REINFORCEMENT LEARNING IN WIRELESS COMMUNICATION NETWORKS

Tufts LiNKS – Laboratory for communIcation in NetworKed Systems

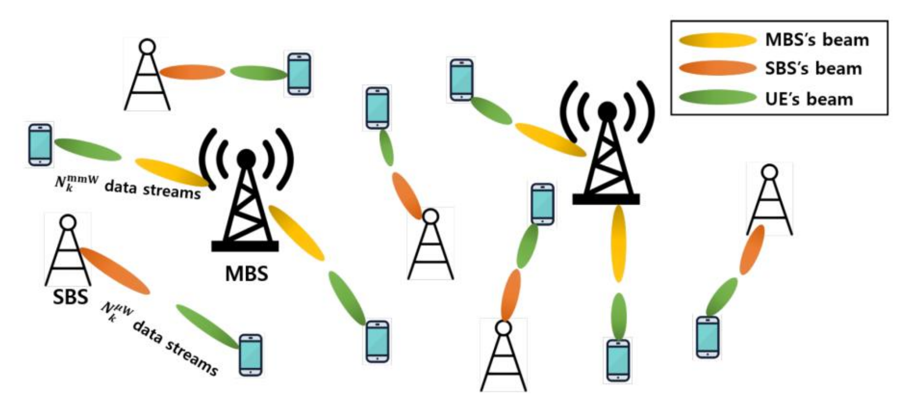
Mai Vu
Tufts University

MathWorks Research Summit – June 2024

Modern Wireless Systems

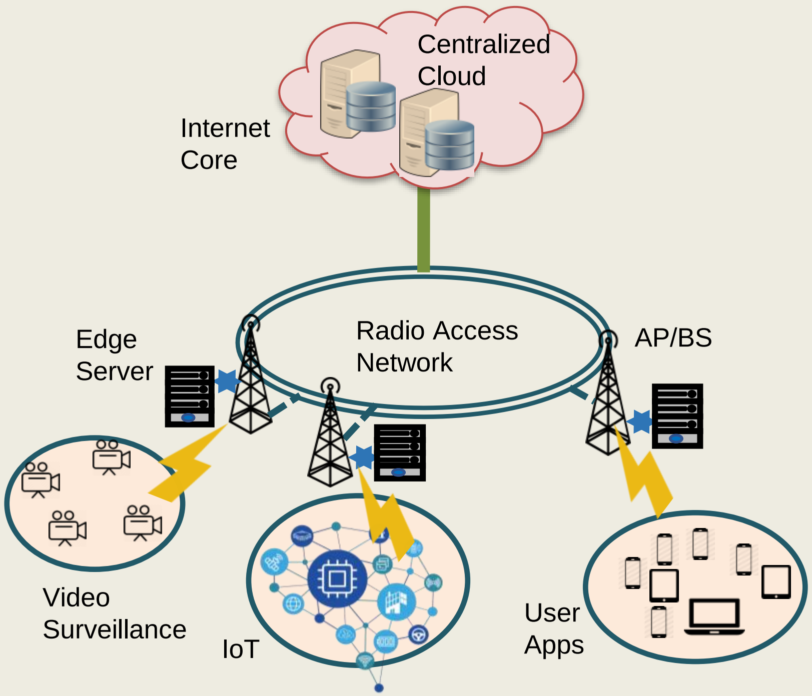
5G and 6G systems

- Higher frequency ranges: mmWave and sub-THz communications
- Dense, dynamic networks
- Range of user mobilities



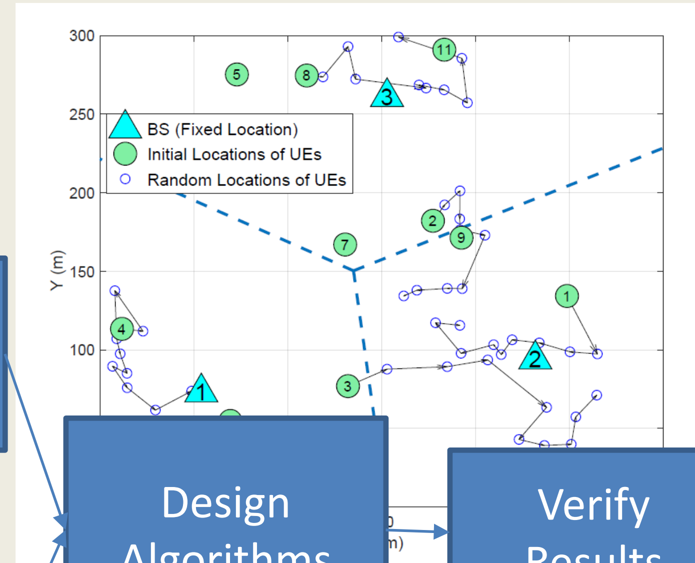
RL for resource management:

- Beamforming design and prediction:
  - Multicell multiuser mobile network
  - Integrated sensing (imaging) and communication
  - Reconfigurable intelligent surfaces
- Network management:
  - User-base association
  - Hand-over
  - Scheduling, load balancing
- Edge networking:
  - Caching
  - Computation offloading

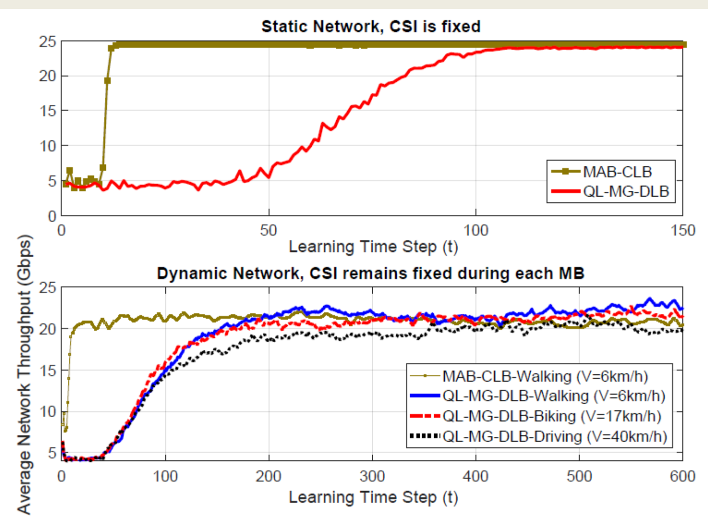


Research Methodology

$$\begin{aligned}
&\underset{\boldsymbol{\beta}(t)}{\text{maximize}} && U(\mathbf{r}(t)) \\
&\text{subject to} && \sum_{j \in \mathcal{J}} 1_{\beta_k(t)}(j) \leq 1, \quad \forall k \in \mathcal{K} \\
& && \sum_{k \in \mathcal{K}} 1_{\beta_k(t)}(j).n_k \leq D_j, \quad \forall j \in \mathcal{J}
\end{aligned}$$



Model the system using Math → Formulate the research problem → Optimization / Machine Learning → Design Algorithms → Verify Results



Algorithm 3: Centralized MAB Load Balancing User Association

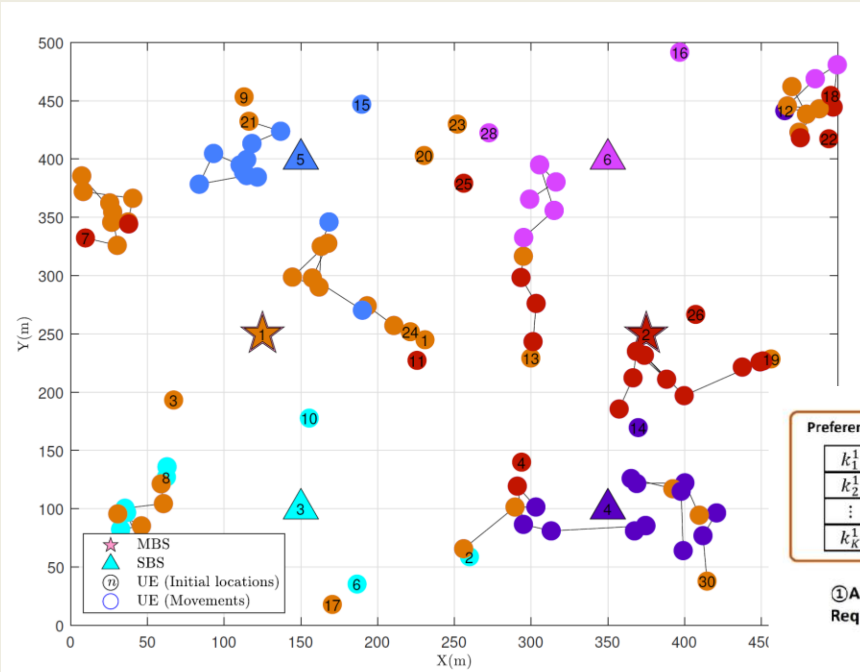
Input: Learning rates α, BSs' quota vector $\mathbf{q}$, initial reward matrix $\mathbf{\Gamma}^{(0)}$, initial matrix of number of BS selection $\mathbf{T}^{(0)} = \mathbf{0}$

for $t = 1 : T$ **do**
- **Central load-balancer (CLB):**
  - Applies UCB formula to obtain input reward matrix $\mathbf{\Omega} = \mathbf{\Gamma}^{(t-1)} + \sqrt{\frac{2\ln(t)}{\mathbf{T}^{(t-1)}}}$;
  - Executes Alg. 1 to obtain $\boldsymbol{\eta}^{(t)}$ and $\boldsymbol{\beta}$;
  - Informs VUEs of their $\eta^{(t)}_{k_v}$, and β_{k_v} if changed;
- **Each VUE** k_v:
  - Connects to BS $j = \eta^{(t)}_{k_v}$;
  - Receives reward $R^{(t)}_{k_v,j}$ and reports it to CLB;
- **CLB:**
  - Executes an updating rule: **BL** (Alg. 2), or **RTL** (Alg. 5), or **BGL** (Alg. 6) to obtain $\mathbf{\Gamma}^{(t)}$ and $\mathbf{T}^{(t)}$;

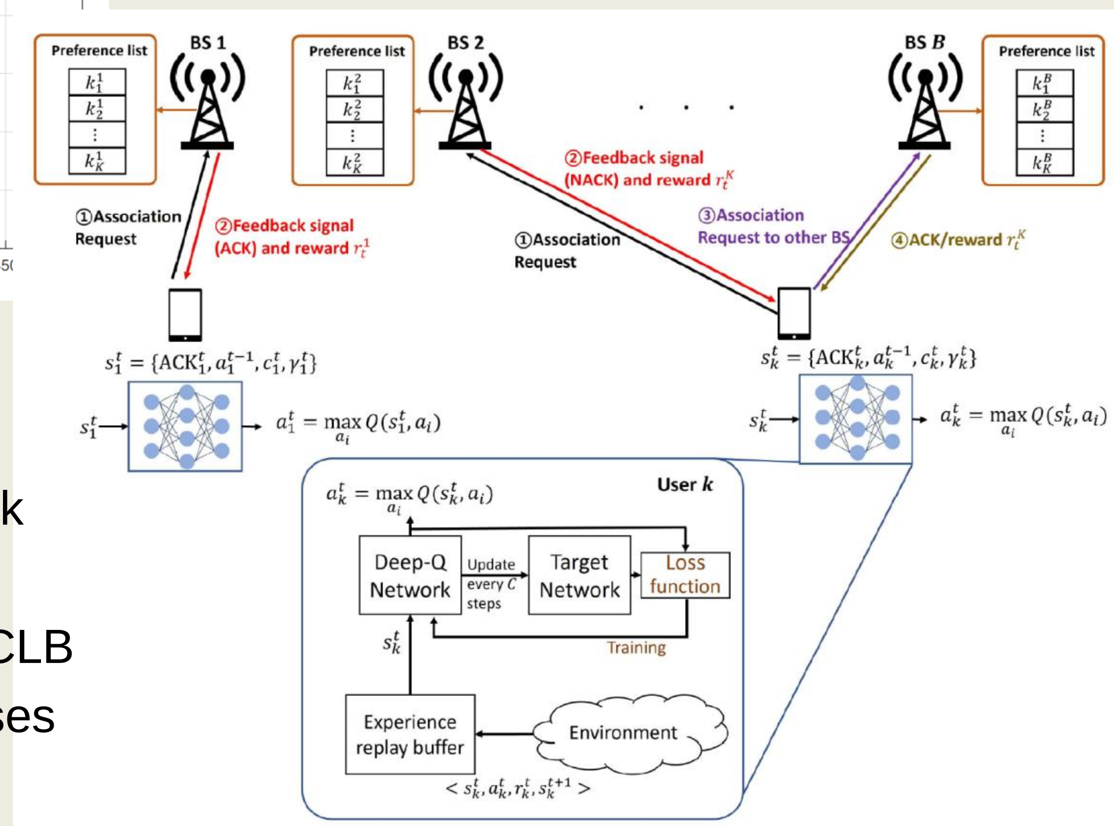
end

Output: Best-to-date association vector $\boldsymbol{\beta}$ (up to time step T), $\mathbf{\Gamma}^{(T)}$ and $\mathbf{T}^{(T)}$

RL for Association and Handover



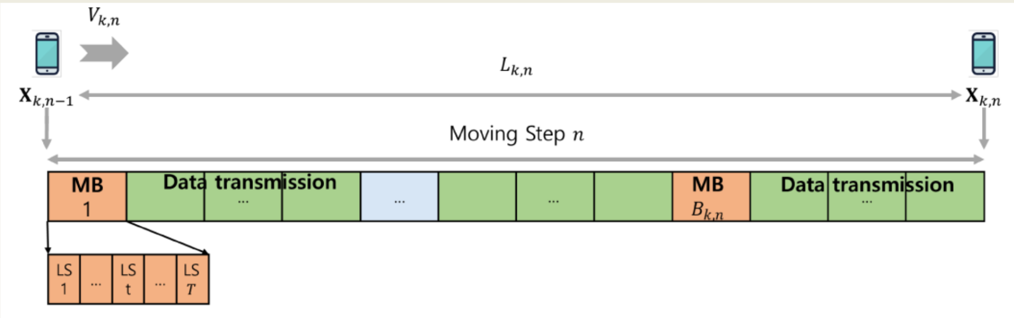
- Wireless network
  - 6 base stations
  - 20 users (UE), some are mobile
  - User's color shows association



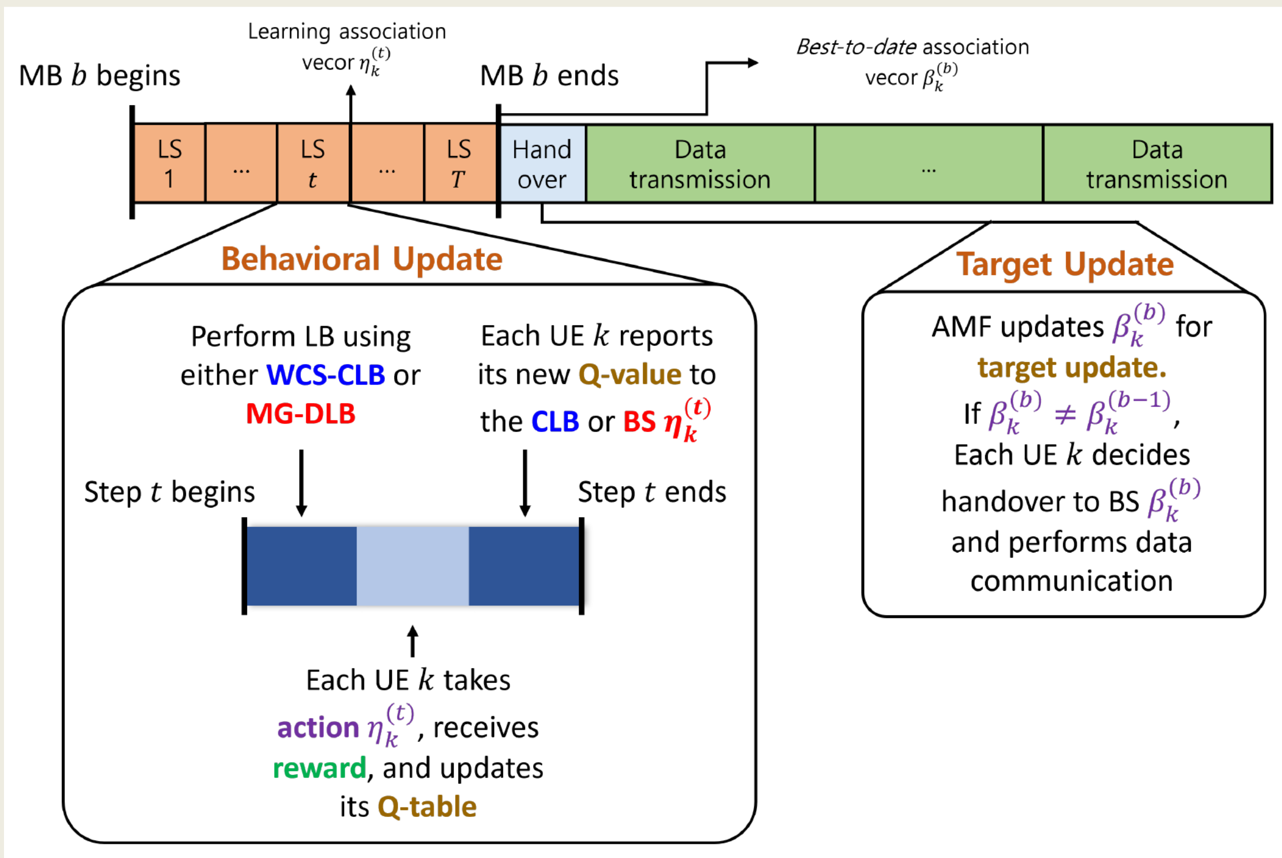
Designed a multi-agent DRL

- Each UE is an agent
- Each UE runs a Deep-Q Network
- Two versions of interaction:
  - Centralized: UEs talk with a CLB
  - Distributed: UEs talk with bases

How is RL integrated into the System?

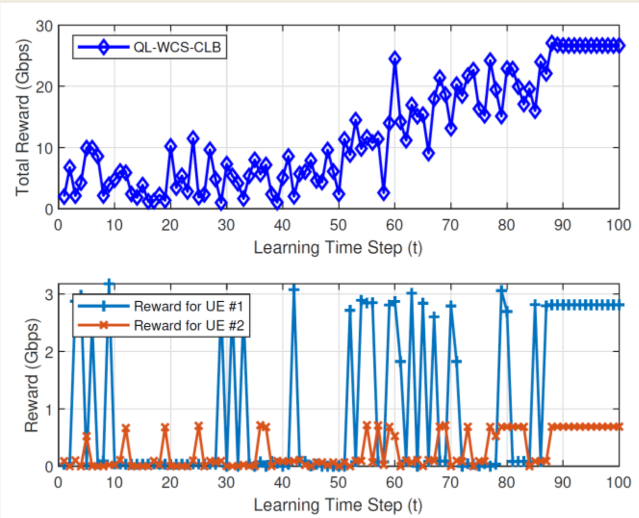


Alizadeh, Alireza, Byungju Lim, and Mai Vu. "Multi-Agent Q-Learning for Real-Time Load Balancing User Association and Handover in Mobile Networks." *IEEE Transactions on Wireless Communications* (2024).

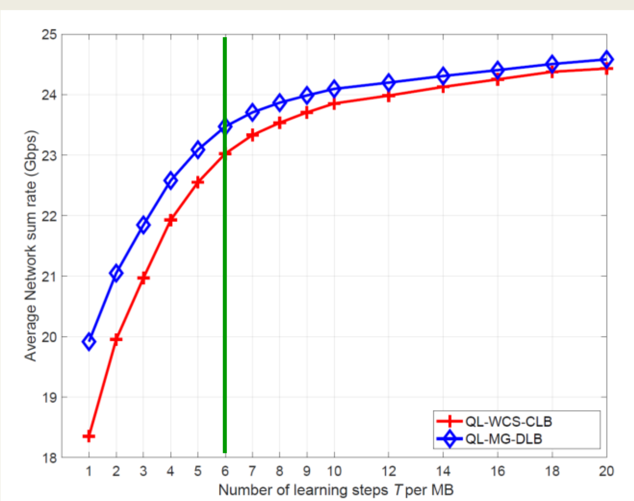


Performance of DRL for Handover in Wireless

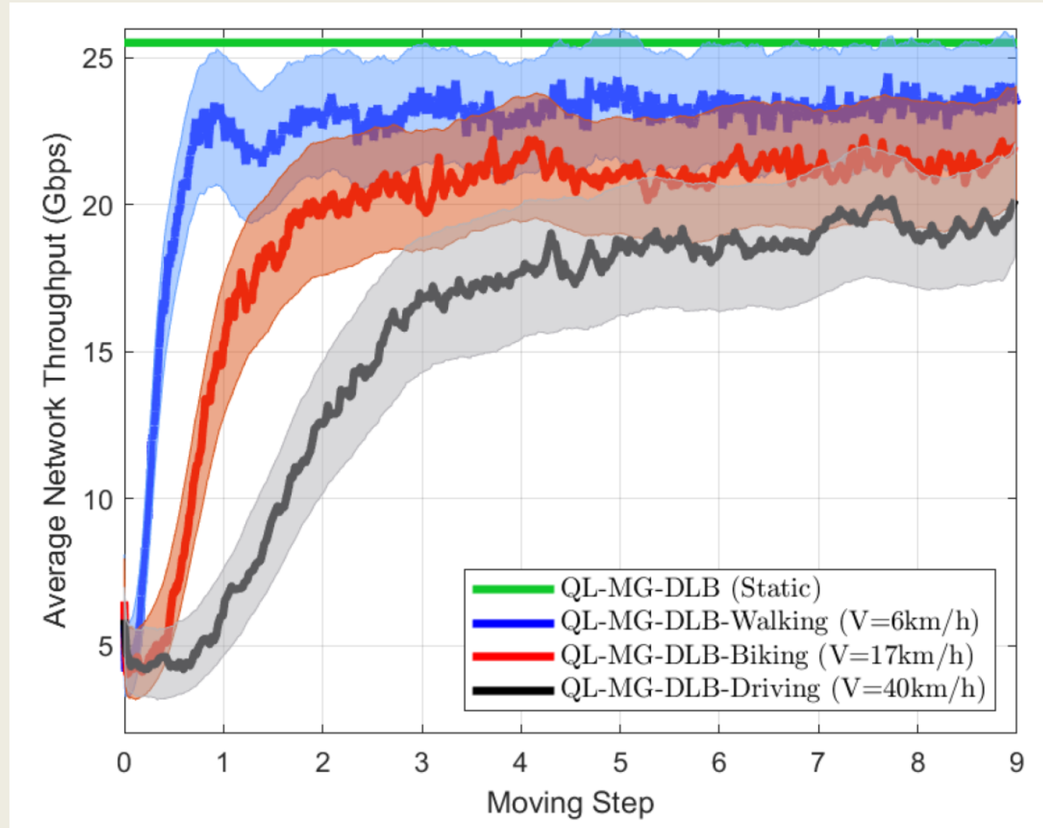
Reward vs. Learning Step at the beginning



Average reward vs. LS after 5 Moving Steps



Learning Progress vs User Mobility



- Has a "ramp-up" time but becomes stabilized after 2-3 moving steps
- Can use pre-training or meta-learning

Challenges in Applying RL in Wireless

- Testing in a real environment is costly
  - Need to have mechanism/algorithms in place and drive around to test
  - Or an extensive database of wireless measurements
  - Current results are evaluated using simulation

- Measurement frequency vs. RL update runtime (overheads)
  - Signaling and computation time overheads vs data transmission time
  - Need to be evaluated in practical system contexts

- New mobility patterns or unseen dynamics
  - Need to understand impacts on system performance
  - Effect on ramp-up time

- System constraints
  - Load balancing: centralized vs distributed
  - Any communication between agents?

RL in Wireless Systems is gaining popularity

Deep **reinforcement learning** for dynamic multichannel access in **wireless** networks
S Wang, H Liu, PH Gomes… - IEEE Transactions on …, 2018 - ieeexplore.ieee.org
… computation, we apply the concept
Finally, we propose an adaptive DQN a
☆ Save 99 Cite Cited by 462 Rel

Deep **reinforcement learning**-based edge caching in **wireless** networks
C Zhong, MC Gursoy… - IEEE Transactions on …, 2020 - ieeexplore.ieee.org
… at the **wireless** network edge using a deep **reinforcement learning** framework with Wolpertinger architecture. In particular, we propose deep actorcritic **reinforcement learning** based …
☆ Save 99 Cite Cited by 140 Related articles All 4 versions Import into BibTeX

Deep **reinforcement learning** for 5G networks: Joint beamforming, power control, and interference coordination
…, BL Evans, A **Alkhateeb** - IEEE Transactions on …, 2019 - ieeexplore.ieee.org
… online **learning** based algorithm b
communications was studied in [6], [
☆ Save 99 Cite Cited by 223 R

Multi-agent deep **reinforcement learning** for dynamic power allocation in **wireless** networks
YS Nasir, D Guo - IEEE Journal on Selected Areas in …, 2019 - ieeexplore.ieee.org
… **reinforcement learning** to power control [8]. Sun et al. [9] proposed a centralized supervised **learning** … 5) Using simulations, we compare the **reinforcement learn**ing outcomes with state-of…
☆ Save 99 Cite Cited by 534 Related articles All 7 versions Import into BibTeX

Deep-**reinforcement learning** multiple access for heterogeneous **wireless** networks
Y Yu, T Wang, SC Liew - IEEE journal on selected areas in …, 2019 - ieeexplore.ieee.org
… to the traditional **reinforcement learning** (RL) [5] for **wireless** … affords us with two essential properties to **wireless** MAC: (i) fast … to **wireless** networks because the **wireless** environment may …
☆ Save 99 Cite Cited by 349 Related articles All 6 versions Import into BibTeX

and many more …