Multi-Agent Reinforcement Learning in Coordinated Multi-Point (CoMP) Technologies within Mobile Networks

Alberto Castro[1]*, Shaharyar Kamal* [2]*, César Azurdia M.*[3]
Affiliation 1, 2, 3 - Department of Electrical Engineering, University of Chile, Chile

open6GRIT
OpenRIT 6G Workshop Chile
April 1-3, 2025

1. Introduction

- Use of Deep Reinforcement Learning (DRL) algorithms in Coordinated Multi-Point (CoMP) technologies to improve user experience in mobile networks, particularly in urban areas.
- Improve resource management by dynamically selecting cells based on user behavior and network conditions. The research highlights the importance of self-learning algorithms and their role in real-time decision making, addressing the limitations of traditional techniques.

2. Methodology, simulation tests

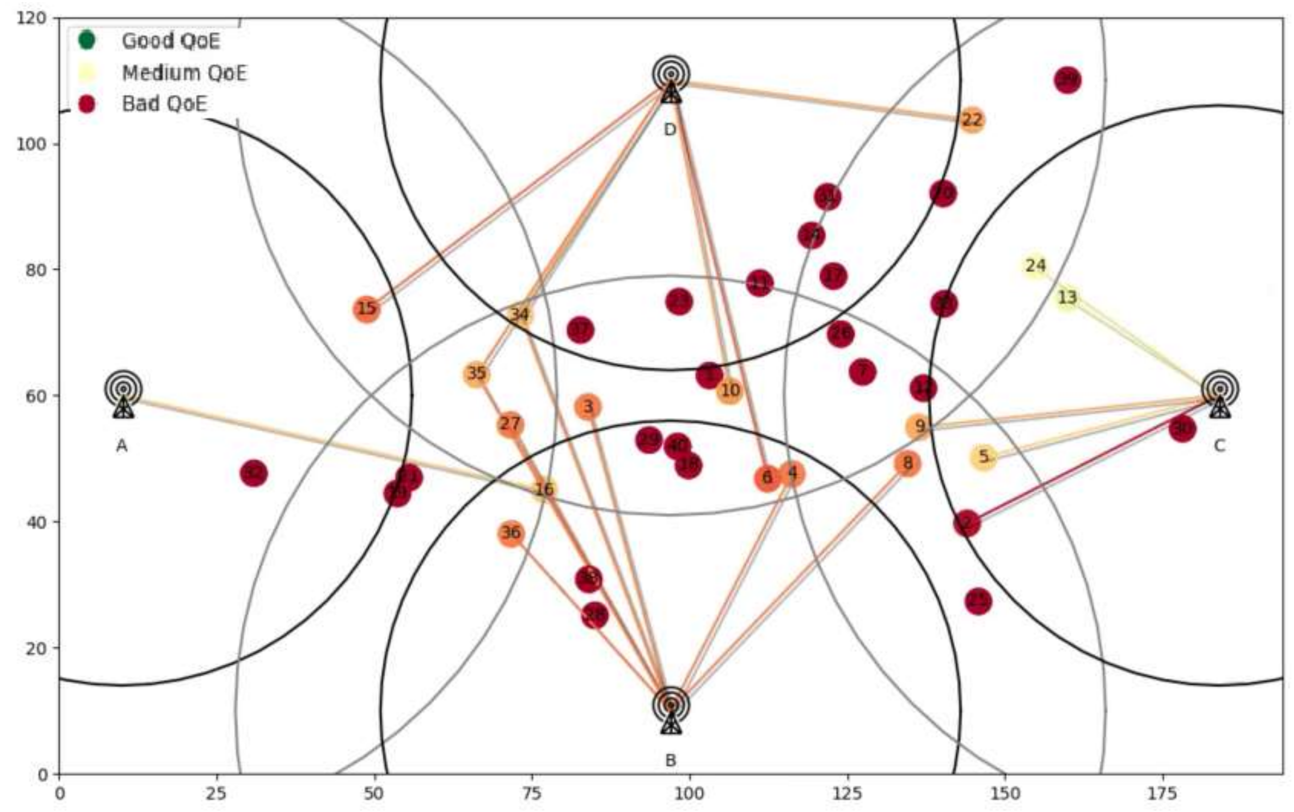

Fig. 1 - Example of 40 UEs simulation scenario.

2.1 Quality of Experience (QoE) metric

- As mobile networks expand with new technologies like Coordinated Multi-Point (CoMP) and machine learning, it is important to know how different things affect Quality of Experience (QoE) metric.

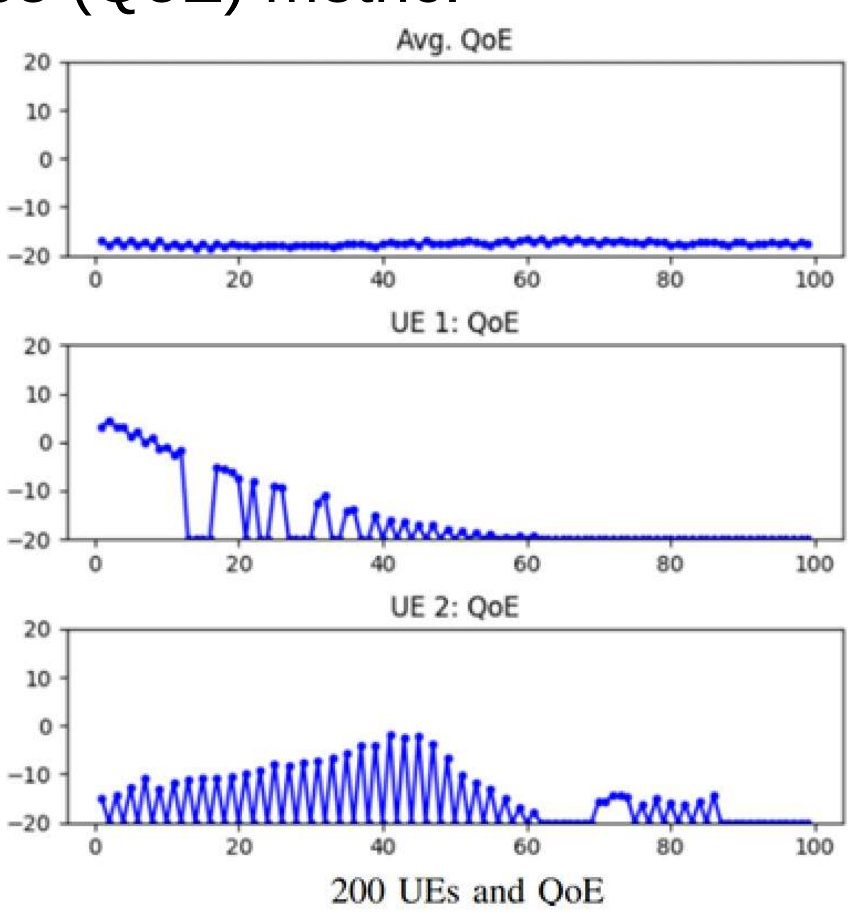

Fig. 2 – Running 200 UEs simulation scenario with Avg. QoE metric

2.2 Least Squares Channel Estimation and QoE optimization

- QoE of UE u_j by its utility $U_j(t)$.

$$U_j(t) = \min\left\{U_j^{\max}, \max\left\{U_j^{\min}, w_1 \log_{w_2}(w_3 + r_j(t))\right\}\right\} \tag{1}$$

- Maximize long-term QoE, averaged over all UEs (Avg. QoE) and time steps.

$$\text{Avg. QoE} = \max \lim_{T\to\infty} \frac{1}{T}\frac{1}{M} \sum_{t\in T, j\in 1,\ldots,M} U_j(t) \tag{2}$$

3. Results

- Using Reinforcement Learning methods, especially for cell selection, has been foundto improve QoE by better allocating resources according to the demands of the live user.
- Table I shows a relevant result that throughput worsens when more UEs are added, but does not continue to worsen after 200 UEs. In the range of 300 to 400 UEs, the total throughput remains the same, compared to not using reinforcement learning optimization.

# UEs	Reward	Step reward mean	Sample throughput [kbit]	Load throughput [kbit]
2	-4.234	-0.334	4.182.534	179.048.644
20	-1068.59	-17.71	729.333	119.149.594
40	-2624.23	-30.238	408.636	145.408.355
80	-5868.49	-61.394	153.29	125.056.894
100	-7724.09	-79.312	110.483	116.915.302
140	-11413.1	-116.059	67.612	77.663.051
180	-15260.7	-159.367	43.78	78.114.532
200	-17307.8	-177.484	36.596	72.601.252
300	-27124.2	-272.539	17.321	46.746.027
350	-32176.1	-323.691	12.614	55.113.621
400	-36941.5	-373.632	10.645	46.414.588

Table. I - Throughput and rewards by utility

- Deciding which cells serve which UEs may be completely driven by the network, or UEs may trigger or assist cell connections themselves.

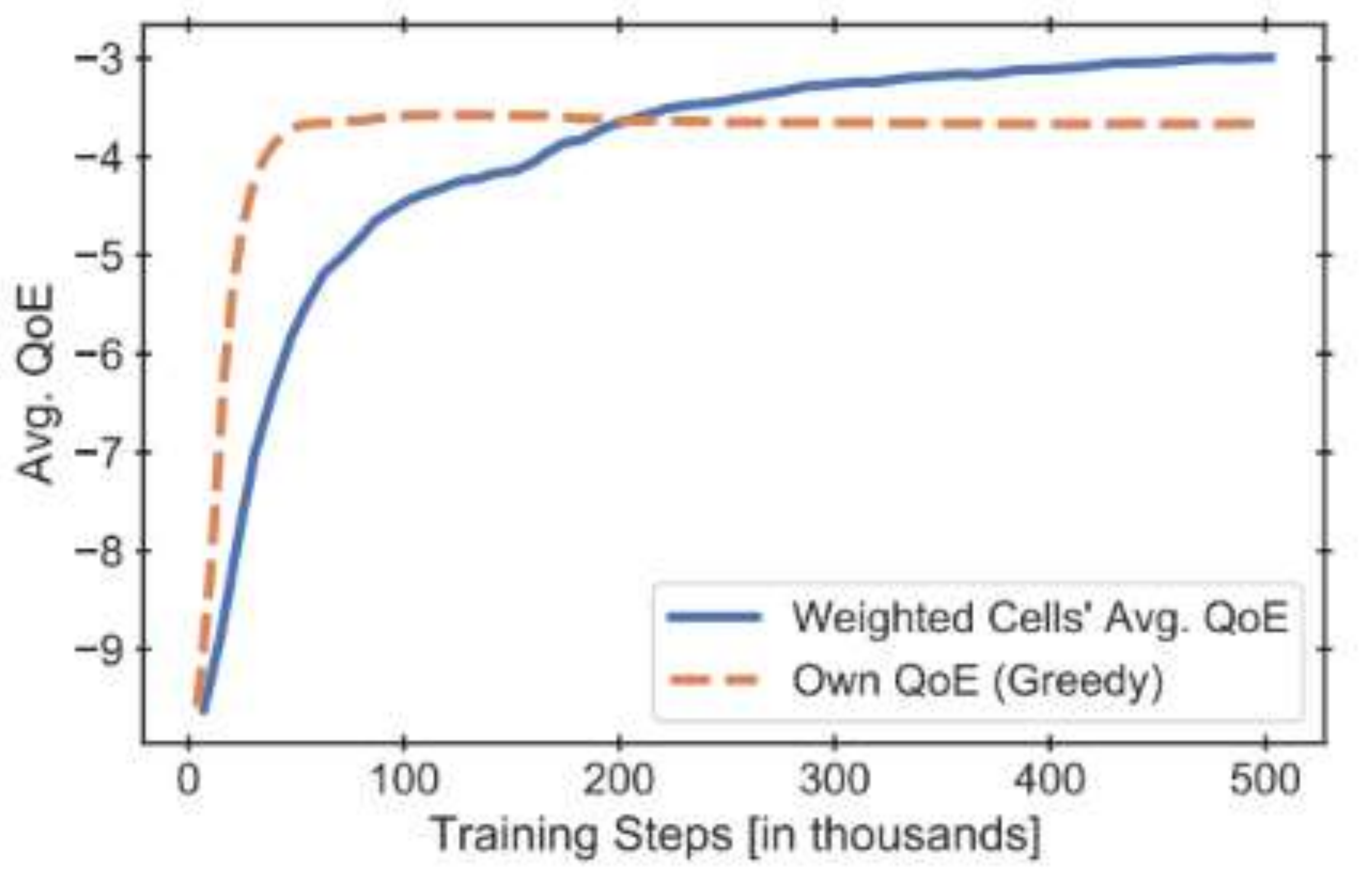

Fig. 3 - Comparison without using Reinforcement Learning optimization

Summary / Conclusions

- The QoE metric based on throughput and channel throughput value (CQI) shows that QoE, divided into quality levels, total quality, and total throughput decrease, but can be stabilized thanks to the use of reinforcement learning optimization, in our case, in a centralized manner.

Bibliography or Website Project

- *https://deepcompresults.ai-signals.tech/ (Results, via docker off/on)*
- *https://www.ai-signals.tech/*
- *Cell Selection with Deep Reinforcement https://bit.ly/3DZLvct*