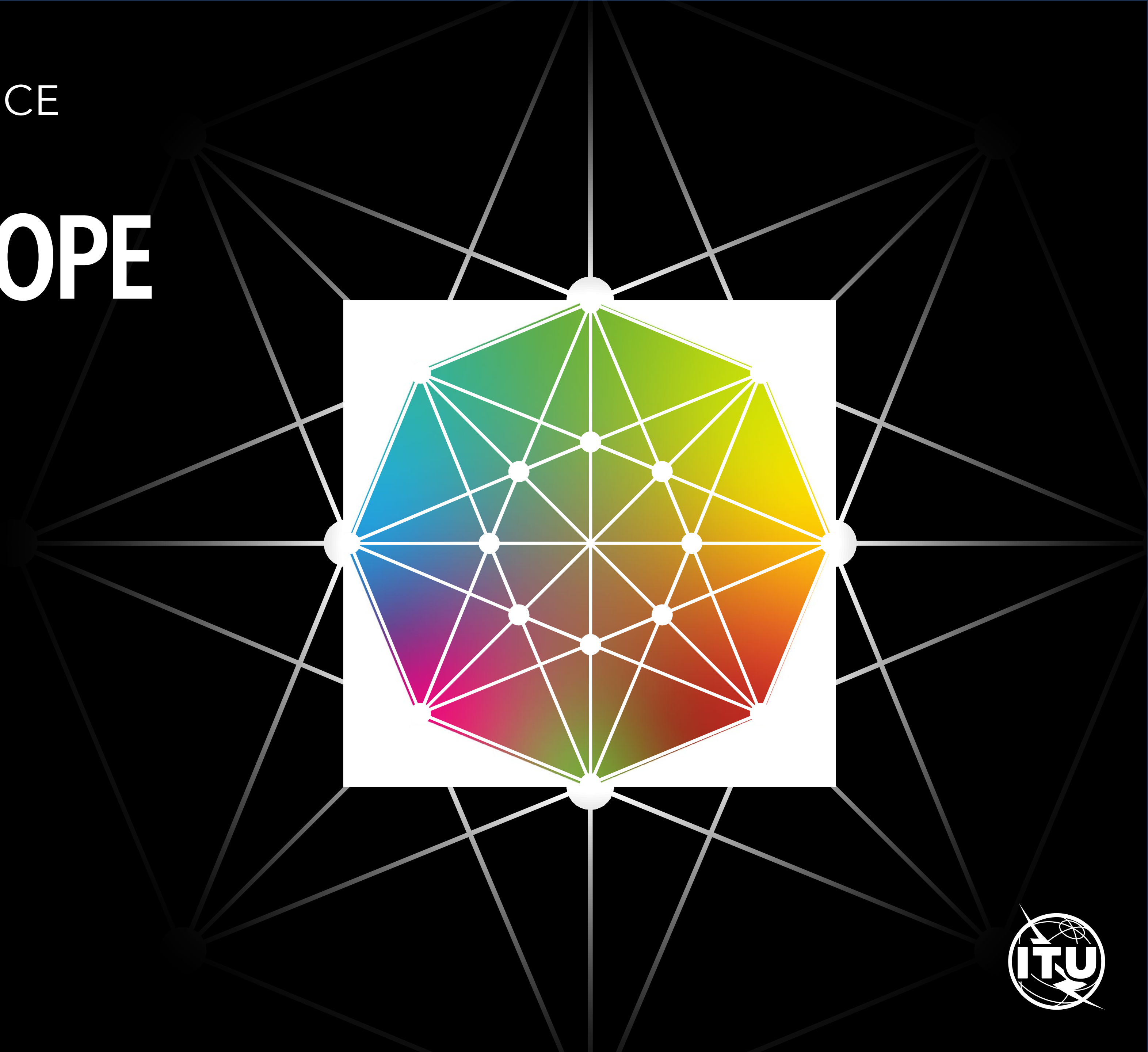
16TH ITU ACADEMIC CONFERENCE

ITU KALEIDOSCOPE
GENEVA 2026

AI and frontier technologies for good

7-9 July 2026
Geneva, Switzerland

Hosted by AI for Good Global Summit



HIERARCHICAL MULTI-OBJECTIVE LEARNING FOR CONTEXT-AWARE 5G RAN SLICE RESOURCE ALLOCATION

09 July 2026

Charles Ssengonzi

Ericsson & University of KwaZulu Natal South Africa

Session #6.4

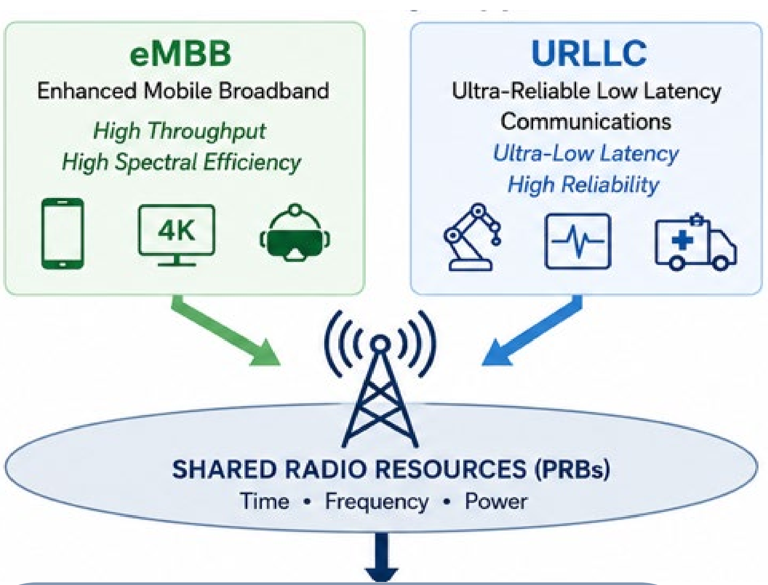
UNIVERSITY OF KWAZULU-NATAL™
INYUVESI YAKWAZULU-NATALI

ERICSSON

ITU KALEIDOSCOPE GENEVA2026

Motivation & Research Gap

5G/6G Networks Must Simultaneously Support Conflicting Service Requirements



Key Challenges

- Resource Contention between eMBB & URLLC slices
- Dynamic & Unpredicted Traffic Demand
- Non-stationary & time-varying network conditions
- Conflicting QoS & SLA requirements
- Need for autonomous and SLA-aware resource orchestration

Existing Approaches

- **Static RAN Slicing**: Fixed resource partitioning [1]
- **3GPP Dynamic Multiplexing**: Puncturing and superposition [2]
- **Flat RL/MORL**: Single-timescale optimization[3]
- **Rule-Based Scheduling**: Proportional Fair and latency-aware schedulers[4]

Limitation of Existing Approaches

- Fixed or single-timescale decision making
- Limited adaptability to dynamic traffic conditions
- Difficulty balancing throughput and latency objectives

KEY RESEARCH QUESTIONS

? **RQ1: Can hierarchical AI dynamically allocate RAN resources between eMBB and URLLC slices while satisfying SLA requirements?**

? **RQ2: Can context-aware reward adaptation improve learning stability under dynamic network conditions?**

? **RQ3: Can O-RAN-aligned hierarchical control improve network efficiency while maintaining URLLC reliability guarantees?**

Research Objective

- Develop a Context-Aware Hierarchical AI Framework for adaptive O-RAN slice resource orchestration.

Two-Timescale HRW Framework for O-RAN Slice Resource Orchestration

Non-RT RIC (Slow Timescale ≥ 1s)

Long Term Reward & Policy Adaptation
- Context-Aware Reward Adaptation
- Utility Prediction (rApp)
- Policy Adaptation

NEAR-RT RIC (Intermediate Timescale 10 ms – 1 s)
- Slice Orchestration (xApps)
- Policy Translation/Policy Enforcement
- Real-Time Control Coordination

O-RAN
Intelligent Controller

E2 Interface

KPI Feedback /Network Telemetry

Network Context Monitoring

CQI/SINR | Queue State | HOL Delay | Traffic demand | Resource utilization

eMBB Slice
- High Throughput traffic
- High Spectral Efficiency

Real-Time Resource Allocation
gNode B O-DU/ O-CU Scheduler (Fast Timescale < 10 ms)

Dynamic PRB Allocation | User Scheduling | Resource Assignment

URLLC Slice
- Ultra-Low Latency
- High Reliability

Key Innovations
- O-RAN-aligned hierarchy
- Context-aware reward adaptation
- Two-timescale control
- SLA-aware multi-objective learning

The proposed HRW framework separates long-term policy adaptation from real-time resource allocation through a hierarchical O-RAN control architecture.

CMDP-Based Learning Framework

State space (S)

- ✓ CQI/SINR
- ✓ Queue State
- ✓ Traffic Demand
- ✓ User Mobility
- ✓ PRB Utilization

Action Space (A)

Near-RT RIC(policy level)

- ✓ Slice Resource Orchestration Policy

O-DU Scheduler (execution level)

- ✓ PRB Allocation
- ✓ Power Allocation
- ✓ User Resource Assignment
- ✓ User Scheduling
- ✓ URLLC Preemption Decisions

Reward Function (R)

- ✓ eMBB Utility U^e
- ✓ URLLC Utility U^u
- ✓ Adaptive Reward Weights (w_τ^e, w_τ^u)
- ✓ Preemption Penalty

Constraints

- ✓ eMBB QoS Guarantees
- ✓ URLLC Reliability Bounds
- ✓ PRB Capacity Limits
- ✓ Power Budget constraints

Two-Timescale Learning

Slow Timescale (Non-RT RIC)

- ✓ Reward Adaptation

Fast Timescale (O-DU)

- ✓ Resource Scheduling
- ✓ PRB Allocation

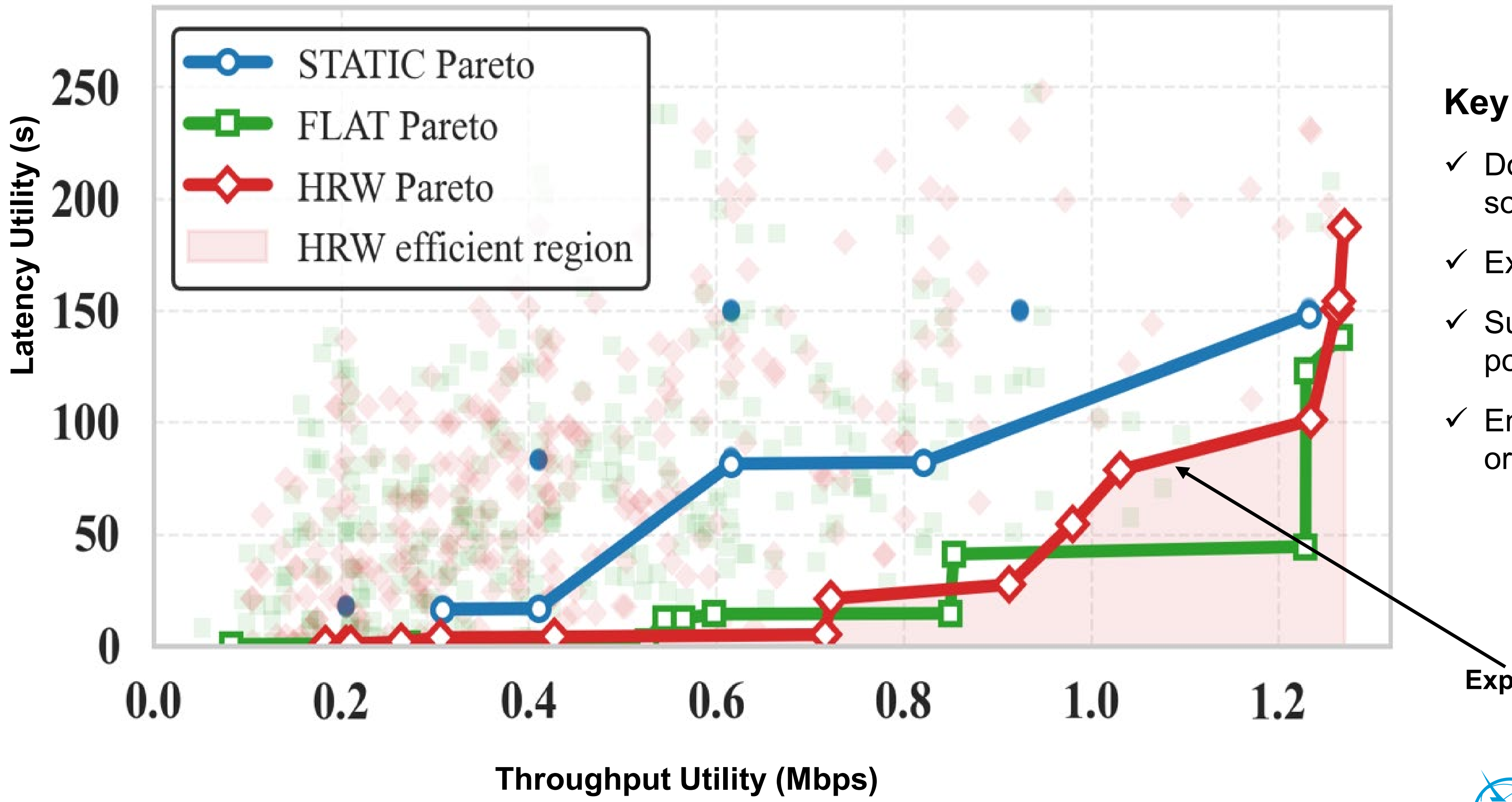
Global Objective

- $\max_{w_\tau} \mathcal{L}_{global}(\theta_\tau, w_\tau) = w_\tau^e U^e(\tau) + w_\tau^u U^u(\tau) - \rho_\tau \sum_{b \in \mathcal{B}} z_b(\tau)$

Adaptive Multi-Objective Optimization

- eMBB Utility + URLLC Utility – Preemption Cost $\left(\sum_{b \in \mathcal{B}} z_b(\tau)\right)$
- $w_\tau^e + w_\tau^u = 1, w_\tau^e, w_\tau^u \geq 0$ Adaptive reward preferences remain normalized and feasible.

HRW Achieves Superior Throughput-Reliability Trade-Offs

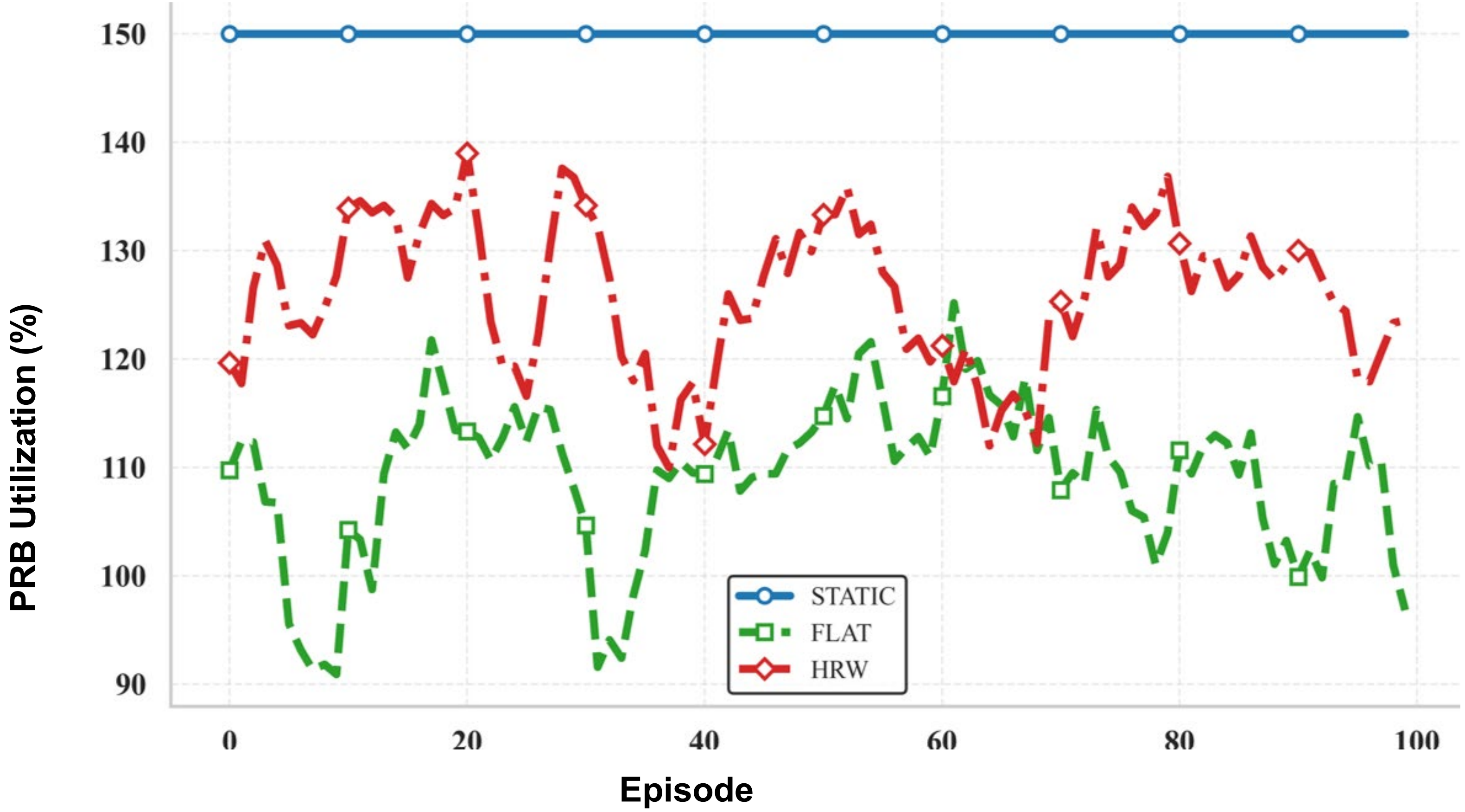


Key Findings

- ✓ Dominates Static and Flat RL solutions.
- ✓ Expands the achievable utility region.
- ✓ Supports eMBB/URLLC operating points.
- ✓ Enables adaptive resource orchestration.

- *O-RAN-aligned hierarchical control outperforms static and flat RL approaches*

HRW Improves Resource Utilization through Dynamic PRB Adaptation



Key Findings

- ✓ Continuous PRB adaptation under dynamic traffic.
- ✓ Higher average resource utilization than Flat RL.
- ✓ Improved responsiveness to changing network conditions.
- ✓ More efficient resource usage than static allocation.

- *Context-aware reward adaptation improves learning stability and resource allocation responsiveness.*

Conclusion

Key Contributions

- ✓ Proposed a Two-Timescale Hierarchical Reward Weighting (HRW) framework for O-RAN slicing.
- ✓ Integrated context-aware reward adaptation through the Non-RT RIC.
- ✓ Separated long-term policy adaptation from real-time resource allocation.
- ✓ Demonstrated improved throughput-latency trade-offs and URLLC SLA compliance.

Key Takeaways

- ✓ O-RAN-aligned hierarchy enables scalable AI-driven slicing.
- ✓ Context-aware reward adaptation improves responsiveness to dynamic traffic.
- ✓ Two-timescale learning balances throughput, latency, and reliability objectives.

HRW maps hierarchical reward adaptation onto the O-RAN control hierarchy, separating slow-timescale policy adaptation from fast-timescale resource allocation to enable autonomous and SLA-aware RAN slicing.

Industry Relevance

- Beyond the performance gains, an important implication of this work is its alignment with the emerging AI-native 6G vision, where intelligence is embedded directly into network control and resource orchestration functions.

Future Work

✓Multi-cell O-RAN deployments with inter-cell interference coordination, mobility management, and CoMP transmission.

✓Distributed and Federated Multi-Agent Reinforcement Learning for scalable slice orchestration across multiple network domains.

✓Benchmarking against advanced DRL architectures, including PPO and Hierarchical Actor-Critic methods.

The proposed HRW framework demonstrates how context-aware, two-timescale AI control can enable AI-native and autonomous resource orchestration in future O-RAN and 6G networks

Thank you!

References

[1] Foukas, X., Patounas, G., Elmokashfi, A. and Marina, M.K., 2017. Network slicing in 5G: Survey and challenges. *IEEE communications magazine*, *55*(5), pp.94-100.

[2] Anand, A., De Veciana, G. and Shakkottai, S., 2020. Joint scheduling of URLLC and eMBB traffic in 5G wireless networks. *IEEE/ACM Transactions On Networking*, *28*(2), pp.477-490.

[3] Alsenwi, M., Tran, N.H., Bennis, M., Pandey, S.R., Bairagi, A.K. and Hong, C.S., 2021. Intelligent resource slicing for eMBB and URLLC coexistence in 5G and beyond: A deep reinforcement learning based approach. IEEE Transactions on Wireless Communications, 20(7), pp.4585-4600.

[4] Bairagi, A.K., Munir, M.S., Alsenwi, M., Tran, N.H., Alshamrani, S.S., Masud, M., Han, Z. and Hong, C.S., 2020. Coexistence mechanism between eMBB and uRLLC in 5G wireless networks. *IEEE transactions on communications*, *69*(3), pp.1736-1749.