ICUFN2026
The 17th International Conference on Ubiquitous and Future Networks
July 07-10, 2026, Milan, Italy
https://icufn.org/

2026.07.07

Self-Directed Spectrum Allocation Framework for Integrated TN-NTN 6G Networks

TN-NTN Integration · Spectrum Allocation · Q-Learning · 6G Networks

Vaskar Chakma, and Wooyeol Choi
Chung-Ang University, Seoul, Republic of Korea

OUTLINE

Why Spectrum Allocation in TN-NTN 6G is Hard?

Rapid Topology Changes

LEO satellites travel at 7–8 km/s, forcing frequent handovers and continuously shifting interference profiles. Static allocation schemes have no mechanism to track these transitions in real time.

Heterogeneous QoS Demands

Ground base stations, HAPs, and UAVs each carry fundamentally different traffic profiles with conflicting QoS requirements. A single policy that satisfies all nodes simultaneously is far from trivial.

Co-Channel Interference

When TN and NTN nodes share spectrum, load imbalances across channels create inter-network interference that simultaneously erodes throughput and degrades fairness for all users.

- *Conventional approaches cannot jointly model TN load, NTN load, interference, and temporal dynamics - an integrated, self-adaptive framework is essential for 6G TN-NTN deployments.*

2 Related Work

Prior Work & Research Gap

1 **DQN for Spectrum Sharing**
Adebayo et al., IoT 2025
Dynamic spectrum sharing between LTE and NB-IoT via deep Q-networks, improving throughput while preserving Jain's fairness - but limited to single-network terrestrial scenarios.

2 **DRL in TV Whitespace**
Ukpong et al., Sci. African 2025
Achieved 96.34% interference avoidance in cognitive radio networks using DRL for dynamic spectrum access - again confined to single-tier, homogeneous network environments.

3 **Multi-Agent DRL for LEO Uplink**
Cho et al., IEEE Commun. Lett. 2023
Multi-agent DRL for interference-aware channel allocation in LEO satellite uplink using sequential training to tackle non-stationarity. Closest to our setting, but no TN-NTN co-existence modelling.

- Elhachmi (IET Networks, 2022) showed distributed Q-learning optimises dynamic spectrum allocation in IoT with careful state discretization - validating our tabular approach for TN-NTN.

Why Q-Learning?

- Model-free - no explicit environment model needed
- Interpretable Q-table, ideal for intermediate-sized state spaces
- Tabular precision outperforms DQN at $|S| = 2{,}000$ states

Research Gap

No prior Q-learning work jointly encodes TN-NTN co-channel interference and temporal traffic asymmetry into the state space.

This paper fills that gap.

3 System Model & MDP Formulation

Network Setup, State Space & Reward Design

Network Configuration

N_{TN} = 15 terrestrial users
N_{NTN} = 8 non-terrestrial users
K = 10 channels, 100 MHz bandwidth
LEO altitude = h = 550 km, v = 7.5 km/s
Traffic model = Sinusoidal: μ_{TN} = 50 Mbps, μ_{NTN} = 30 Mbps

MDP State Space |S| = 2000

TN Load (avg) = 10 bins
NTN Load (avg) = 10 bins
Interference Level = 5 bins
Time Period = 4 bins

Co-Channel Interference: $I = \frac{1}{K}\sum_{K=1}^{K} |L_{TN}^{(k)} - L_{NTN}^{(k)}|$ where L = percentage channel load on channel k

Multi-Objective Reward Function

$$R(s,a) = Throughput(s,a) - 0.3 * LoadImbalance(s) - 0.2 * I(s) + 10 * F(s)$$

Throughput
Maximise spectral efficiency

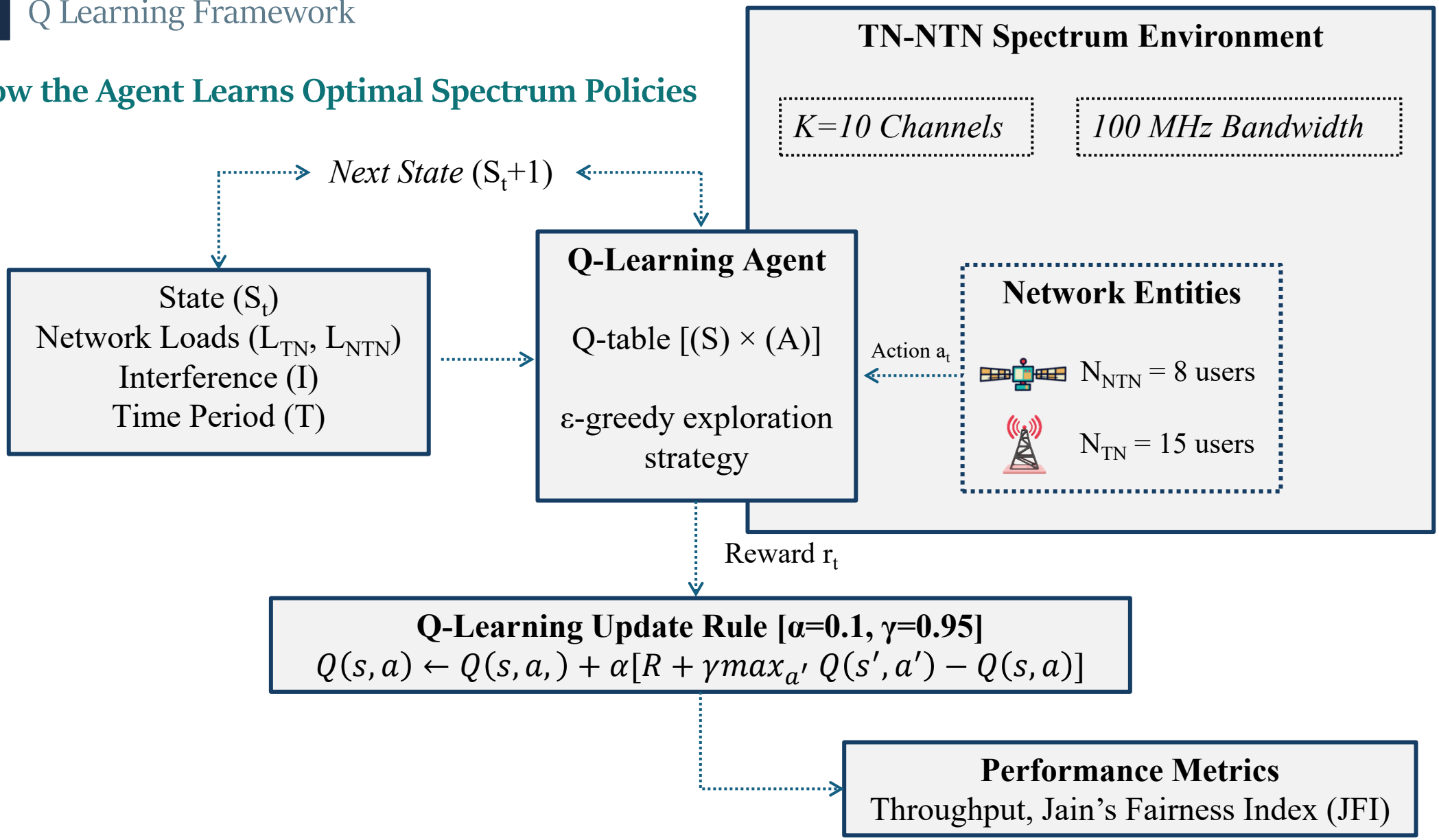
Load Balance (0.3)
Penalise uneven channel loading

Interference (0.2)
Suppress co-channel interference

Jain's Fairness (10)
Equity across TN & NTN users

4 Q Learning Framework

How the Agent Learns Optimal Spectrum Policies



4 Simulation Results

Training Convergence & Policy Behavior

~2,000
Episodes
Convergence point

37.5
Avg Reward
Per episode (stable)

28.5
Mbps
Average throughput

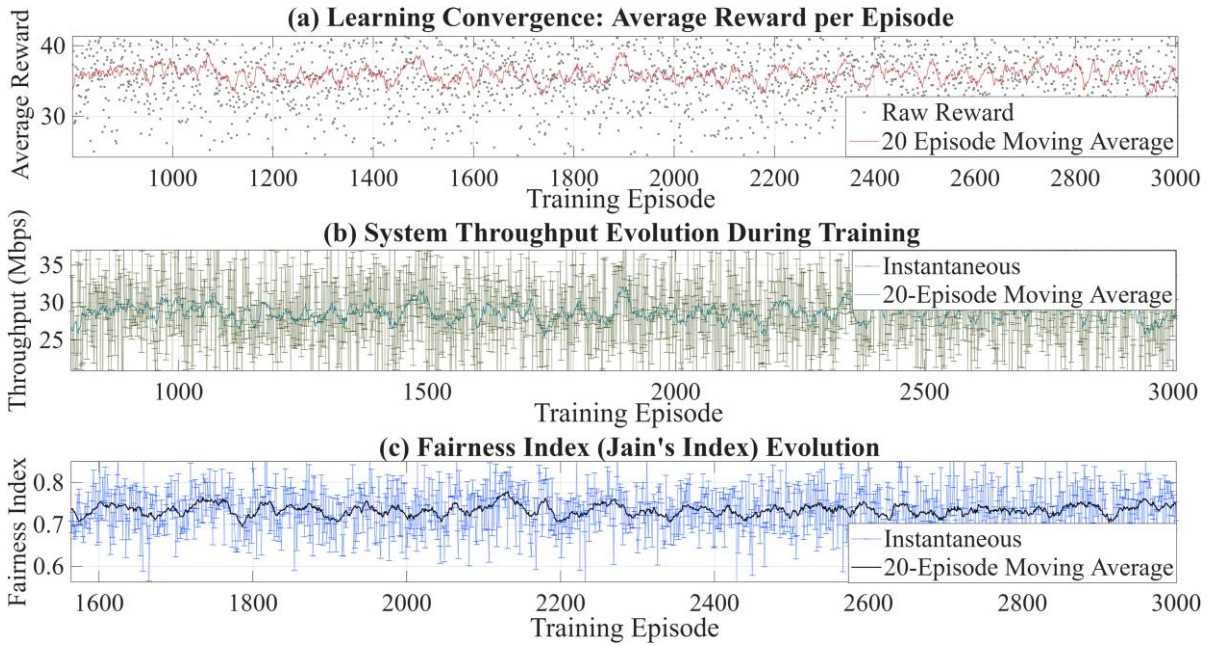
0.75
JFI ± 0.08
Jain's Fairness Index

Reward Convergence

The 20-episode moving average smooths the high variance of instantaneous rewards and reveals stable convergence around episode 2,000 with an average reward of 37.5. Prior to convergence the agent is actively exploring via ε-greedy decay.

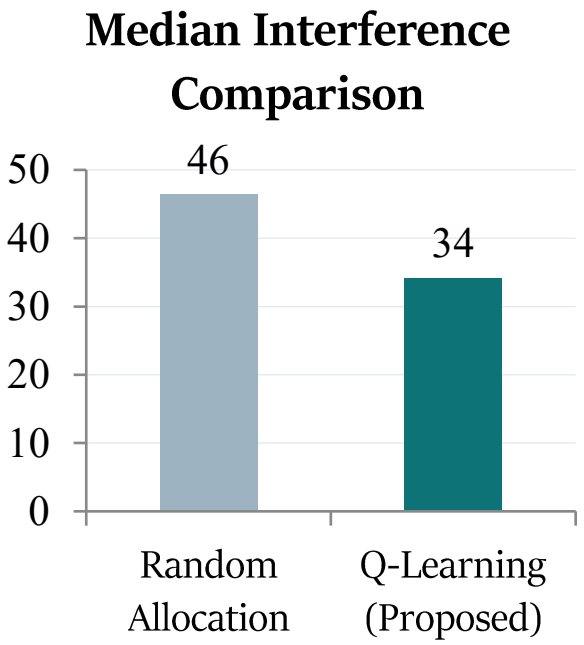
Throughput Stability

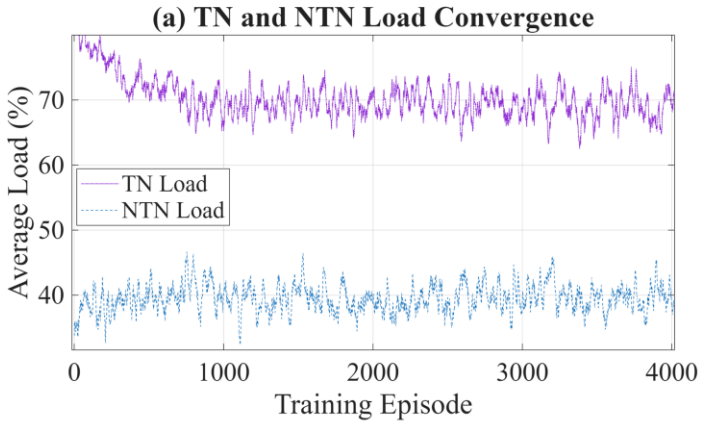
System throughput settles consistently at 28.5 Mbps. As the ε-greedy policy shifts from exploration to exploitation, the agent increasingly selects less-congested channels - directly translating to higher spectral efficiency.

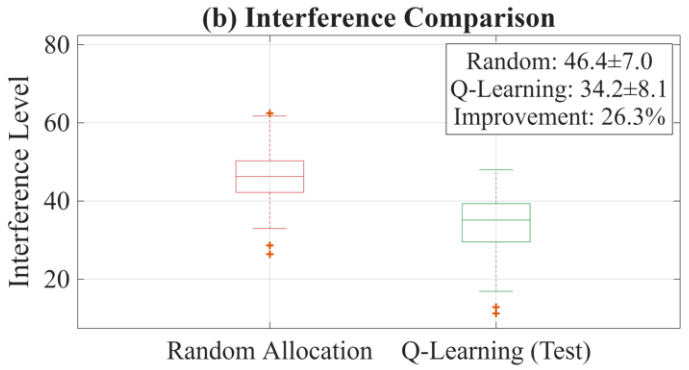


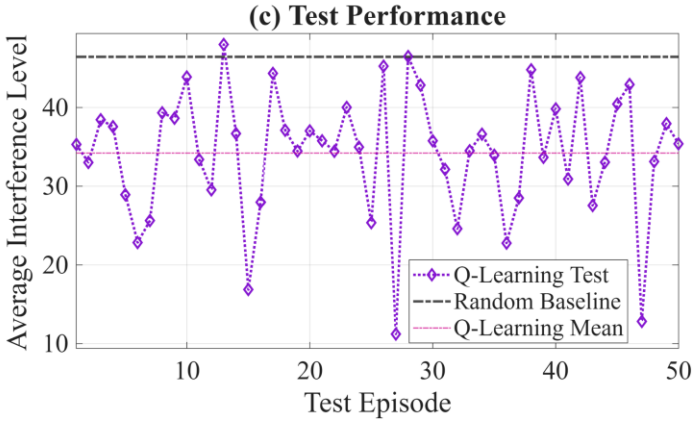
4 Simulation Results

Interference Mitigation & Load Balancing









TN-NTN Load Convergence

TN Load Starts >80% → stabilises ≈ 70%

NTN Load Holds consistently ≈ 40%

26.3% Reduction in average interference versus random allocation baseline

Test-Phase Performance (50 episodes)

Q-learning mean ≈ 34.2 vs random baseline ≈ 46.4. Interference dips below 20 in episodes with favourable network states - the policy actively exploits low-interference windows.

4 Simulation Results

Channel Utilization & Q-Value Policy Structure

◆ Channel Utilisation:

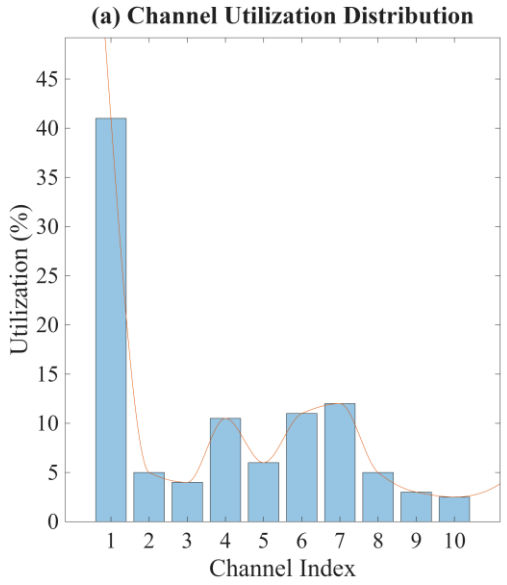
Channel 1 absorbs 39% of traffic. Remaining channels share 4-12% each, confirming effective load distribution. Channels 1, 8, and 10 recur most in the test-phase heatmap, reflecting temporal adaptation by the learned policy.

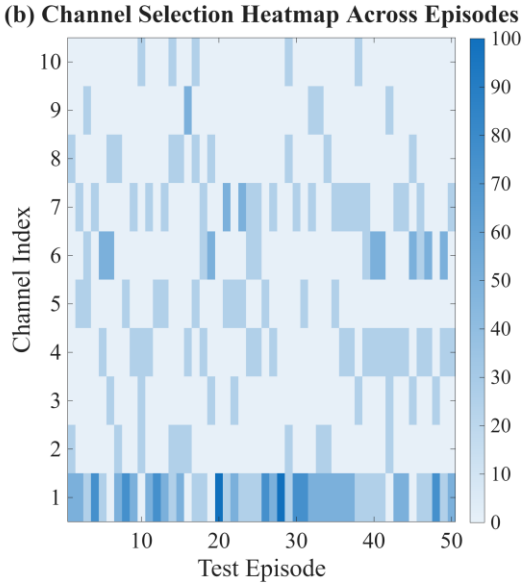
◆ Q-Values Differentiate Actions:

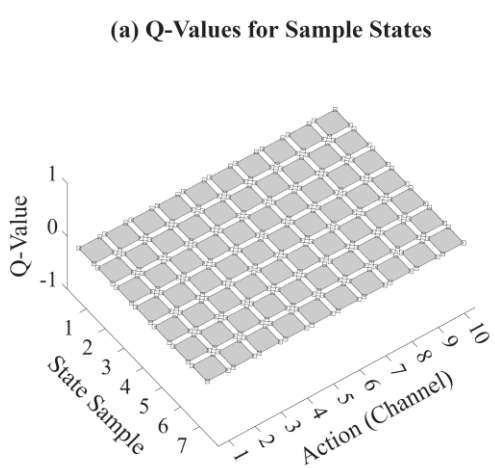
For sample states, Q-values trend upward with channel index. The agent has learned a non-uniform preference grounded in real channel quality differences, not random selection.

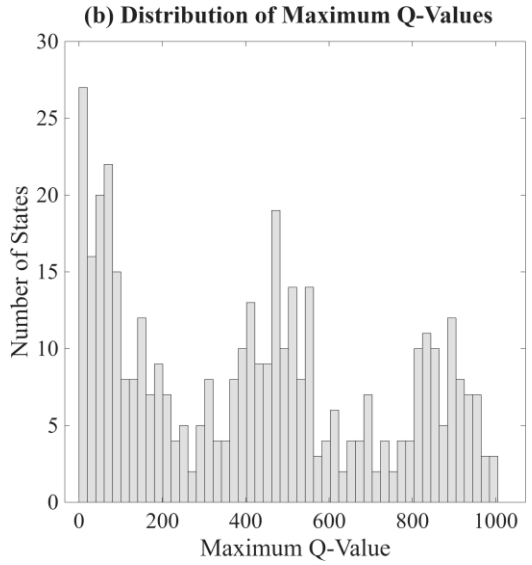
◆ Skewed Q-Value Distribution:

Most of the 2,000 states carry low maximum Q-values; a small subset yields significantly higher returns - confirming the agent exploits high-reward state-action pairs effectively.









Simulation Summary vs Baseline

Metric	Random Allocation	Q-Learning (Proposed)	Improvement
Avg. Reward / Episode	~15–20 (unstable)	37.5 (stable)	↑ ~2× gain
System Throughput	~18–22 Mbps	28.5 Mbps	↑ ~30% increase
Jain's Fairness Index	0.58–0.65 (variable)	0.75 ± 0.08	↑ More equitable
Median Interference	46.4	34.2	↓ 26.3% reduction
Policy Consistency	None (random each step)	Stable after ~2,000 eps	Converges & holds

📶 Throughput up 30% over baseline, reaching a stable 28.5 Mbps after convergence.

⚖️ Fairness maintained at 0.75 - equitable service for both TN and NTN users throughout.

📉 26.3% median interference reduction - improved signal quality and lower packet loss.

5 Discussion

Strengths & Limitations

Strengths

Unified TN-NTN State Encoding

First Q-learning framework to jointly encode co-channel interference and temporal traffic asymmetry for integrated TN-NTN - no prior tabular RL work addresses this combination.

Multi-Objective Optimization

Simultaneously balances throughput, load equity, interference suppression, and Jain's Fairness in a single reward signal - no post-hoc trade-off tuning required.

Interpretable & Efficient Policy

A 2,000-state tabular Q-table offers full interpretability. Unlike DQN, the policy can be inspected, audited, and deployed with negligible inference cost at runtime.

Verified Convergence

Stable convergence confirmed after ~2,000 episodes across 5,000-episode runs. Fairness and throughput metrics both hold steady post-convergence with low variance.

Limitations

Single Shared Channel Action

Each action selects one channel shared across all users. Per-user channel assignment, more realistic for heterogeneous networks is not yet modelled.

Tabular State Space Scalability

A 2,000-state discretization is practical here, but denser traffic scenarios or finer bin resolution would require deep RL (DQN) to avoid the curse of dimensionality.

Single-Agent Decision Scope

The current agent makes centralized decisions. Distributed satellite-level decision-making requires multi-agent RL, which introduces non-stationarity challenges.

Single Representative Run

Results are from one representative MATLAB simulation run. Statistical averaging across multiple seeds is needed for stronger empirical claims.

5 Future Work

Where We Go from Here

Near-Term 🧠 **Deep Q-Networks (DQN)**

Replace the tabular Q-table with a neural network to handle larger, continuous state spaces, enabling finer-grained load, interference, and mobility modelling across denser TN-NTN topologies.

→ *Extension: Double DQN & Dueling DQN for variance reduction*

Near-Term 🛰️ **Multi-Agent RL Across Satellites**

Deploy independent Q-learning (or MARL) agents at each satellite or base station, enabling distributed spectrum decisions - while addressing non-stationarity through sequential or cooperative training.

→ *Challenge: coordinated channel assignment without centralized control*

Mid-Term 📊 **Rigorous Baseline Comparisons**

Evaluate against greedy allocation, DQN, Double DQN, and Dueling DQN under identical simulation conditions across multiple random seeds for statistically valid performance claims.

→ *Metrics: throughput CDF, fairness distribution, interference PDF*

Mid-Term 🔬 **Digital Twin-Based Validation**

Reproduce the TN-NTN environment as a digital twin - embedding real-world LEO orbital mechanics, HAP station-keeping, and UAV trajectory models - for policy validation beyond synthetic sinusoidal traffic.

→ *Target: validate on real KPI traces from 5G NTN testbed*

Thank You!

Intelligent Networking Lab (INL)
Department of Computer Science and Engineering
Chung-Ang University
Seoul, Republic of Korea

vaskar@cau.ac.kr