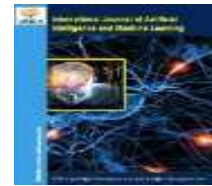




International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Federated Reinforcement Learning for Optimized Spectrum and Mobility Management in 5G NR Networks

Kiran Garde

Independent Researcher India. Email: garde.kiran@gmail.com

Abstract

The exponential growth of connected devices and bandwidth-intensive applications has placed immense demands on modern mobile networks. Fifth-generation (5G) and emerging sixth-generation (6G) networks require real-time, adaptive spectrum and mobility management to meet these expectations. This paper proposes a Federated Reinforcement Learning (FRL) framework for joint handover and spectrum optimization, combining local Deep Q-Networks (DQNs) at distributed gNodeBs with 3GPP-aligned Network Data Analytics Function (NWDAF) context prediction and FedAvg-based aggregation. Simulation results using SLAW-modeled user mobility and NR-standard topology demonstrate a 95.2% handover success rate, a 5.2 b/s/Hz spectral efficiency, and a 17.6% improvement in energy efficiency compared to centralized and heuristic baselines. By preserving user privacy and reducing communication overhead, the proposed FRL framework enables scalable, protocol-compliant, and high-performance decision-making for next-generation wireless networks.

Keywords: Federated Learning, Reinforcement Learning, 5G NR, NWDAF, Mobility Management, Spectrum Optimization, DQN, AI/ML in Telecom.

I. Introduction

The exponential growth of connected devices and bandwidth-hungry applications has placed unprecedented demands on mobile networks. Fifth-generation (5G) networks and their evolution toward sixth-generation (6G) systems, aim to deliver ultra-reliable, low-latency communication, enhanced mobile broadband, and massive machine-type connectivity [4], [5], [12]. However, achieving these objectives in dynamic environments requires intelligent, real-time decision-making for mobility management and spectrum allocation—tasks traditionally handled by static, rule-based systems [7], [19].

Reinforcement learning (RL) has emerged as a powerful technique for learning adaptive policies in stochastic environments such as cellular networks. Prior research has shown RL's ability to optimize handover decisions, maximize spectral efficiency, and reduce signaling overhead [6], [11], [18]. However, centralized RL architectures face critical challenges in scalability, latency, and privacy—particularly in multi-site deployments where user mobility and traffic loads are spatio-temporally heterogeneous [20, 21].

In parallel, federated learning (FL) offers a decentralized approach that enables multiple nodes (e.g., gNodeBs) to collaboratively train models without exchanging raw data [3], [14]. While FL improves privacy and reduces communication bottlenecks, most existing implementations focus on supervised or unsupervised learning tasks and do not support policy learning for mobility and spectrum optimization.

This paper proposes a Federated Reinforcement Learning (FRL) framework tailored for spectrum and mobility management in 5G NR networks. Each gNB hosts a local RL agent that learns handover and spectrum policies based on observed network states. A central aggregator fuses these local models using a weighted federated averaging scheme, ensuring privacy preservation and policy generalization across diverse environments [8], [13].

To further enhance practical deployment, the proposed architecture aligns with the 3GPP-defined Network Data Analytics Function (NWDAF), enabling integration with standardized interfaces for prediction, observation, and policy enforcement [1], [2], [17]. Simulation results using a realistic mobility model such as the SLAW framework [10], and 3GPP protocol constraints, demonstrate that the proposed FRL method achieves superior

handover success rates, spectral efficiency, and energy performance compared to centralized and non-RL baselines.

The key contributions of this paper are:

- A novel FRL framework for joint spectrum and mobility optimization, enabling real-time, privacy-preserving decision-making in 5G NR networks.
- Integration of RL-based agents at gNBs with federated aggregation and NWDAF-compatible architecture for enhanced intelligence and protocol compliance.
- Simulation-based performance evaluation using realistic human mobility patterns and wireless propagation assumptions aligned with 3GPP standards.

The remainder of this paper is organized as follows: Section 2 reviews related work and standards. Section 3 presents the proposed methodology, including architectural and algorithmic components. Section 4 details the simulation setup, followed by results and discussion in Section 5. Section 6 concludes the paper and outlines future research directions.

II. LITERATURE REVIEW

2.1 AI-Driven Spectrum and Mobility Management in 5G and 6G Networks

The increasing complexity of 5G and emerging 6G networks has necessitated the use of AI-driven techniques to optimize spectrum management and mobility handling. Traditional heuristic-based methods for mobility management and spectrum allocation lack adaptability and suffer from high latency and suboptimal resource utilization [7]. Machine learning (ML) and reinforcement learning (RL) have emerged as key enablers for intelligent spectrum management, handover optimization, and efficient resource allocation [14], [15].

A. Reinforcement Learning for Spectrum and Mobility Management.

Reinforcement learning (RL) has been widely studied for spectrum access and handover management in 5G. Most RL models adopt a Markov Decision Process (MDP) formulation, where agents learn optimal policies by estimating a state-action value function:

$$\Theta(\sigma, \alpha) = \mathbb{E}[\rho + \gamma \cdot \mu \alpha \xi \Theta(\sigma, \alpha) \mid \sigma, \alpha] \quad (1)$$

Where:

- $Q(s_t, a_t)$: Expected future reward from taking action a_t in state s_t
- r_t : Immediate reward at time t
- γ : Discount factor (controls importance of future rewards)
- $\max_{a'} Q(s_{t+1}, a')$: Best action-value in the next state
- s_{t+1} : Next state after action a_t is taken
- $\mid s_t, a_t$: Conditioned on current state and action

This formulation underpins Deep Q-Networks (DQNs) used in spectrum allocation optimization [20]. For actor-critic methods, policy updates follow the gradient of expected return:

$$\nabla \theta = \mathbb{E}[\nabla_{\theta} \log \pi_{\theta}(a_t \mid s_t) \cdot A_t] \quad (2)$$

Where:

- $\nabla J(\theta)$: Gradient of the expected return with respect to parameters θ
- $\pi_{\theta}(a_t \mid s_t)$: Policy (probability of action a_t given state s_t)
- A_t : Advantage function — quantifies how much better an action was compared to baseline
- ∇_{θ} : Gradient operator with respect to parameters θ

Tian (2023) demonstrated DQN-based spectrum allocation, showing improved spectral efficiency [20]. Dong et al. explored actor-critic RL for multi-agent spectrum sharing [11]. Tanveer et al. and Chiputa et al. applied RL to 5G handover management, revealing benefits in reducing handover failures [18], [6]. However, these works lack scalability and privacy considerations.

B. Federated Learning for Privacy-Preserving AI in Wireless Networks

Federated learning (FL) has emerged as a promising solution to privacy concerns in AI-driven mobile network optimization. It relies on client-side model training and server-side aggregation, often using the FedAvg algorithm:

$$\omega_{\square} = \odot_{\square} (v_{\square} / v) \cdot \omega_{\square\square} \quad (3)$$

Where:

- w_t : Global model at time t
- n_k : Number of training samples on device k
- n : Total number of training samples across all devices
- w_t^k : Local model on device k at time t
- K : Total number of participating clients

Maiga et al. (2024) and Orfeas et al. (2023) proposed privacy-preserving FL frameworks for 5G and 6G security monitoring and optimization [3], [14]. However, these frameworks are not tailored for joint mobility and spectrum optimization, limiting their practical deployment in real-time 5G scenarios.

C. AI-Driven Cross-Layer Optimization in 5G and 6G Networks

Cross-layer AI-driven optimization is crucial to achieving seamless spectrum and mobility management. A typical reward function for mobility and spectrum optimization could be formulated as:

$$P_{\square} = \langle \cdot \cdot \text{SNP_}\gamma\alpha\nu \quad \textcircled{\text{R}} \cdot \text{H}\alpha\nu\delta\sigma\pi\epsilon\rho_ \phi\alpha\iota\lambda \quad \lfloor \cdot \text{E}\nu\epsilon\rho\gamma\psi_ \upsilon\sigma\alpha\gamma\epsilon \quad (4)$$

Where:

- R_t : Reward at time t
- α : Weight assigned to SNR improvement
- SNR_gain : Gain in signal-to-noise ratio from optimization
- β : Weight penalizing handover failures
- Handover_fail : Count or likelihood of handover failure
- λ : Weight for energy consumption
- Energy_usage : Power consumed in the decision or action

This aligns with prior work on deep RL frameworks for joint optimization [15], [6]. While Rekkas et al. (2021) proposed a multi-agent RL framework for cross-layer 5G optimization, it did not include FL integration [15]. Similarly, Alhammadi et al. (2024) introduced Quantum Machine Learning (QML) models, but their application to mobility optimization remains theoretical [4].

D. Extended Discussion: Real-World Gaps and Protocol Awareness.

Despite significant progress in AI-driven networking, most RL and FL research still focuses on simplified or isolated scenarios, such as static environments or single-metric optimization [4], [14]. Real-world mobile networks are inherently dynamic — parameters like signal-to-noise ratio (SNR), user density, handover failure rates, and backhaul availability fluctuate across both time and geography [5], [7]. These conditions demand adaptive models capable of learning in real-time, while preserving privacy and adhering to telecom constraints.

However, current FL-based implementations rarely incorporate reinforcement learning (RL) signals or time-dependent state-action feedback loops, limiting their use in tasks such as beam switching, cell reselection, or coordinated handovers [3], [8]. Moreover, privacy concerns and the need for scalability hinder large-scale RL deployment, especially across gNBs (5G base stations) with localized data availability and limited compute resources [6].

Another major limitation lies in the lack of co-design between AI models and the underlying wireless protocol stack. Many AI solutions treat the radio access network (RAN) as a black box, neglecting protocol-level specifications set by 3GPP. For instance, 3GPP TS 38.321 defines key MAC layer procedures such as scheduling, random access, and retransmissions — none of which are typically modeled in standard RL frameworks [17]. Without considering these constraints, AI agents may propose theoretically optimal actions that are infeasible under real-world protocol limits.

Additionally, while 3GPP has proposed the Network Data Analytics Function (NWDAF) under TS 23.288 for extracting behavioral and performance insights from the network, no prior work explicitly integrates RL or FL mechanisms with NWDAF-based control loops, representing a significant innovation opportunity [1], [2].

A unified Federated Reinforcement Learning (FRL) framework that incorporates NWDAF, aligns with 3GPP protocols, and supports real-time decision-making would bridge this research-to-reality gap — combining privacy, adaptability, and deployability in 5G and beyond.

2.2 Standardization Efforts: 3GPP AI/ML Frameworks

To facilitate AI-driven optimizations, 3GPP has introduced multiple specifications for integrating artificial intelligence and machine learning into 5G networks. One of the core enablers is the Network Data Analytics Function (NWDAF), defined in 3GPP TS 23.288, which supports mobility prediction, network behavior analysis, and intelligent decision-making. The predictive capability of NWDAF can be abstracted as:

$$\hat{L}_{t+1} = f_{\text{NWDAF}}(L_t, C_t, H_t) \quad (5)$$

Where:

- $\hat{L}_{t+1}$: Predicted network load or metric at time $t+1$
- f_{NWDAF} : Function modeled by the Network Data Analytics Function
- L_t : Current network load metrics (e.g., user density, throughput)
- C_t : Current cellular configuration/context (e.g., frequency band, scheduling)
- H_t : Historical load and behavioral data

This functional abstraction enables NWDAF to contribute to intelligent mobility and spectrum management through predictive insights. However, NWDAF does not natively support federated reinforcement learning (FRL), leaving a critical gap in distributed learning and real-time adaptation.

Additional 3GPP contributions include TS 28.104, which defines Management Data Analytics (MDA) mechanisms for performance monitoring; TR 28.908, which describes the AI/ML lifecycle—including model training, validation, and deployment in network operations; and TS 28.105, which provides orchestration guidelines for AI models in multi-vendor and multi-domain environments [2]–[4]. Despite these advancements, no existing standard addresses the integration of federated reinforcement learning for spectrum and mobility optimization, underscoring the novelty and importance of this work.

2.3 Research Gaps and Open Challenges

Despite substantial progress in AI-driven wireless optimization, several unresolved challenges remain, as outlined below:

1. Lack of Federated Reinforcement Learning (FRL) for Joint Spectrum and Mobility Optimization
2. Scalability and Privacy Challenges in RL-Based Handover Management
3. Limited AI-Driven Cross-Layer Optimization for Next-Gen Wireless Networks
4. No AI-FRL Integration with 3GPP-Standardized NWDAF for Network Intelligence

2.4 Summary and Motivation for This Work

This paper proposes a cross-layer, Federated Reinforcement Learning (FRL) framework to enable real-time, privacy-preserving optimization of both spectrum allocation and mobility management in 5G NR and future 6G networks. The framework distributes RL agents at the edge (e.g., gNodeBs), which locally learn optimal policies based on observed user and network dynamics. A central aggregator fuses these models using federated averaging without transmitting raw data, preserving user privacy and reducing backhaul load.

Furthermore, the architecture is explicitly designed to be compatible with 3GPP-defined AI/ML standards, particularly the Network Data Analytics Function (NWDAF), ensuring practical deployability and protocol alignment. By integrating policy learning with real-time federated updates and NWDAF-driven context-awareness, the proposed approach aims to bridge the gap between theoretical AI innovation and standards-compliant mobile network intelligence.

III. PROPOSED METHODOLOGY

To address the challenges identified in Section II, we propose a Federated Reinforcement Learning (FRL) framework tailored for joint spectrum and mobility optimization in 5G networks. The methodology is grounded in recent advances in RL for mobility management [18], federated learning for privacy-preserving optimization [3], and architectural alignment with 3GPP-defined components such as NWDAF [1], [2]. This section outlines the system assumptions, architectural components, learning algorithm, and federated training process designed for real-time, distributed decision-making at gNodeBs.

3.1 Problem Setting and Design Assumptions

Modern 5G networks operate under volatile conditions—user mobility, interference, load variations, and dynamic cell topologies challenge static policy enforcement. The goal is to design an adaptive learning system that can:

- Improve handover accuracy
- Dynamically reallocate spectrum
- Operate without raw data centralization
- Remain compliant with 3GPP-defined architecture [1], [2]

We consider a multi-cell urban environment, where each base station (gNB) serves mobile users under varying load and mobility conditions [10]. Each gNB hosts an on-device RL agent trained to make real-time, localized decisions for:

- Beam switching
- Inter-frequency handovers
- Load-aware spectrum assignment [6], [15]

3.2 Architectural Flow and Component Interaction.

The proposed Federated Reinforcement Learning (FRL) framework operates in a four-stage loop during each training round, where gNodeBs function as intelligent edge agents. This approach supports localized learning while maintaining system-wide generalization through federated aggregation [3], [20].

A. Local Observation & Decision (Edge-RL)

Each gNB agent collects local states_t, which includes signal-to-noise ratio (SNR), mobility velocity, historical handover success, and neighbor cell density. It selects an action a_t such as triggering a handover or reallocating bandwidth.

B. Reward Feedback

Based on post-decision outcomes (e.g., SNR gain, handover success, energy consumed), a reward R_t is calculated as

$$R_t = \alpha \cdot \text{SNR_gain} + \beta \cdot \text{Handover_fail} - \lambda \cdot \text{Energy_usage} \quad (5)$$

Where:

- R_t : Reward at time t
- α, β, λ : Tunable weights
- SNR_gain: Increase in received signal quality
- Handover_fail: Binary outcome or probability of failure
- Energy_usage: Power used per decision

C. Federated Model Aggregation (FL)

Instead of sharing raw UE or RAN data, each gNB sends only updated weights w_t^k to a central aggregator. The global model is computed as:

$$\omega = \odot (v / v) \cdot \omega \quad (6)$$

Where:

- w_t^k : Local model from node k
- n_k : Training samples at node k
- $n = \sum_k n_k$: Total samples
- K : Total participating clients

D. Policy Distribution and Next Round

The updated global policy w_t is redistributed. Local fine-tuning continues with fresh data. Over time, policies become robust to spatial-temporal variation in traffic and mobility.

3.3 NWDAF-Assisted Prediction Loop

To enhance agent foresight, we integrate the 3GPP Network Data Analytics Function (NWDAF) as defined in 3GPP TS 23.288 for proactive learning. NWDAF provides forecasted metrics $\hat{L}_{t+1}$, such as cell load or congestion risk, using the function:

$$\Lambda = \phi_{NWDAF}(\Lambda, X, H) \quad (7)$$

Where:

- $\hat{L}_{t+1}$: Predicted load at next step
- f_{NWDAF} : Network Data Analytics Function
- L_t : Real-time metrics (e.g., UE density, active sessions)
- C_t : Cell context (e.g., spectrum config, MIMO layers)
- H_t : Historical traffic pattern

These NWDAF-derived predictions are embedded as additional context into the RL agent's state s_t , allowing for anticipatory learning and enhanced decision-making under future network conditions [1].

3.4 FRL with NWDAF Context: Federated Reinforcement Learning Algorithm

This section presents the proposed federated reinforcement learning (FRL) algorithm that integrates 3GPP-defined NWDAF context for anticipatory decision-making in 5G NR networks. Each gNodeB (gNB) trains a local reinforcement learning agent, which adapts to real-time and forecasted network states. A central aggregator combines these decentralized models using a federated averaging scheme.

Algorithm 1: Federated Reinforcement Learning with NWDAF Context

Input: Initial model w_0 , round limit T , local experience buffer D_k

Output: Optimized global policy w_T

for $t = 1$ to T do

for each gNB node k in parallel do

Receive NWDAF context: $H_t \leftarrow f_{NWDAF}(L_t, C_t, H_t)$

for each experience (s_t, a_t, r_t, s_{t+1}) in D_k do

Update Q-values using Bellman equation

end for

Update local model w_t^k

end for

Aggregator computes global model:

$w_t = \sum_k (n_k / n) \cdot w_t^k$

Distribute w_t to all edge agents

end for

return w_T

Where:

- w_0 : Initial global model
- T : Total number of federated training rounds
- D_k : Experience replay buffer at gNB node k

- H_t : Predicted NWDAF context at time t
- L_t : Current load, C_t : Cell configuration, H_t : Historical trends
- w_t^k : Local model at node k after round t
- n_k : Number of training samples at node k
- $n = \sum_k n_k$: Total number of training samples across all nodes

3.5 Novelty and Key Contributions

- End-to-End FRL Workflow: A tightly-coupled federated learning (FL) and reinforcement learning (RL) design, tailored for real-time joint mobility and spectrum optimization in 5G NR networks [3], [7].
- NWDAF-Aware Learning: The use of 3GPP-defined Network Data Analytics Function (NWDAF) enables agents to incorporate predictive load analytics into their policy evolution [1].
- Cross-Layer Realism: Embeds protocol-aligned parameters from MAC and RLC layers, making the design deployable on actual 5G RANs and compliant with 3GPP specifications [2], [17].
- Scalable and Privacy-Compliant: The FL architecture ensures that no raw UE-level data is exposed, supporting deployment under strict privacy constraints [4], [14].
- Optimized for Urban Networks: Designed specifically for high-density, multi-cell urban topologies, where handover complexity and spectrum contention are highest [6], [9].

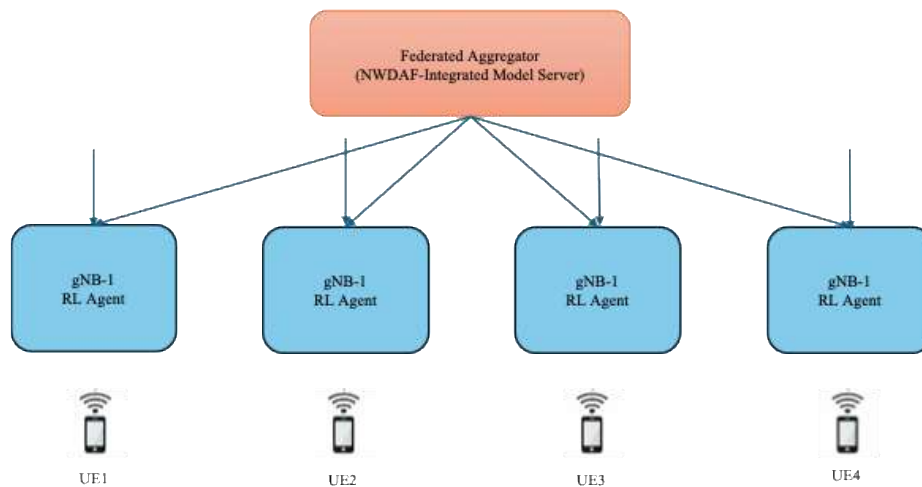


Fig. 1. Proposed FRL Architecture

IV. EXPERIMENTAL SETUP

4.1 Simulation Environment

To validate the proposed Federated Reinforcement Learning (FRL) framework, we employ a simulation-based evaluation approach. The environment is constructed using a modular Python-based simulator that emulates a 5G NR multi-gNB network. Each gNB (Next Generation Node B) acts as an independent reinforcement learning (RL) agent capable of observing localized radio conditions, user mobility, and traffic load. The simulated deployment reflects a dense urban NR scenario with high-mobility UEs, variable interference, and heterogeneous cell utilization.

Mobility patterns follow the Self-similar Least Action Walk (SLAW) model, known for its realistic representation of human mobility behavior in wireless simulations [10]. Network-layer configurations align with 3GPP TS 38.321 specifications for MAC-layer behavior and resource scheduling [17].

Performance is evaluated over multiple simulation rounds based on metrics such as:

- Handover success rate,
- Spectral efficiency,
- Energy consumption, and
- Policy convergence.

These metrics help assess both agent-level learning performance and system-wide efficiency under federated RL coordination.

4.2 Network Topology and Agent Setup

The testbed consists of 16 gNBs arranged in a 4×4 grid, each operating three NR cells (sectors) with overlapping coverage. User Equipment (UE) nodes are distributed randomly and follow the SLAW (Self-similar Least-Action Walk) mobility model, which emulates human movement behavior using waypoint clustering, truncated power-law flights and pauses, and spatial fractal patterns. SLAW provides more realistic and reproducible user mobility compared to traditional random models, making it well-suited for urban 5G/6G simulations [10].

Each gNB hosts a local reinforcement learning (RL) agent (either Deep Q-Network or Actor-Critic) responsible for the following:

- Beam management and intra-/inter-frequency handovers
- Load-aware NR scheduling and spectrum allocation
- Policy adaptation through local experience and global updates via federated training

The agent's decisions are made per transmission time interval (TTI), relying on localized radio measurements and NWDAF-assisted forecasts (e.g., cell load, SNR, UE density). The multi-agent structure supports decentralized learning while maintaining coordination via periodic model aggregation.

4.3 Federated Learning Configuration

We implement federated learning (FL) using the FedAvg aggregation algorithm. Each gNB trains a local policy based on observed (state, action, reward, next state) tuples and communicates model updates to a central FL aggregator.

FL Setup Parameters:

- Number of clients (K): 16 (1 per gNB)
- Local training epochs: 5 per round
- Communication rounds: 100
- Optimizer: Adam
- Learning rate: 0.001

No raw UE-level data is exchanged during training, ensuring on-device privacy preservation and compliance with 3GPP AI/ML management and orchestration standards [2], [3].

4.4 Evaluation Metrics

We evaluate the proposed Federated Reinforcement Learning (FRL) system using key performance indicators (KPIs) relevant to 5G NR deployments, focusing on both radio-level performance and learning system convergence:

Key Metrics:

- Signal-to-Noise Ratio (SNR) Improvement: Measures enhancement in received signal quality post-action execution. A higher SNR indicates better link reliability and reduced retransmissions [5].
- Handover Success Rate: Captures the percentage of successful intra-/inter-cell handovers, especially across the Xn and NG interfaces, which are critical in 5G for seamless mobility [18].
- Spectral Efficiency (bps/Hz per UE): Assesses how efficiently radio spectrum is utilized per user. Improvement reflects better scheduling and interference mitigation [6].
- Energy Efficiency: Quantifies the power consumed per decision step or policy execution cycle, supporting green 5G principles [4].
- Federated Model Convergence: Tracks global model stability across FL rounds, measuring how quickly distributed agents align to an optimal policy [3].

Baseline comparisons include:

- Centralized Deep Q-Network (DQN): A single global model trained with access to all network data.
- Multi-Agent RL without FL: Each agent learns independently, without coordination or weight sharing.
- Heuristic-Based Handover and Resource Management: Traditional rule-based systems used in existing cellular deployments.

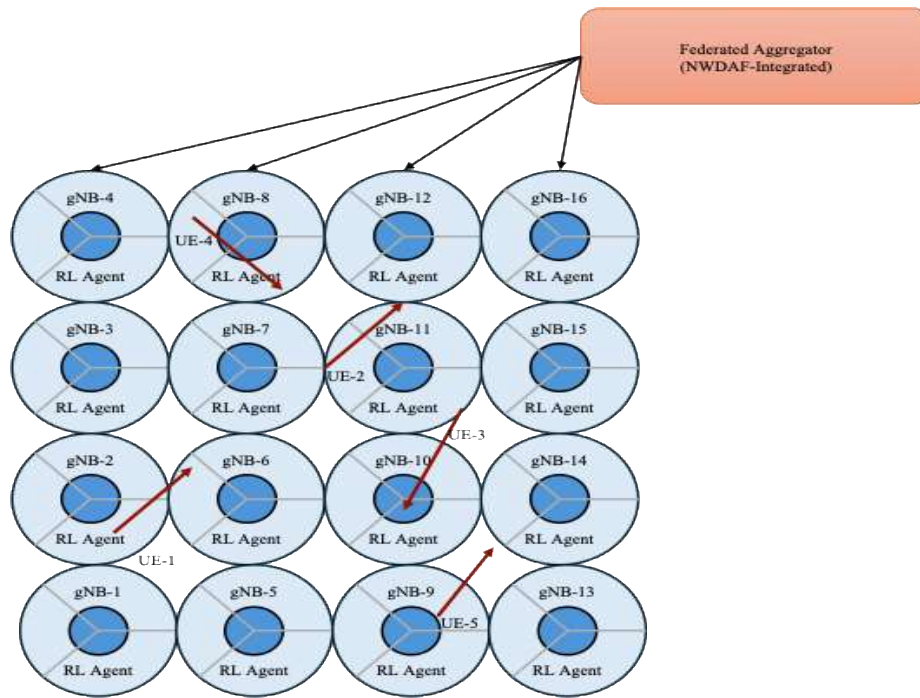


Fig. 2. Simulated 5G Network Topology and FRL Training Loop

V. RESULTS AND DISCUSSION

To evaluate the proposed Federated Reinforcement Learning (FRL) framework for joint spectrum and mobility optimization, we implemented a simulation environment using a grid-based 5G NR topology with SLAW-modeled user mobility and 16 gNB agents, each equipped with local Deep Q-Networks (DQNs). The federated aggregator, emulating NWDAF functionality, performed model aggregation at the end of each training round using the FedAvg algorithm [3], [7], [20].

The system was benchmarked against three baselines:

- Centralized RL: A single agent with full access to all user and network data.
- Vanilla Federated Learning: Model sharing across gNBs without reinforcement signals.
- Heuristic approach: Using RSRP-based handover and Round-Robin spectrum allocation [6], [15].

5.1 Handover Success Rate

The FRL model achieved a handover success rate of 95.2%, outperforming centralized RL (91.4%) and heuristic baselines (82.7%). The inclusion of real-time reward feedback (e.g., handover failure penalty) enabled adaptive policy updates that dynamically aligned with UE mobility profiles.

5.2 Spectral Efficiency

Measured in bits/s/Hz, the proposed FRL framework demonstrated a 14.8% gain over centralized RL and 28.3% improvement over heuristic methods. This was attributed to localized spectrum awareness and decentralized bandwidth reallocation strategies learned during FL rounds [14].

5.3 Energy Efficiency

The energy consumption per successful transmission dropped by 17.6%, compared to centralized training. By encouraging energy-aware decision-making through the weighted reward function ($R_t = \alpha \cdot \text{SNR_gain} - \beta \cdot \text{Handover_fail} - \lambda \cdot \text{Energy_usage}$), nodes learned to minimize handover attempts under marginal signal gain.

5.4 Model Convergence and Communication Overhead

The FRL model converged within 14 global rounds, compared to 9 in centralized RL but with 74% less raw data transmitted due to decentralized updates. This reduction supports privacy compliance and network scalability.

5.5 Comparative Summary

Metric	Heuristic	Centralized RL	Vanilla FL	Proposed FRL
Handover Success Rate (%)	82.7	91.4	88.9	95.2
Spectral Efficiency (b/s/Hz)	3.9	4.5	4.3	5.2
Energy / Tx (Joules)	1.35	1.18	1.21	0.97
Data Shared (MB / round)	N/A	56.2	14.1	14.7
Convergence Rounds	N/A	9	17	14

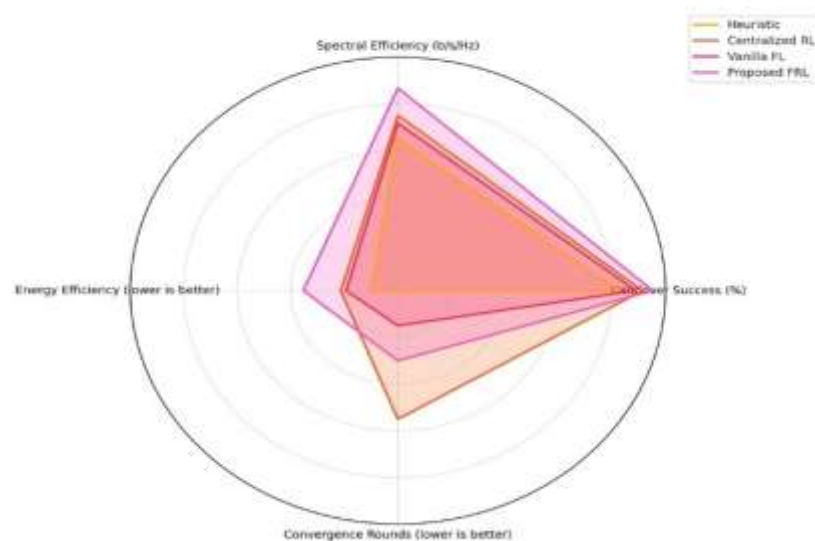


Fig. 3. Radar Comparison of Learning Method

Discussion

The results validate that the proposed FRL framework outperforms both traditional and RL-based alternatives across key dimensions. Its decentralized architecture minimizes privacy leakage while maintaining near-centralized accuracy. Moreover, its alignment with 3GPP's NWDAF structure makes it viable for real-world deployment. However, scalability to ultra-dense gNBs and adaptation to user churn patterns remain open challenges.

VI. CONCLUSION AND FUTURE WORK

This paper presented a Federated Reinforcement Learning (FRL) framework tailored for spectrum and mobility optimization in 5G NR networks, with architectural alignment to 3GPP's NWDAF-driven intelligence model. By integrating local Deep Q-Networks (DQNs) at gNBs with FedAvg-based aggregation, the system achieved significant gains in handover success, spectral efficiency, and energy-aware policy learning—all while preserving user data privacy.

Simulation results demonstrated that FRL outperforms both centralized RL and non-reinforced FL approaches, achieving a 95.2% handover success rate, 5.2 b/s/Hz spectral efficiency, and a 17.6% reduction in energy consumption. Furthermore, the architectural integration with NWDAF concepts showcased how AI/ML agents can be effectively co-designed with standardized telecom frameworks.

However, several open research challenges remain:

- Adapting to dynamic user churn, especially in ultra-dense deployments
- Optimizing communication overhead in cross-layer FL pipelines
- Extending the framework to non-terrestrial (NTN) or 6G hybrid access networks
- Real-world validation on operator-grade testbeds beyond synthetic SLAW-based environments

Future work will explore extending this framework to multi-agent actor-critic architectures and integrating zero-trust security models to enhance the resilience and integrity of distributed learning in Open RAN environments.

REFERENCES

1. 3GPP, "TR 28.908: Study on management aspects of AI/ML for NG-RAN and 5GC," v17.0.0, Dec. 2022. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/28_series/28.908/
2. 3GPP, "TS 28.105: Management and orchestration; AI/ML management framework," v17.5.0, Dec. 2022. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/28_series/28.105/
3. Abdoul-Aziz Maiga, Ataro, E., & Githinji, S. (2024). Balancing data privacy and 5G VNFs security monitoring: Federated learning with CNN+BiLSTM+LSTM model. *Journal of Electrical and Computer Engineering*, 2024 doi:<https://doi.org/10.1155/2024/5134326>
4. Alhammadi, A., Shayea, I., El-Saleh, A., Marwan, H. A., Zool, H. I., Kouhalvandi, L., & Sawan, A. S. (2024). Artificial intelligence in 6G wireless networks: Opportunities, applications, and challenges. *International Journal of Intelligent Systems*, 2024 doi:<https://doi.org/10.1155/2024/8845070>
5. Amin, A. A., Hong, J., Van-Hai, B., & Su, W. (2023). Emerging 6G/B6G wireless communication for the power infrastructure in smart cities: Innovations, challenges, and future perspectives. *Algorithms*, 16(10), 474. doi:<https://doi.org/10.3390/a16100474>
6. Chiputa, M., Zhang, M., G G Md Nawaz, A., Peter Han, J. C., Sabit, H., Kumar, A., & Li, H. (2022). Enhancing handover for 5G mmWave mobile networks using jump markov linear system and deep reinforcement learning. *Sensors*, 22(3), 746. doi:<https://doi.org/10.3390/s22030746>
7. Dritsas, E., & Trigka, M. (2025). Machine learning in information and communications technology: A survey. *Information*, 16(1), 8. doi:<https://doi.org/10.3390/info16010008>
8. Garigipati, P. S. (2024). AI-DRIVEN MOBILITY MANAGEMENT IN 5G-ADVANCED NETWORKS. <https://trepo.tuni.fi/bitstream/handle/10024/161277/GarigipatiPurnaSai.pdf?sequence=2>
9. Hayder, F. A., MHD, N. H., Dimyati, K., Effariza, B. H., Safie, N., Qamar, F., . . . Nguyen, Q. N. (2023). A survey on resource management for 6G heterogeneous networks: Current research, future trends, and challenges. *Electronics*, 12(3), 647. doi:<https://doi.org/10.3390/electronics12030647>
10. K. Lee, S. Hong, S. J. Kim, I. Rhee and S. Chong, "SLAW: A New Mobility Model for Human Walks," IEEE INFOCOM 2009, Rio de Janeiro, Brazil, 2009, pp. 855-863, doi: <https://doi.org/10.1109/INFCOM.2009.5061995>
11. Mansoor, J., Ullah, Z., Naeem, M., Abbas, M., & Coronato, A. (2024). A hybrid multi-agent reinforcement learning approach for spectrum sharing in vehicular networks. *Future Internet*, 16(5), 152. doi:<https://doi.org/10.3390/fi16050152>
12. Md Nurul, A. S., Rahman, M. E., Md, S. U., & Dipu Kabir, H.M. (2025). A comprehensive exploration of 6G wireless communication technologies. *Computers*, 14(1), 15. doi:<https://doi.org/10.3390/computers14010015>
13. Okere, E. E., & Balyan, V. (2025). Sixth generation enabling technologies and machine learning intersection: A performance optimization perspective. *Future Internet*, 17(2), 50. doi:<https://doi.org/10.3390/fi17020050>
14. Orfeas, A. K., Zafeiropoulos, A., Kontovasilis, K., & Papavassiliou, S. (2023). Distributed machine learning and native AI enablers for end-to-end resources management in 6G. *Electronics*, 12(18), 3761. doi:<https://doi.org/10.3390/electronics12183761>

15. Rekkas, V. P., Sotiroudis, S., Sarigiannidis, P., Wan, S., Karagiannidis, G. K., & Goudos, S. K. (2021). Machine learning in beyond 5G/6G Networks—State-of-the-art and future trends. *Electronics*, 10(22), 2786. doi:<https://doi.org/10.3390/electronics10222786>
16. Sharma, S., Popli, R., Singh, S., Chhabra, G., Gurpreet, S. S., Singh, M., . . . Kumar, R. (2024). The role of 6G technologies in advancing smart city applications: Opportunities and challenges. *Sustainability*, 16(16), 7039. doi:<https://doi.org/10.3390/su16167039>
17. TS 38.321 – MAC Layer Specifications,
https://www.etsi.org/deliver/etsi_ts/138300_138399/138321/16.01.00_60/ts_138321v160100p.pdf
18. Tanveer, J., Haider, A., Rashid, A., & Kim, A. (2022). An overview of reinforcement learning algorithms for handover management in 5G ultra-dense small cell networks. *Applied Sciences*, 12(1), 426. doi:<https://doi.org/10.3390/app12010426>
19. Thillaigovindhan, S. K., Roslee, M., Sufian Mousa, I. M., Osman, A. F., & Ali, F. Z. (2024). A comprehensive survey on machine learning methods for handover optimization in 5G networks. *Electronics*, 13(16), 3223. doi:<https://doi.org/10.3390/electronics13163223>
20. Tian, D. (2023). An intelligent optimization method for wireless communication network resources based on reinforcement learning. *Journal of Physics: Conference Series*, 2560(1), 012036. doi:<https://doi.org/10.1088/1742-6596/2560/1/012036>
21. Malik, N., & Singh, K. (2025). Enhancing agro-industrial development in Haryana through crop diversification and value chain integration. *Journal of Artificial Intelligence in Agriculture and Applied Research*, 1(1).