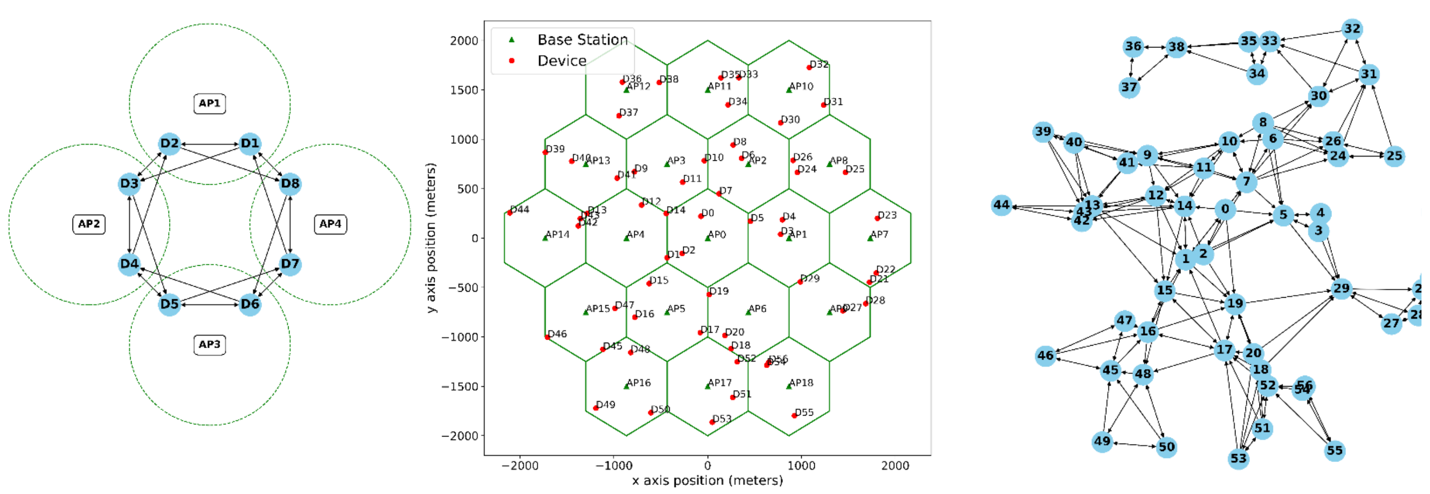
Spectrum Subband Allocation and Packet Scheduling in Conflict Graphs via Multi-agent Reimforcement Learning

Yiming Zhang, ECE Department, yimingzhang2026@u.northwestern.edu
Dongning Guo, ECE Department

Northwestern University

BACKGROUND

Conflict graphs are simple abstractions of wireless networks.



RESEARCH QUESTIONS

Our goal is to design efficient scheduling policies for each age in conflict graphs:

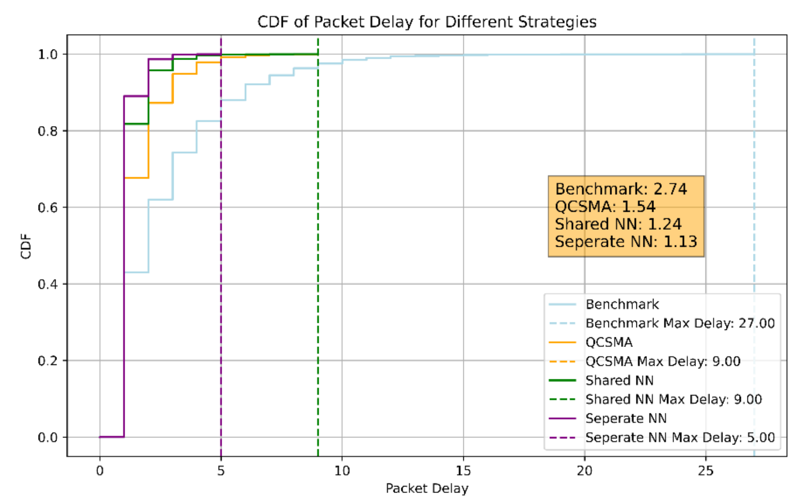
- What performance metric?
- How to handle multiple frequency sub-bands?
- How to scale to large networks?
- How to use dynamic traffic and information sharing between agents to resolve conflicts?

METHODS AND MATERIALS

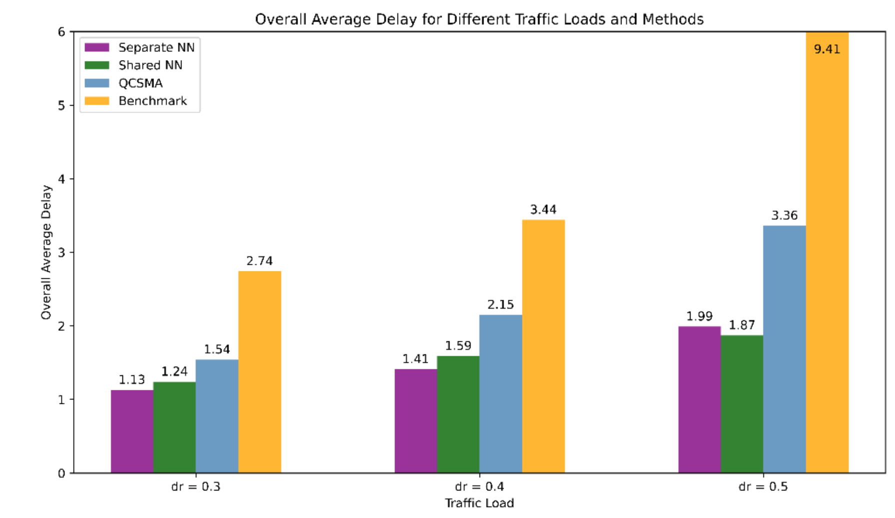
1.Problem Formulation: Cast our challenge as a decentralized partially observable Markov decision process (Dec-POMDP).
2.Algorithm Selection: Utilized the on-policy algorithm, Multi-agent Proximal Policy Optimization (MAPPO).
3.Network Training: Developed distinct neural networks tailored for individual agents.
4.Recurrent Units: Integrated long short-term memory (LSTM) to encode the dynamic of environments and actions.
5.Training & Execution Strategy: Modified the centralized training distributed execution approach to enable learning tailored to our specific scenario.

RESULTS

4-agent 8-link conflict graph with 3 channels

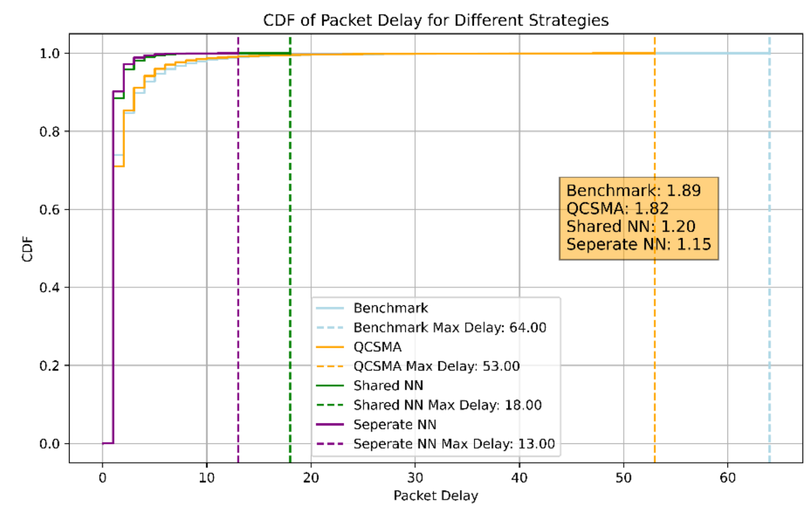


(a) CDF under light traffic load

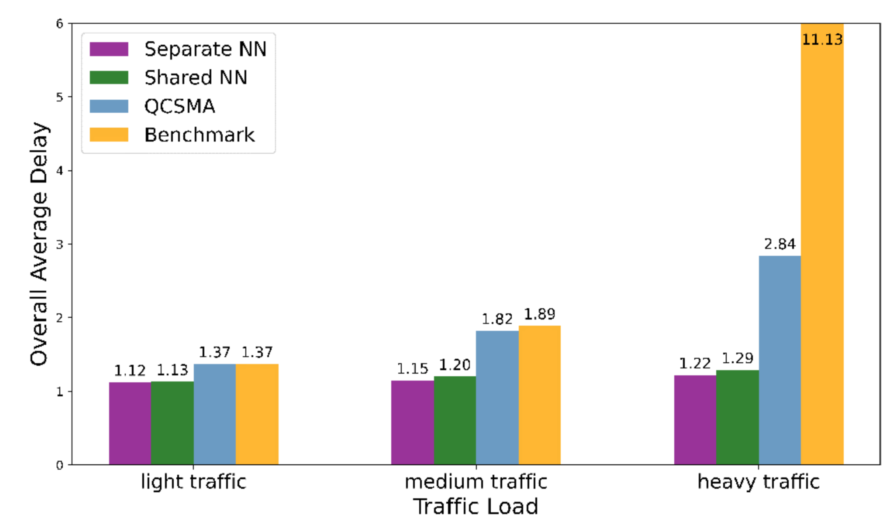


(b) Packet delay under varying traffic load

19-agent 57-link conflict graph with 3 channels



(a) CDF under medium traffic load

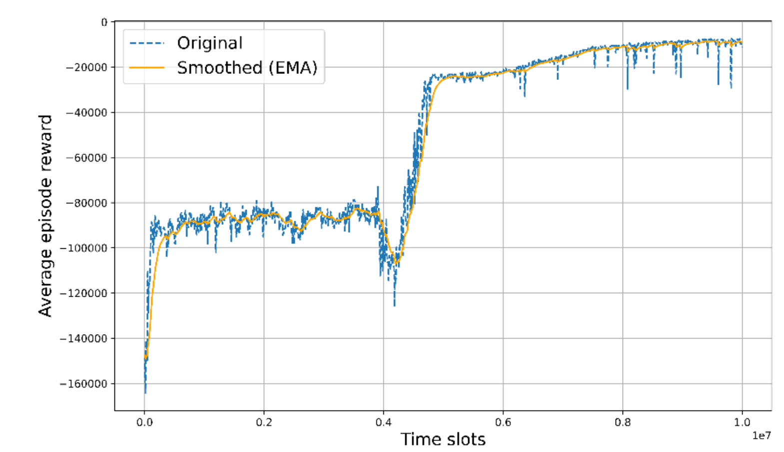


(b) Packet delay under varying traffic load

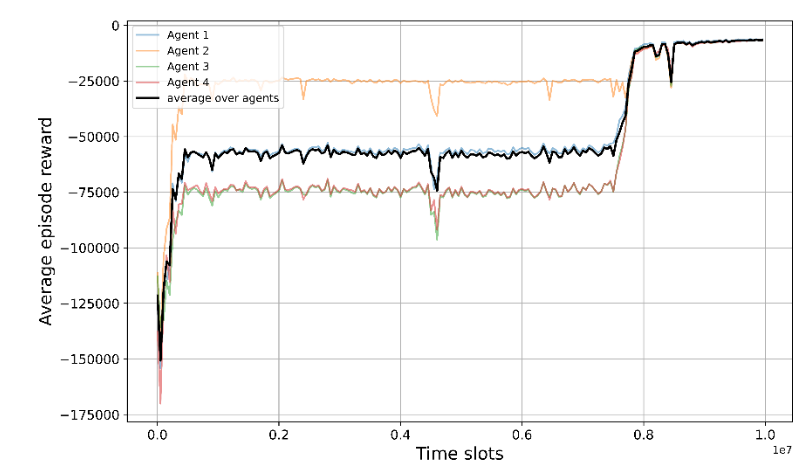
Model mismatch

Testing Traffic Load	Light	Medium	Heavy
trained under light traffic	good	mixed	unstable
trained under medium traffic	good	good	unstable
trained under heavy traffic	good	good	good

Convergence



(a) Shared policy



(b) Separate policies

CONCLUSION

The simulation results showed the effectiveness of the trained policy in achieving significantly improved QoS in capacity region compared to other widely used schedulers.

The proposed method exhibits remarkable robustness in adapting to variations in conflict graph sizes and traffic densities, further validating its efficacy in practical deployment scenarios.

ACKNOWLEDGEMENTS

This work was supported in part by NSF grant No. 2003098, No. 2216970 and Spectrum X.

REFERENCES

[1] Y. S. Nasir and D. Guo. "Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks." *IEEE Journal on Selected Areas in Communications* 37.10 (2019): 2239-2250.
[2] Y. Zhang and D. Guo. "Distributed MARL for scheduling in Conflict Graphs", Allerton Conf. Commun. Control Computing, 2023