *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 2, pp. 389–405, 2023. **DOI**: 10.1109/JSAC.2023.3336183.

[33] F. Kavehmadavani, V.-D. Nguyen, T. X. Vu, and S. Chatzinotas, “Intelligent traffic steering in beyond 5G open RAN based on LSTM traffic prediction,” *IEEE Trans Wirel Commun*, vol. 22, no. 11, pp. 7727–7742, 2023. **DOI**: 10.1109/TWC.2023.3254903

[34] C. Sun, Y. Zhou, G. Jung, T. X. Tran, and D. Pompili, “Carl: Cascade reinforcement learning with state space splitting for o-ran based traffic steering,” *arXiv preprint arXiv:2312.01970*, 2023.

[35] O. Semiari, H. Nikopour, and S. Talwar, “*Graph Reinforcement Learning for QoS-Aware Load Balancing in Open Radio Access Networks*,” arXiv e-prints, p. arXiv-2504, 2025. **DOI**: 10.1109/ICCWorkshops67674.2025.11162460

[36] Raya Adel Kamil , Saif Mohamed Baraa Alsabti, Rusul K. Abdulsattar, Ammar H. Mohammed, and Taha A. Elwi "On the Enhancement Anomaly Detection for RF Bio-Sensors by Computing Artificial Networks Using Machine Learning Techniques", *Infocommunications Journal*, Vol. XVII, No 2, June 2025, pp. 89-95., **DOI**: 10.36244/ICJ.2025.2.11

**Mohammed I. Aal-Nouman** (Member, IEEE) received his Ph.D. in Telecommunications Engineering from the University of Salford, Manchester, UK, in 2017. He got his MSc. In Computer Engineering from Al-Nahrain University in 2007. He is currently a Lecturer in the College of Information Engineering at Al-Nahrain University, Baghdad, where he has served since 2008. His research interests include mobile and wireless communications, heterogeneous networks, and performance optimization in next-generation communication systems.

**Sarmad M. Hadi** is a lecturer at the College of Information Engineering, Al-Nahrain University, Iraq, and the head of the university's computer center. He received his M.Sc. degree from Al-Nahrain University, Iraq (2010), and his Ph.D. degree in Information and Communication Engineering from Al-Nahrain University (2020). His research interests include AI, web applications, multimedia, bioinformatics and modern communication technologies.

**Haider A. H. Alobaidy** (Member, IEEE) is a lecturer at the College of Information Engineering, Al-Nahrain University, Iraq. Prior to his current role, he served as a postdoctoral researcher at Universiti Kebangsaan Malaysia (UKM), where he further honed his expertise in wireless communication systems and related fields. He received an M.Sc. degree from Al-Mustansiriyah University, Iraq (2016), and a Ph.D. degree in Electrical, Electronics, and Systems Engineering from UKM (2022). His research interests include wireless communication, wireless sensor networks, IoT, machine learning, and channel propagation modeling and estimation.