Melike (Mel) Erol-Kantarci, PhD

Chief Cloud RAN AI\ML Data Scientist at **Ericsson**

Canada Research Chair in AI-Enabled Next-Generation Wireless Networks & Associate Professor at **University of Ottawa**

IEEE ComSoc Distinguished Lecturer

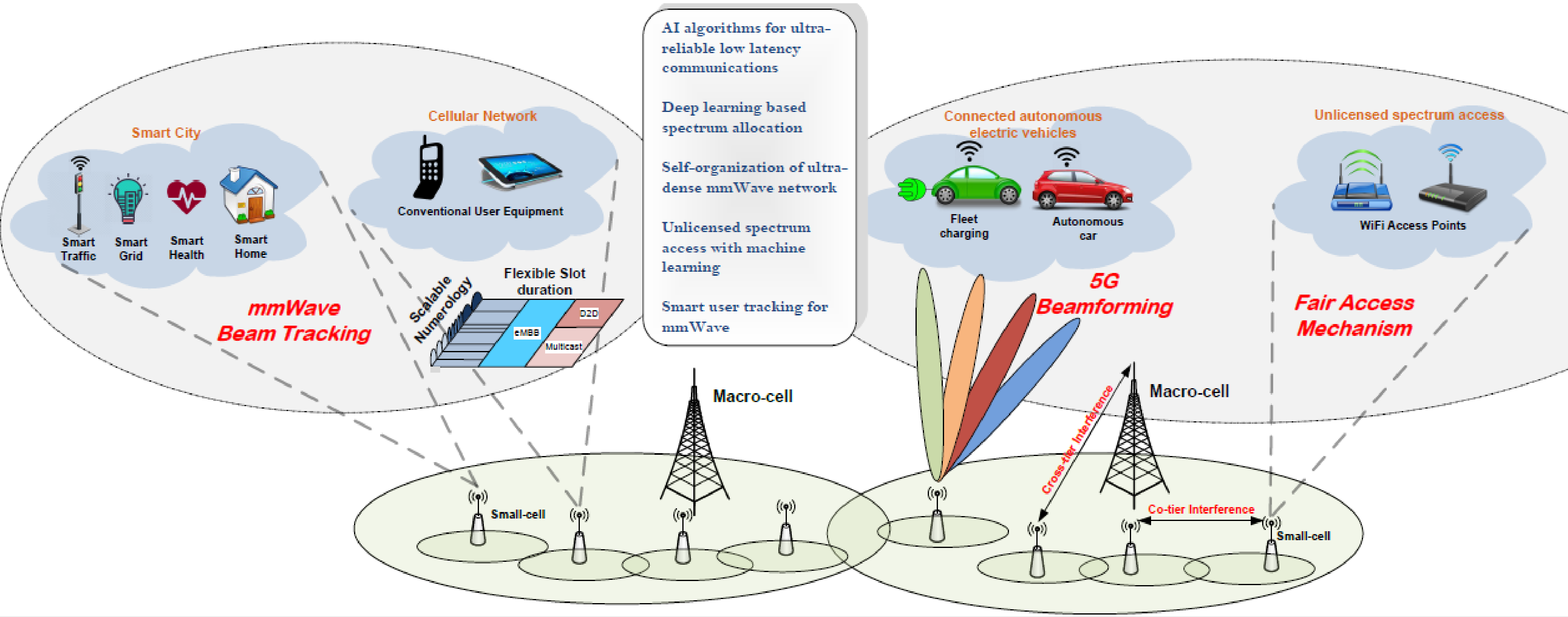
Hosted by Oregon, Santa Clara Valley, Seattle Foothill, San Fernando Valley Chapters

Deep and Reinforcement Learning in 5G and 6G Networks

NETCoRE
Networked Systems and Communications Research Lab

uOttawa

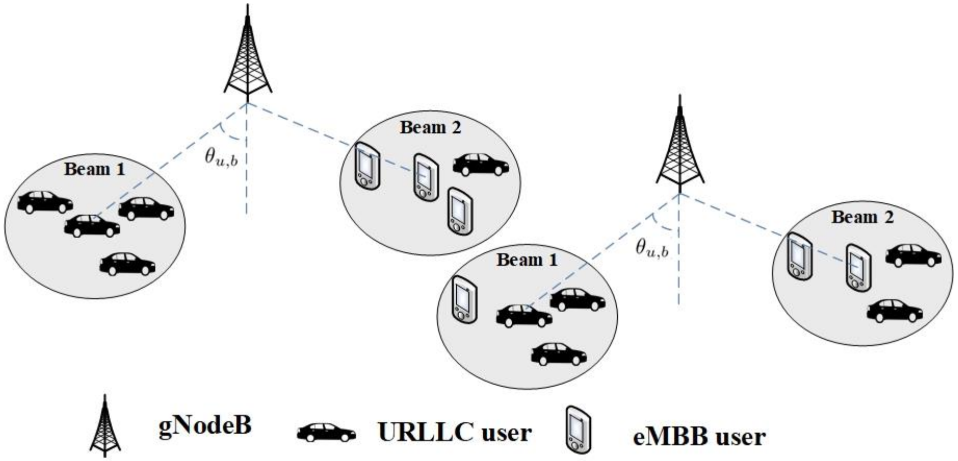
AI-enabled 5G and 6G



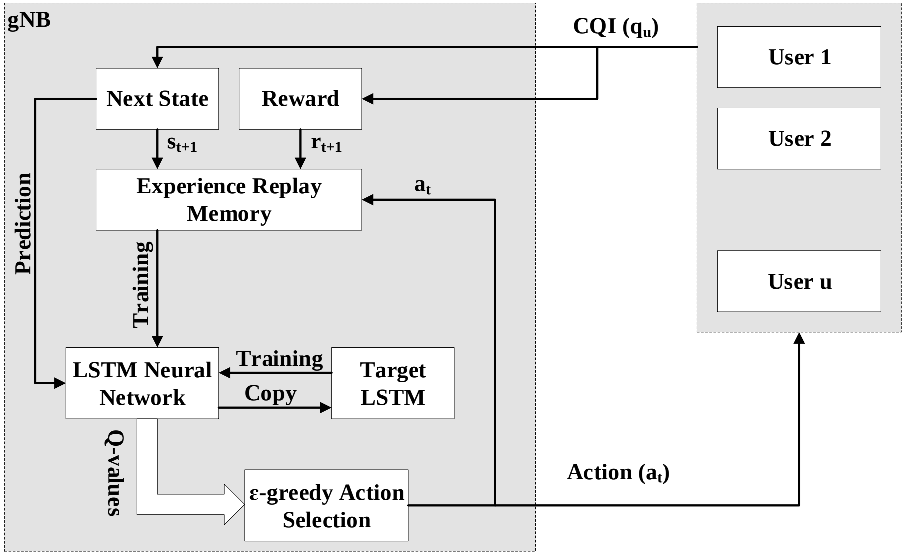
M. Elsayed, M. Erol-Kantarci, "AI-enabled Future Wireless Networks: Challenges, Opportunities and Open Issues," IEEE Vehicular Technology Magazine, Special Issue on 6G: What is Next?, Volume: 14 , Issue: 3 , pp. 70-77, Sept. 2019.

Radio Resource and Beam Management in 5G mmWave

- With mobility, cluster patterns vary
  - Beam management and radio resource allocation becomes more challenging
- Clustering with DBSCAN
- Deep Reinforcement Learning



System model of mmWave network.



Conceptual diagram of LSTM-based Deep Q-learning.

M. Elsayed and M. Erol-Kantarci, "Radio Resource and Beam Management in 5GmmWave Using Clustering and Deep Reinforcement Learning", IEEE Globecom, Dec. 2020.

Deep Q-learning with DBSCAN (DQLD)

- Online clustering using DBSCAN:
  - Clusters users that are close to each other and can be covered by a single beam.
  - DBSCAN helps to find the least number of beams.

- Resource block allocation:
  - LSTM-based deep Q-learning is used to perform RB allocation within each beam.

Latency and Sum Rate

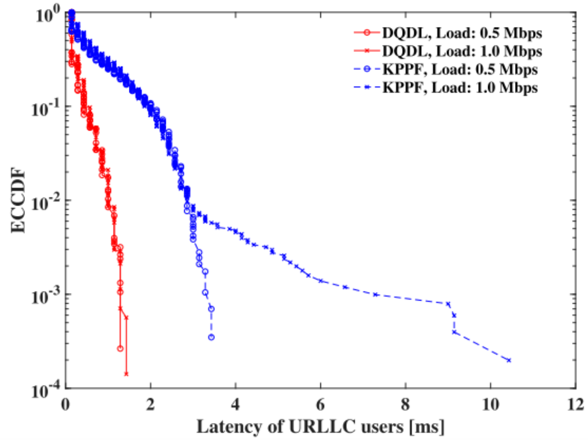


Fig. 3: Latency of URLLC users versus total URLLC offered load ([0.5, 1] Mbps).

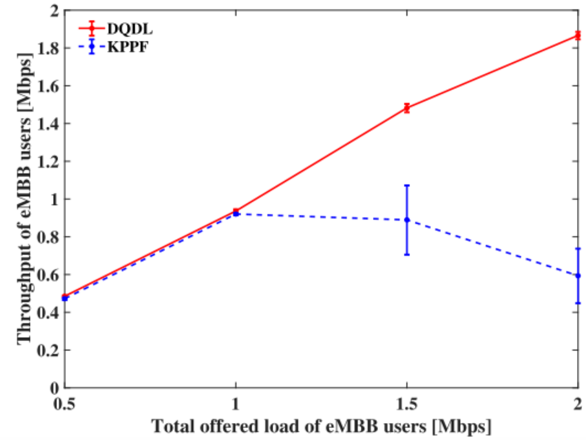


Fig. 7: Sum rate of eMBB users versus total eMBB offered load.

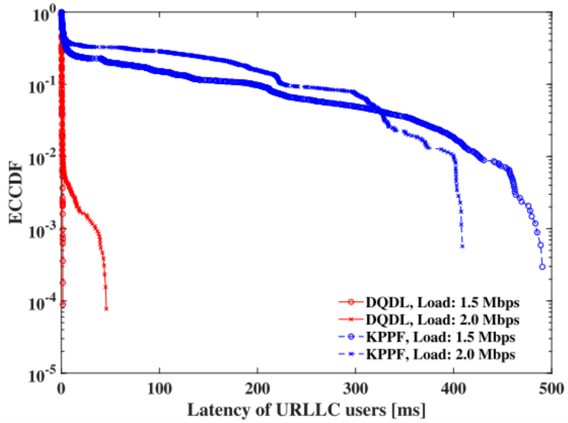


Fig. 4: Latency of URLLC users versus total URLLC offered load ([1.5, 2] Mbps).

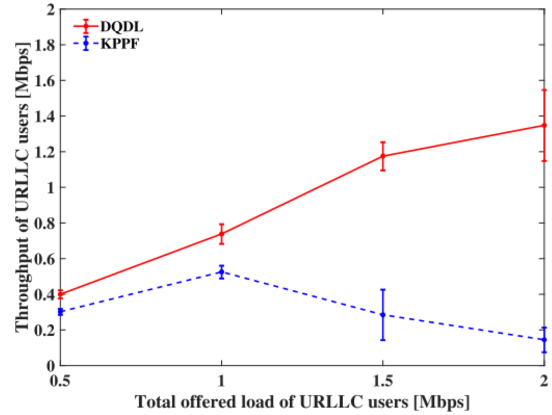
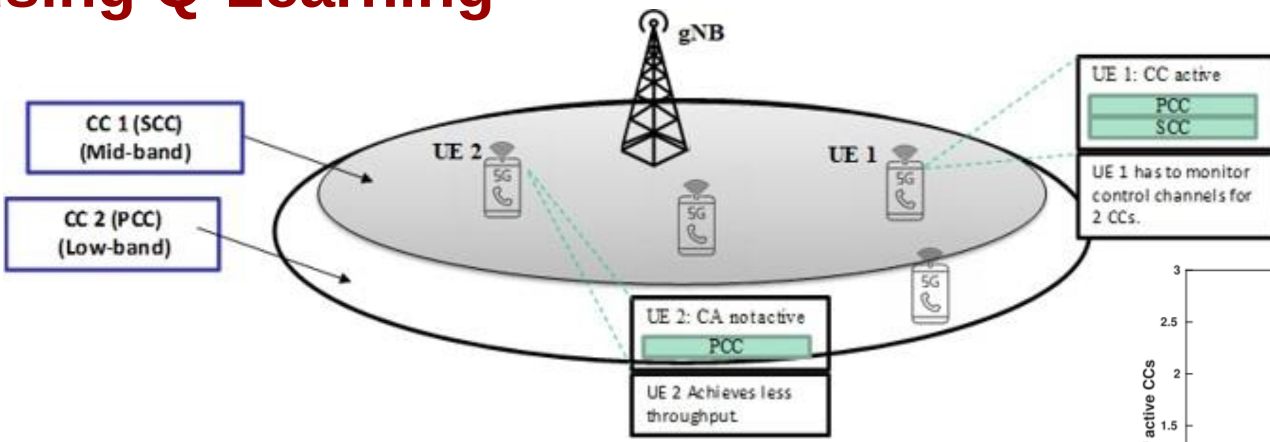
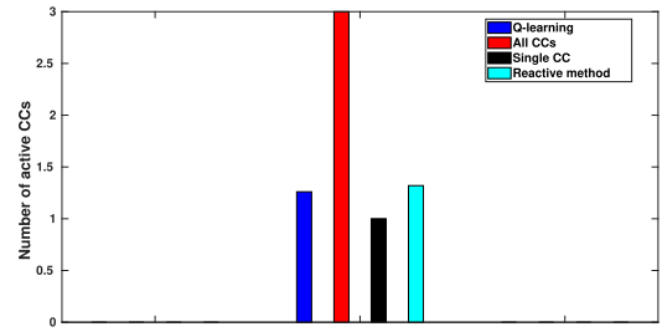


Fig. 6: Sum rate of URLLC users versus total URLLC offered load.

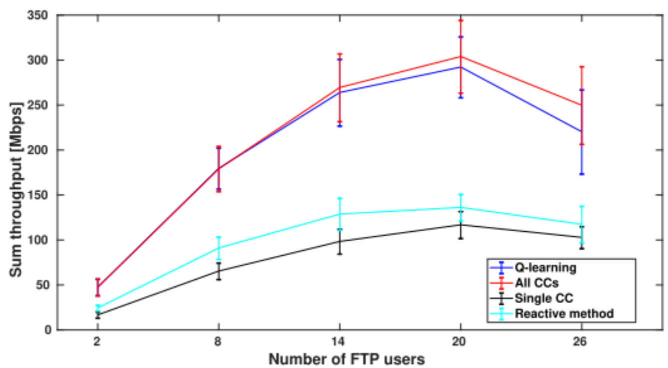
Proposed algorithm is compared to a baseline that uses K-means for clustering and priority-based proportional fairness for RA. Simulation results show that the proposed algorithm outperforms the baseline in terms of latency, reliability and rate.

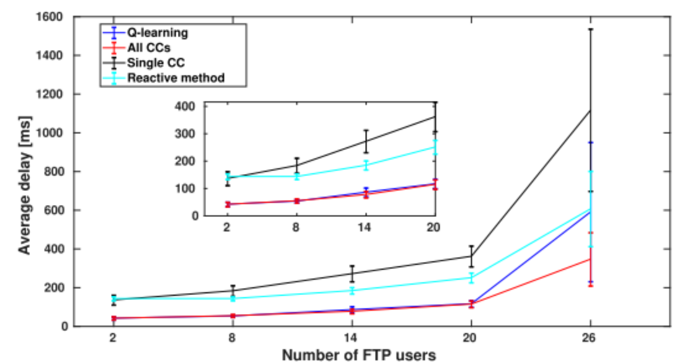
Carrier Aggregation with UE Energy-Efficiency using Q-Learning





✓ Comparable throughput and latency with less energy and control overhead

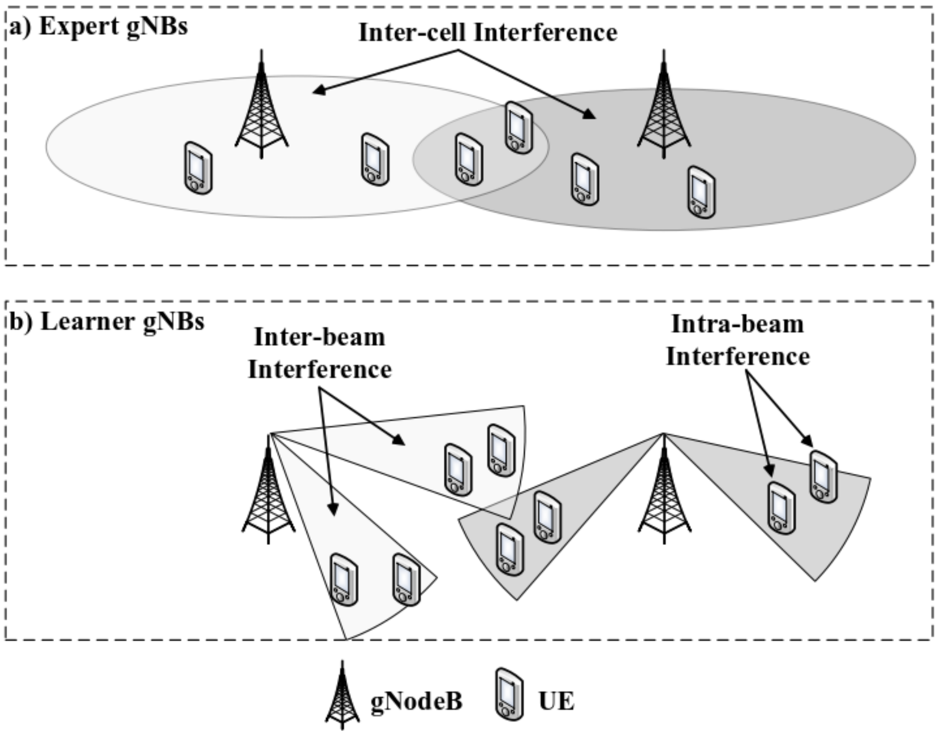




M. Elsayed, R. Joda, H. Abou-zeid, R. Atawia, A. Bin Sediq, G. Boudreau, M. Erol-Kantarci, "Reinforcement Learning Based Energy-Efficient Component Carrier Activation-Deactivation in 5G," IEEE Globecom, 2021.

Transfer Reinforcement Learning

- Expert gNBs
  - User-cell association solely
  - Q-learning
- Learner gNBs
  - Joint user-cell association and selection of number of beams
  - Transfer Q-learning



M. Elsayed, M. Erol-Kantarci, H. Yanikomeroglu, "Transfer Reinforcement Learning for 5G-NR mm-Wave Networks," IEEE Transactions on Wireless Communications, vol. 20, no. 5, May 2021.

Convergence of TL

Convergence of expert gNB

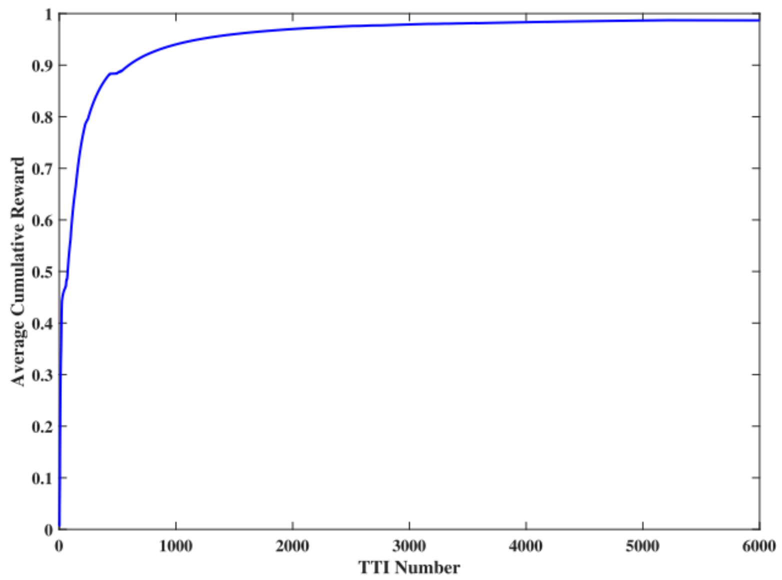


Fig. 4. Convergence of expert gNBs represented by the average cumulative reward.

Convergence of Learner gNB

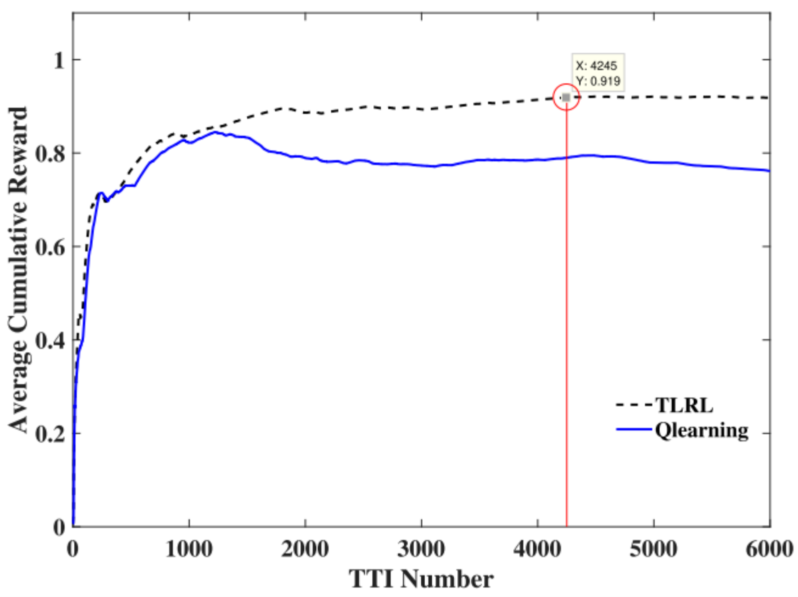
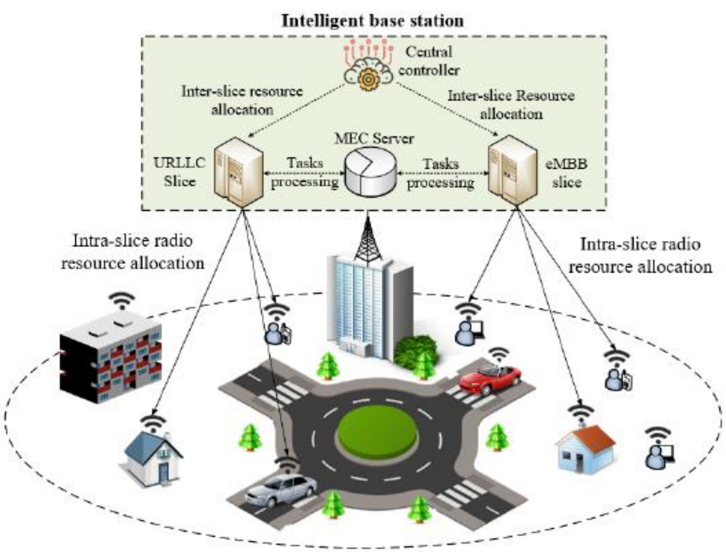


Fig. 5. Convergence of learner gNBs represented by the average cumulative reward. Total offered load is 1.3 Mbps.

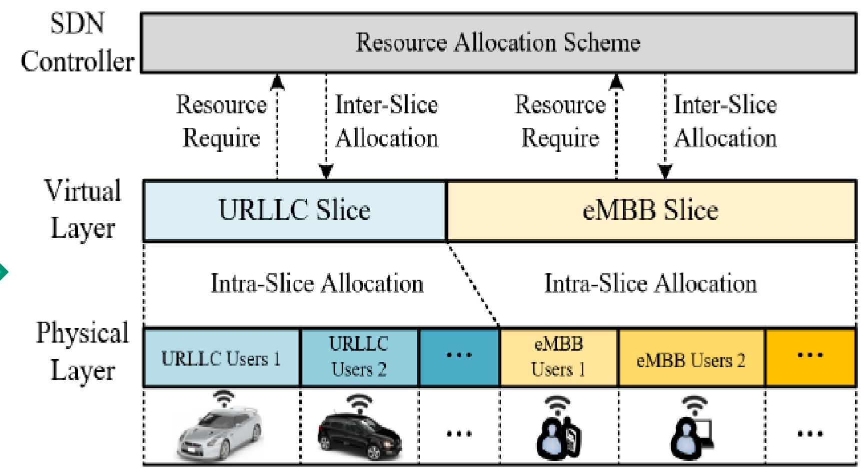
Knowledge Transfer based Resource Allocation for RAN Slicing

- We apply a two-step resource allocation scheme: inter-slice phase and intra-slice phase.
- Resources: PRB and computational resources



Proposed system architecture

Resource allocation scheme



Resource allocation scheme

H. Zhou, and M. Erol-Kantarci, Knowledge Transfer based Radio and Computation Resource Allocation for 5G RAN Slicing in Proceedings of 2022 IEEE conference on CCNC , pp.1-6, Jan. 2022.

Knowledge Transfer based Resource Allocation for RAN Slicing

- eMBB and URLLC slices. The eMBB slice intends to maximize the throughput, while the URLLC slice requires a lower latency.

- In knowledge transfer reinforcement learning, there are two phases: knowledge transfer phase and learning phase.

$$\max \quad w^{embb} b_j^{embb,avg} + w^{urllc}(d^{tar} - d_j^{urllc,avg})$$

$$\text{s.t.} \quad b_j^{embb,avg} = \frac{\sum_{u \in \mathcal{M}_j^{embb}} b_{j,u}^{embb}}{|\mathcal{M}_j^{embb}|}$$

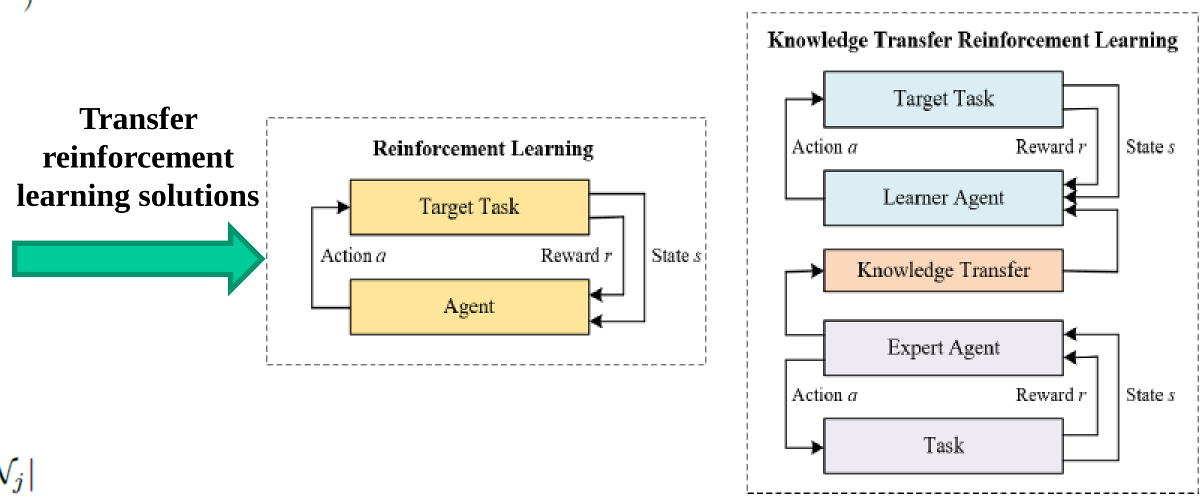
$$b_j^{urllc,avg} = \frac{\sum_{v \in \mathcal{M}_j^{urllc}} d_{j,v}^{urllc}}{|\mathcal{M}_j^{urllc}|}$$

$$(1)\,(2)\,(3)\,(4)\,(6)$$

$$\sum_{u \in \mathcal{M}_j^{embb}} x_{j,u,r'} + \sum_{v \in \mathcal{M}_j^{urllc}} x_{j,v,r'} = 1$$

$$\sum_{r' \in \mathcal{N}_{j'}} \Big(\sum_{u \in \mathcal{M}_j^{embb}} x_{j,u,r'} + \sum_{v \in \mathcal{M}_j^{urllc}} x_{j,v,r'} \Big) \leq |\mathcal{N}_j|$$

$$C_j^{embb} + C_j^{urllc} \leq C_j$$



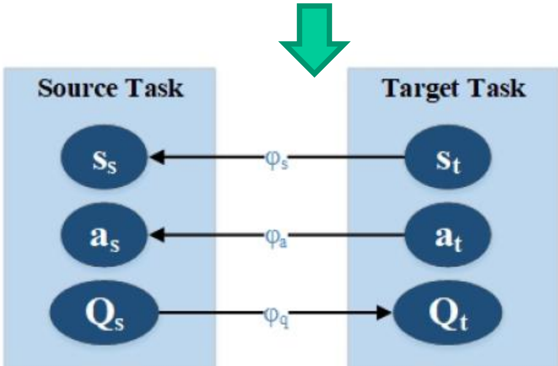
Knowledge transfer reinforcement learning

NETCoRE Networked Systems and Communications Research Lab

Knowledge Transfer based Resource Allocation for RAN Slicing

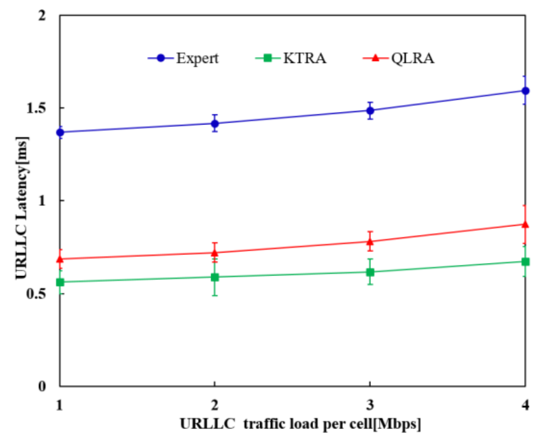
- We use the prior knowledge of experts to improve the exploration efficiency: finding a specific Q-value of expert agent to represent the potential reward of learner actions.

$$Q^T(\mathcal{F}(s_{l,t}), \mathcal{F}'(a_{l,t})) = Q^e(s_e, a_e).$$

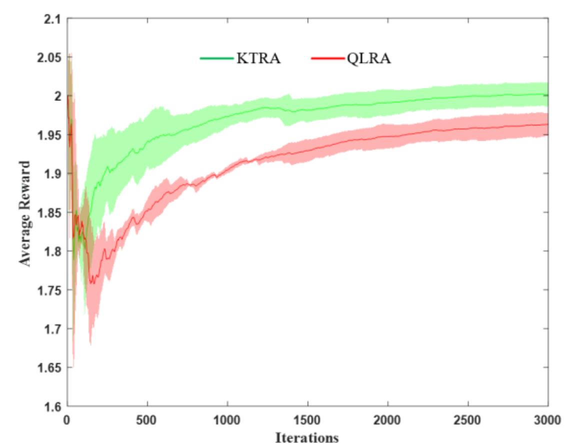


Map function definition

Faster convergence and better network metrics

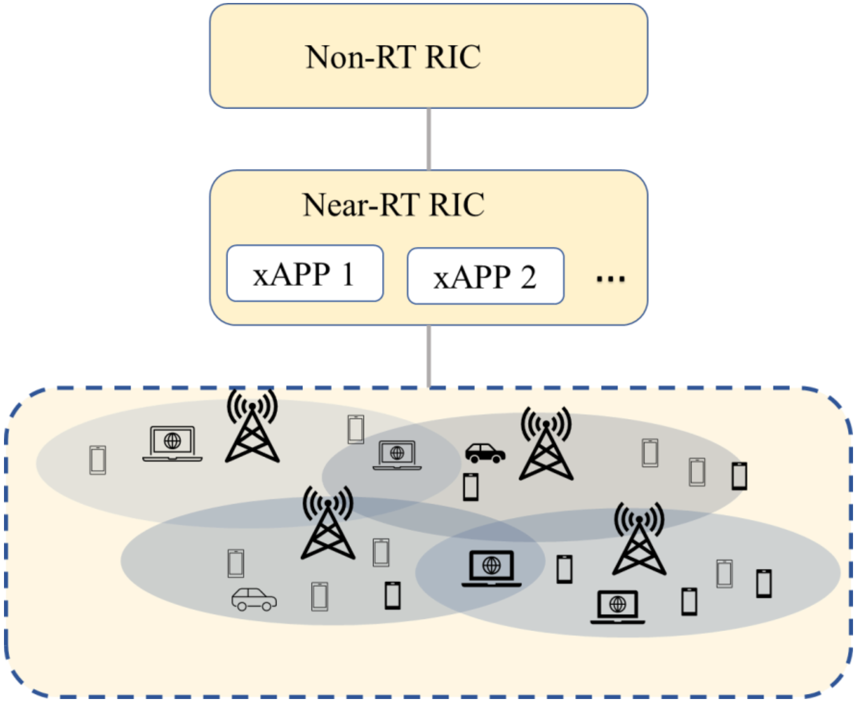


Average URLLC latency [ms] under various URLLC traffic.



Convergence performance of KTRA (knowledge transfer) and QLRA (Q-learning).

Team Learning in DRL-based xApps in O-RAN



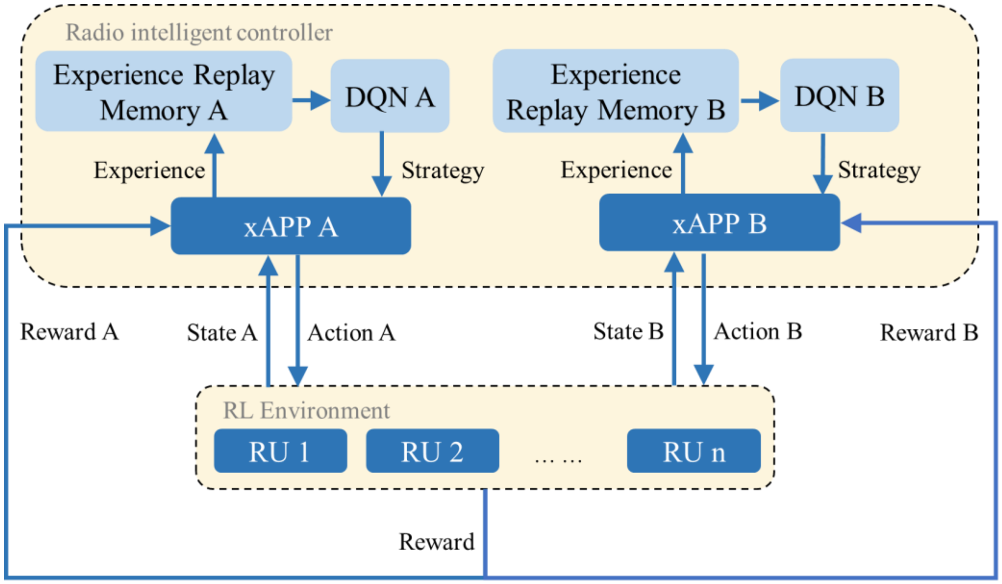
- Multiple network functions in the Near-RT RIC deployed by different vendors
- There can be conflicts between xAPPs
- <u>*How to address these conflicts?*</u>

- Team learning: A team of agents are in the same environment and share part of the observational information.
- Learn and choose actions in a distributed manner and cooperate for the same team goal.
- Team goal: maximize the system throughput.

H. Zhang, H. Zhou, M. Erol-Kantarci, "Team Learning-based Resource Allocation for Open Radio Access Networks (O-RAN), in Proc. of IEEE ICC, May 2022.

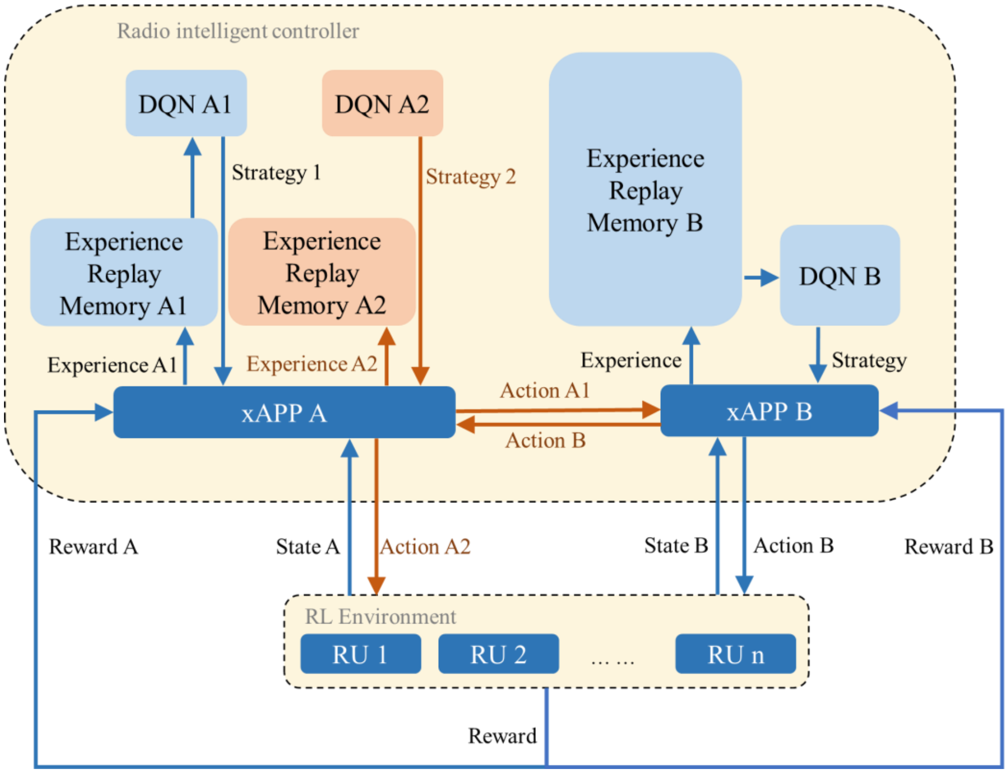
DRL-based xAPP Interactions without Team Learning

- Potential problem: when an xAPP acts, it does not take into account the actions of other xAPPs.
  - The selected action may be the optimal choice in its own view, but the performance of other xAPPs may be affected.

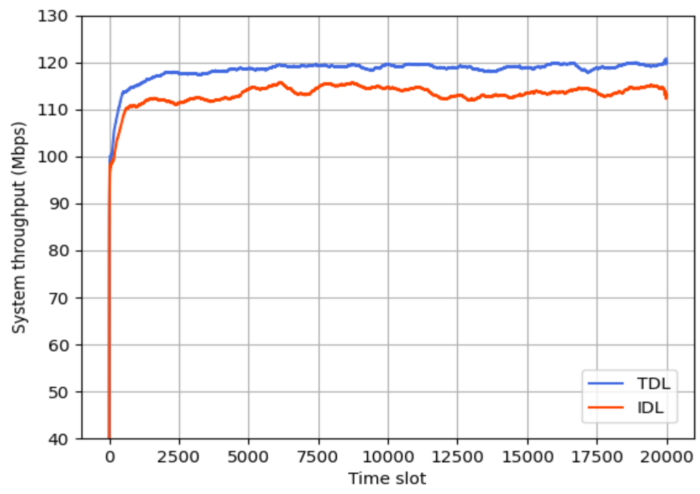


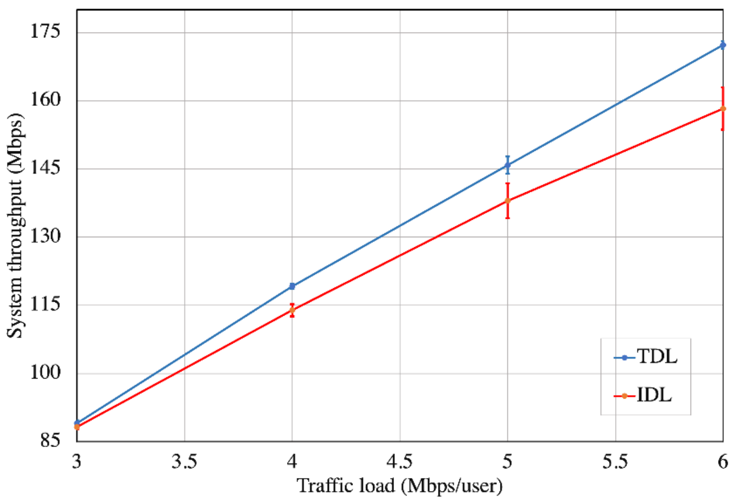
DRL-based xAPP Interactions with Team Learning

- xAPPs will include the actions of other xAPPs in its own state



Performance Evaluation





TDL: team deep Q-learning
IDL: independent deep Q-learning

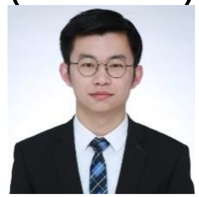
Varying traffic load, V=20m/s
8.8% higher throughput when traffic load = 6 Mbps

NETCORE Lab - http://www.site.uottawa.ca/~merolka2/NETCORE

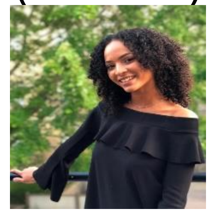
Roghayeh Joda
(Researcher)

Omer Melih Gul
(Postdoc)

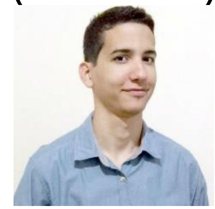
Hao Zhu
(PhD Student)



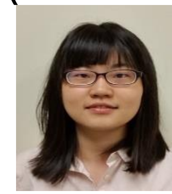
Hind Mukhtar
(PhD Student)



Ycaro Dantas
(MASc student)



Tiffany Nien Fang Cheng
(MSc Student)



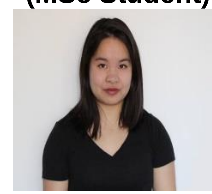
Turgay Pamuklu
(Postdoc)

Mikhak Samadi
(PhD Student)

Fahime Khoramnejad
(PhD Student)

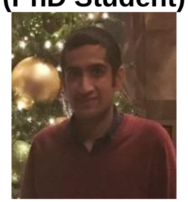
Han Zhang
(PhD Student)

Anne Catherine Nguyen
(MSc Student)



Mohammad Akbari
(Postdoc)

Mohammad Sadeghi
(PhD Student)



Pedro Iturria Rivera
(PhD Student)

Arafat Habib
(PhD Student)

Yujie Yao
(MASc Student)

NSERC CRSNG

NSF

Mitacs

Ontario Centres of Excellence

ENCQOR 5G

ALWAYS LOOKING FOR NEW POSTDOCS AND PHD STUDENTS!

NETCORE
Networked Systems and Communications Research Lab

Thank you!

Melike.erolkantarci@uottawa.ca
Melike.erol-kantarci@ericsson.com

For more talks subscribe to my YouTube channel:
https://www.youtube.com/channel/UCMSyl_tepY0s7LtJNgq-kiw