DeepCoMP: Self Learning Dynamic Multi Cell Selection for CoMP

EC354 Cellular Mobile Communication Project

Under the guidance of
Dr. Mukesh Kumar Mishra.
Assistant Professor at the Department of
ECE, IIIT Dharwad

Submitted by:
Abhishek Singh Kushwaha (19BEC001)
Sriram Shivganesh (19BEC043)

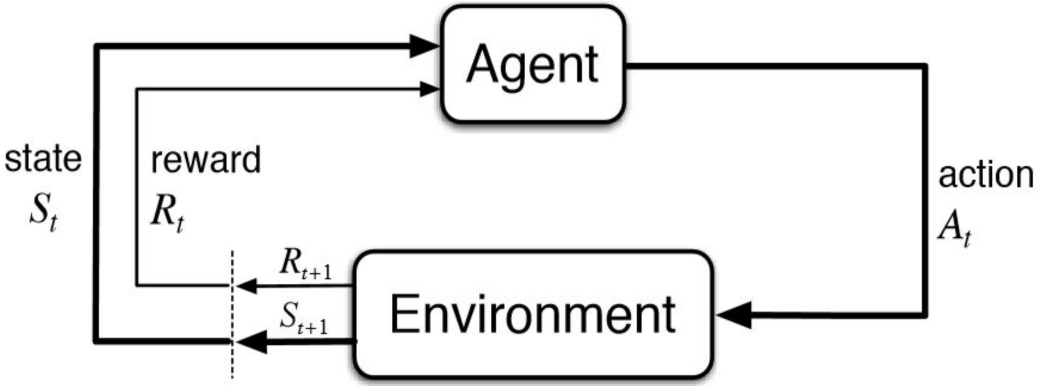
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY
Dharwad
ज्ञानेन विकासः

What is Reinforcement Learning?

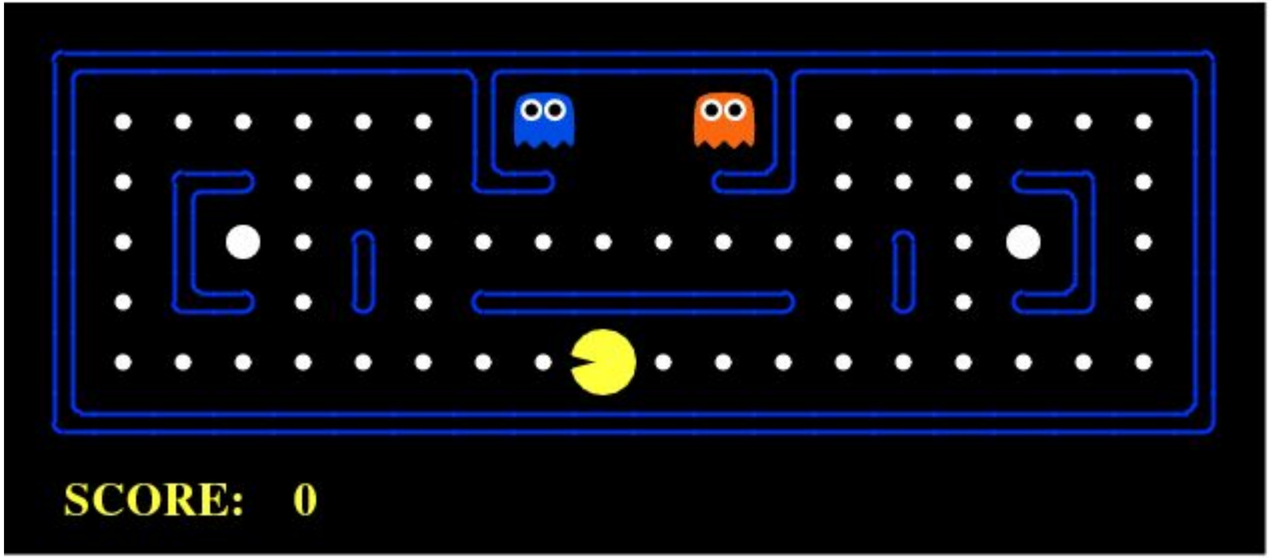
- Reinforcement Learning is a feedback-based Machine learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions.
- For each good action, the agent gets positive feedback or reward, and for each bad action, the agent gets negative feedback or penalty.
- The agent interacts with the environment and explores it by itself. The primary goal of an agent in reinforcement learning is to improve the performance by getting the maximum positive rewards.

What is Reinforcement Learning?

- We do not need to pre-program the agent, as it learns from its own experience without any human intervention.
- RL solves a specific type of problem where decision making is sequential, and the goal is long-term solution.
- Current use cases include the following:
  - **resource management**
  - gaming
  - robotics



Example of Reinforcement Learning



Reinforcement Learning = Machine Learning + Control Theory

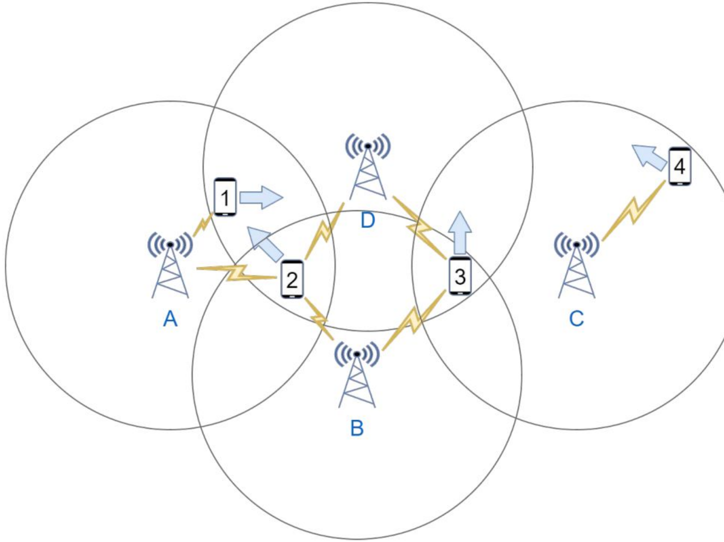
CoMP - Coordinated Multipoint

- Coordinated MultiPoint (CoMP) is based on transmission and reception at multiple separated sites with dynamic coordination among them.
- CoMP started to be used more aggressively in LTE Advanced, as a way of improving service at the cell edge.
- Makes better utilization of network by providing connections from several base stations at once.

Problem statement

Wireless mobile scenario

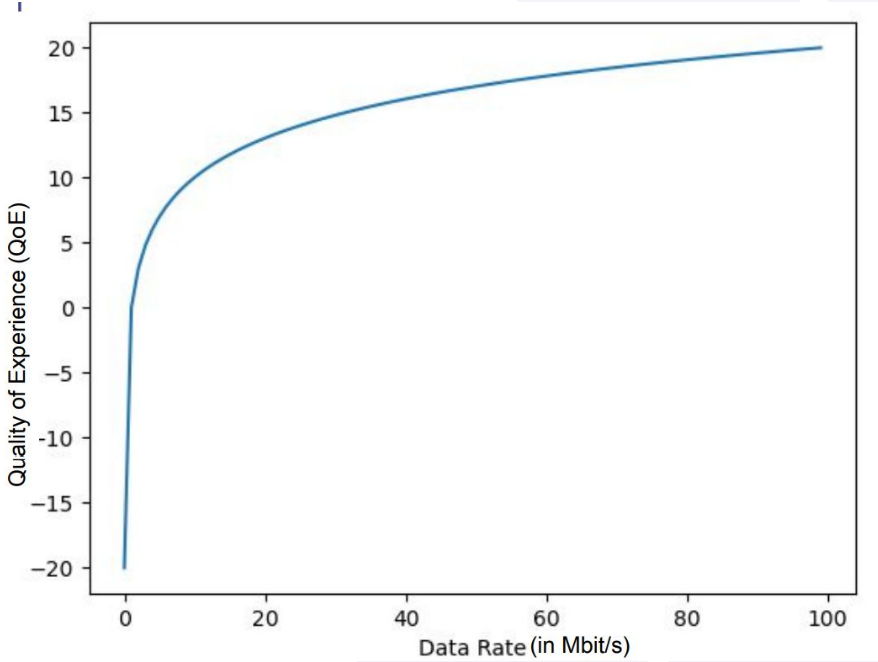
- dense cells, moving users
- requirement of high data rate
- Users compete for resources
- Heterogeneous resource allocation



Motivation and Approach

Parameter :
Quality of Experience (QoE) = log(data rate)

Goal is to **Maximize QoE of all users**



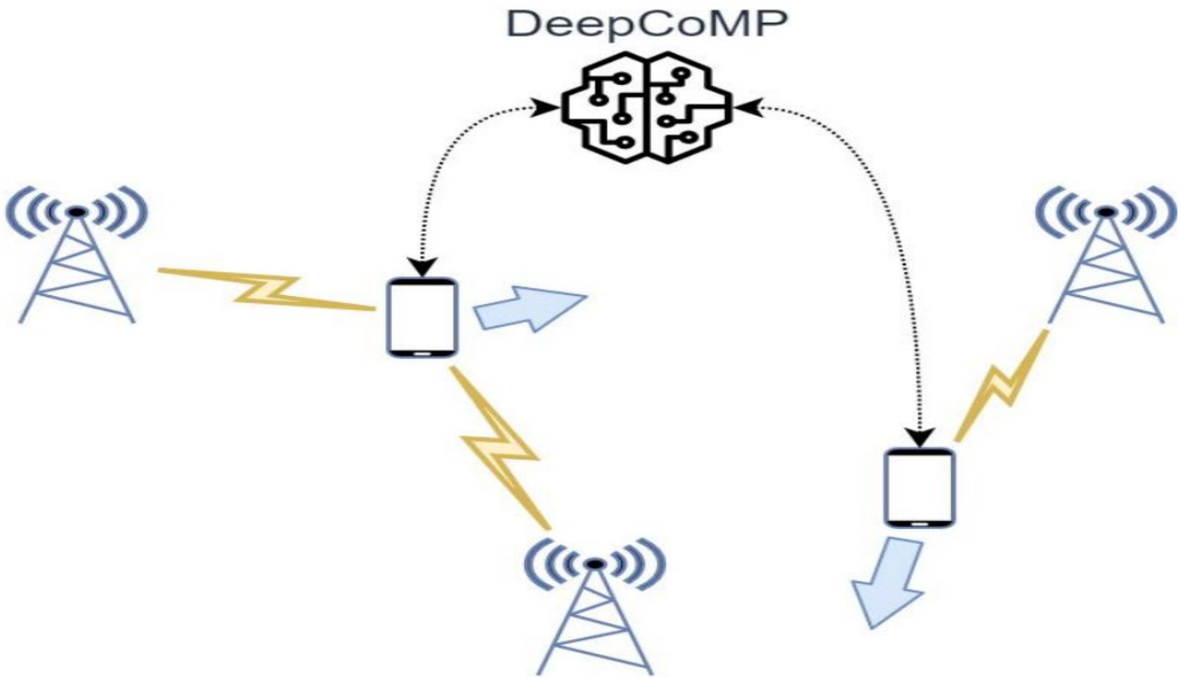
Types of self learning DRL Approaches:

Training	**Inference**	**Name**
Centralized	Centralized	DeepCoMP
Distributed	Distributed	D3-CoMP
Centralized	Distributed	DD-CoMP

DeepCoMP : Central observation and control of all users

- Requires global view and control of all users
- Large action space
- Complex
- But allows fine grain cooperation between users

DeepCoMP



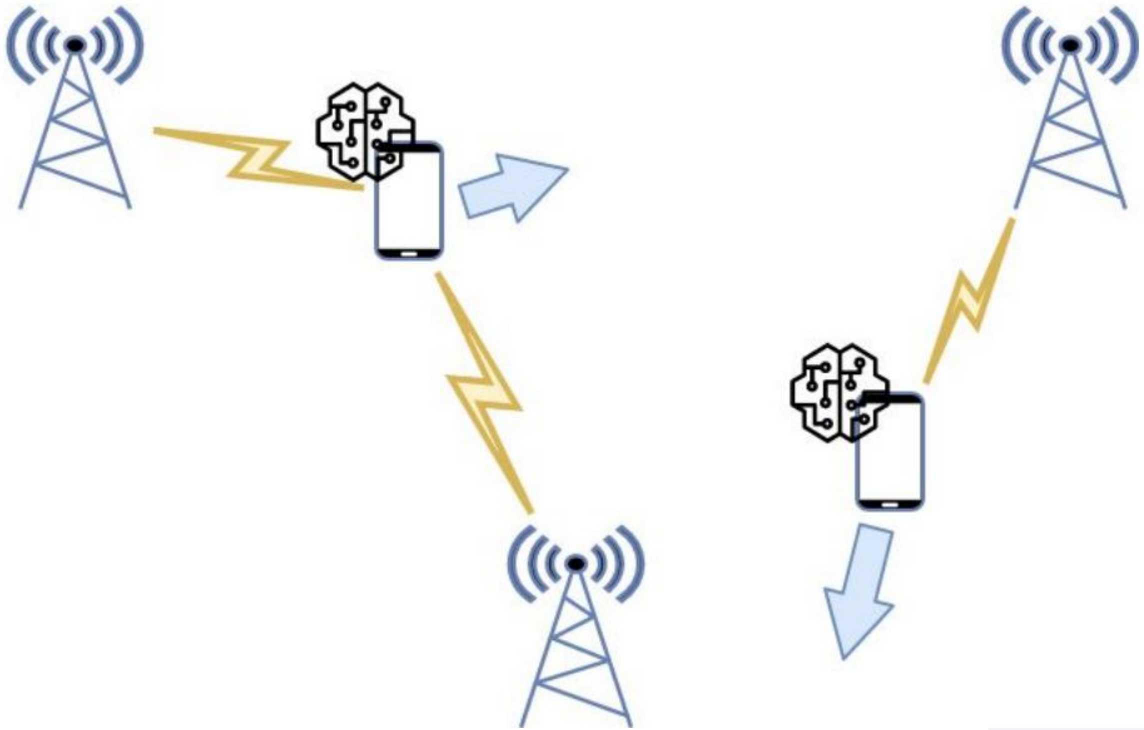
DeepCoMP

- Observations:
  - Current connections
  - Signal strength between all cells and users
  - Users' QoE
- Actions:
  - Either keep all current connections/disconnect
  - Or connect/disconnect a certain cell
- Reward:
  - Sum of users' QoE

Distributed DRL:

- Separate DRL agents for each user
- Local observations and control
- It is Simpler and Faster
- Prone to greedy behaviour

Distributed DRL:



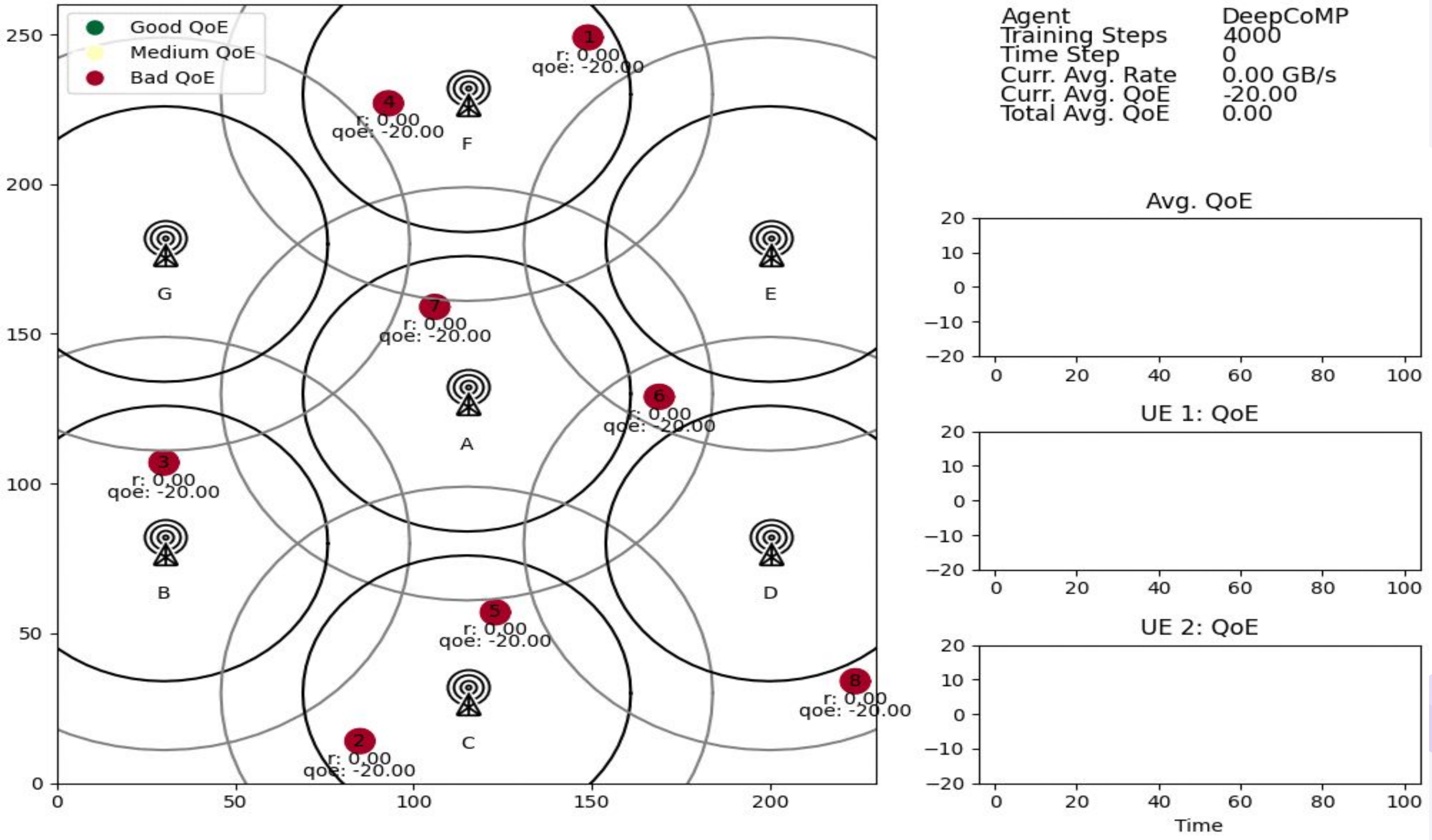
D3-CoMP

- Fully distributed
- independent DRL agents
- No communication between DRL agents for training
- Can learn heterogeneous cell selection policies per user

DD-CoMP

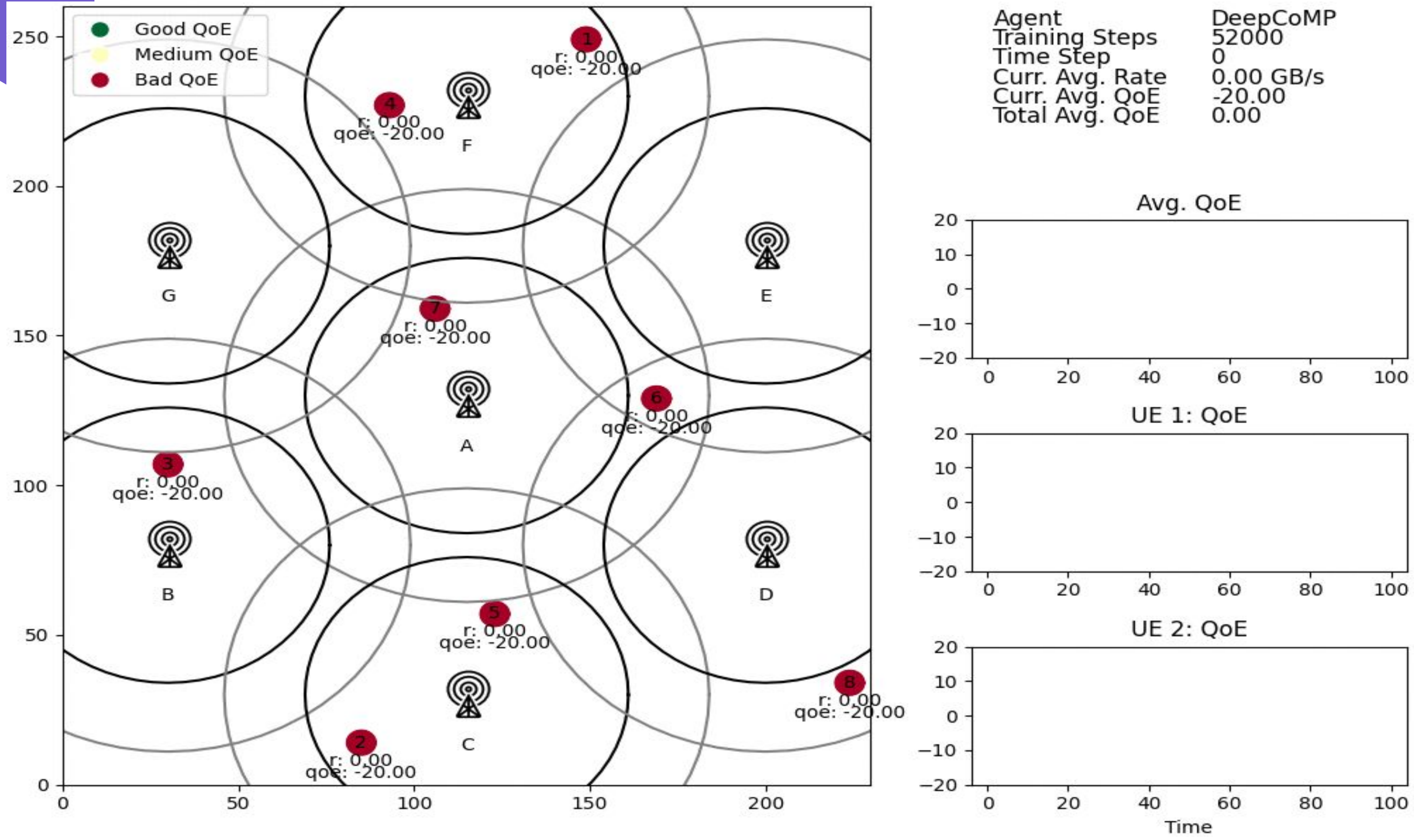
- Central policy and training, but distributed inference
- Distributed inference with local observations and actions
- But DRL agents share their experience
- Leverage data from other users
- Slightly better than D3-CoMP (more robust)

Outputs



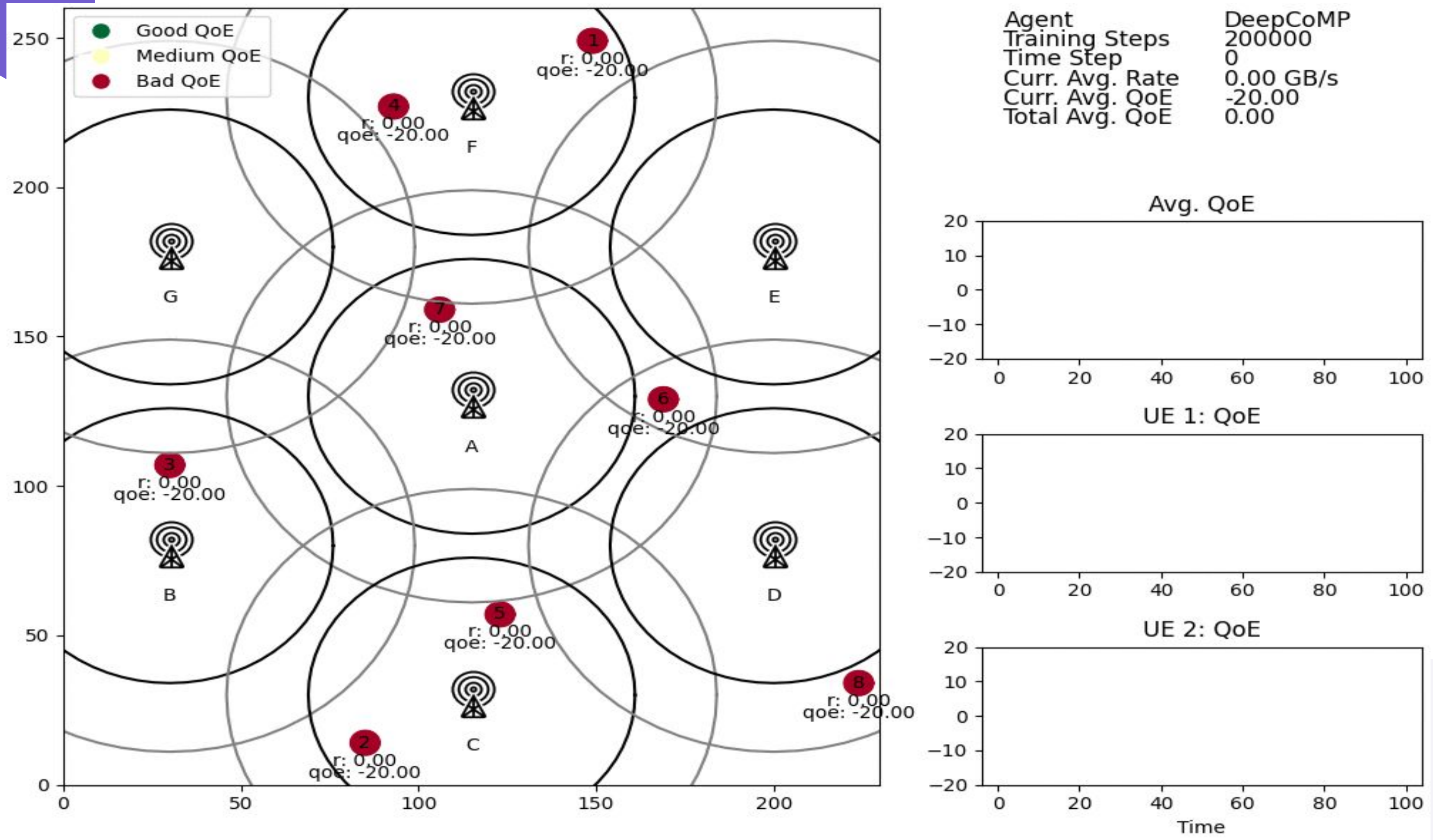
Name DeepCoMP:
Steps: 500
QoE: bad

Outputs



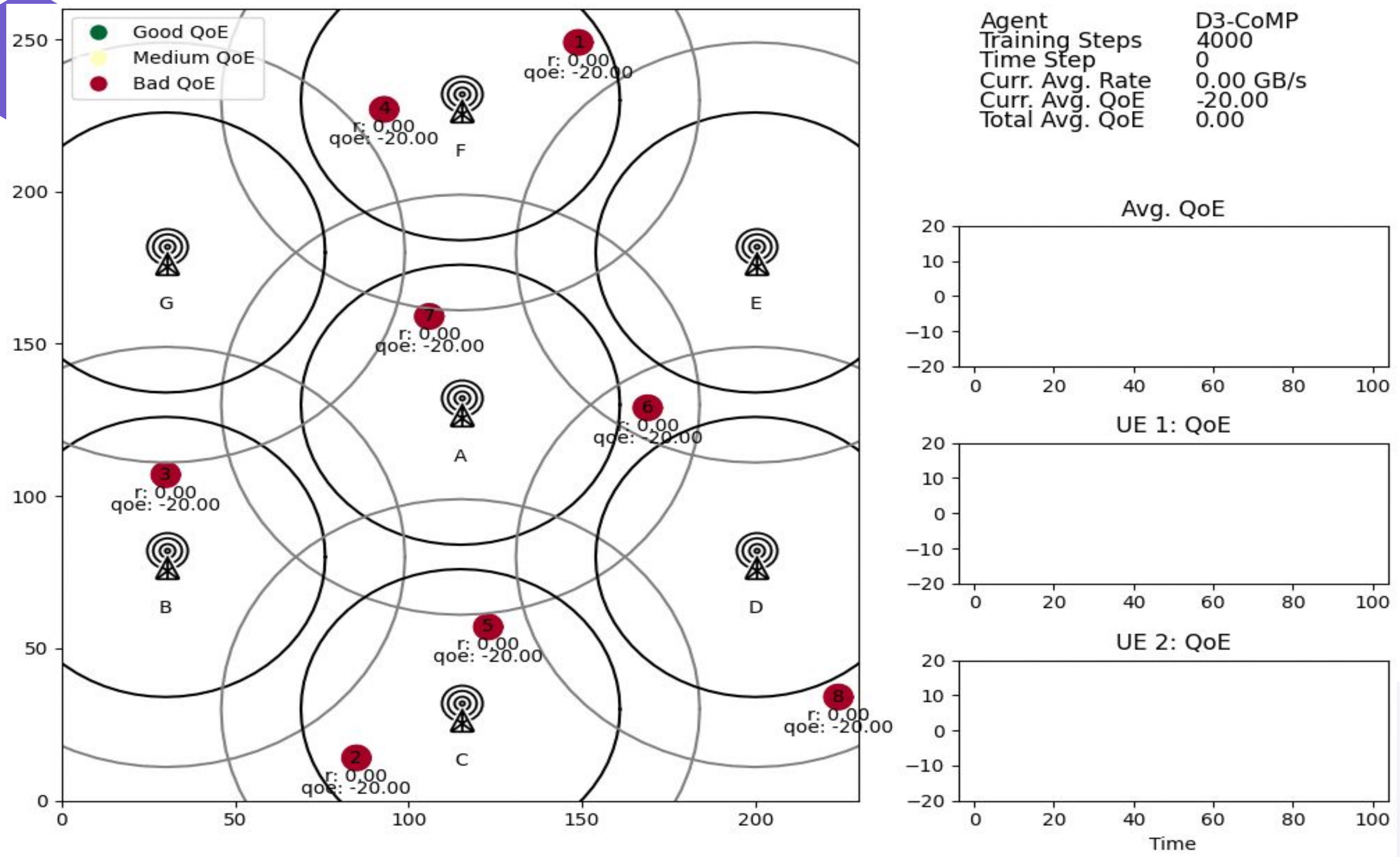
Name: DeepCoMP
Steps: 50,000
QoE: decent

Outputs



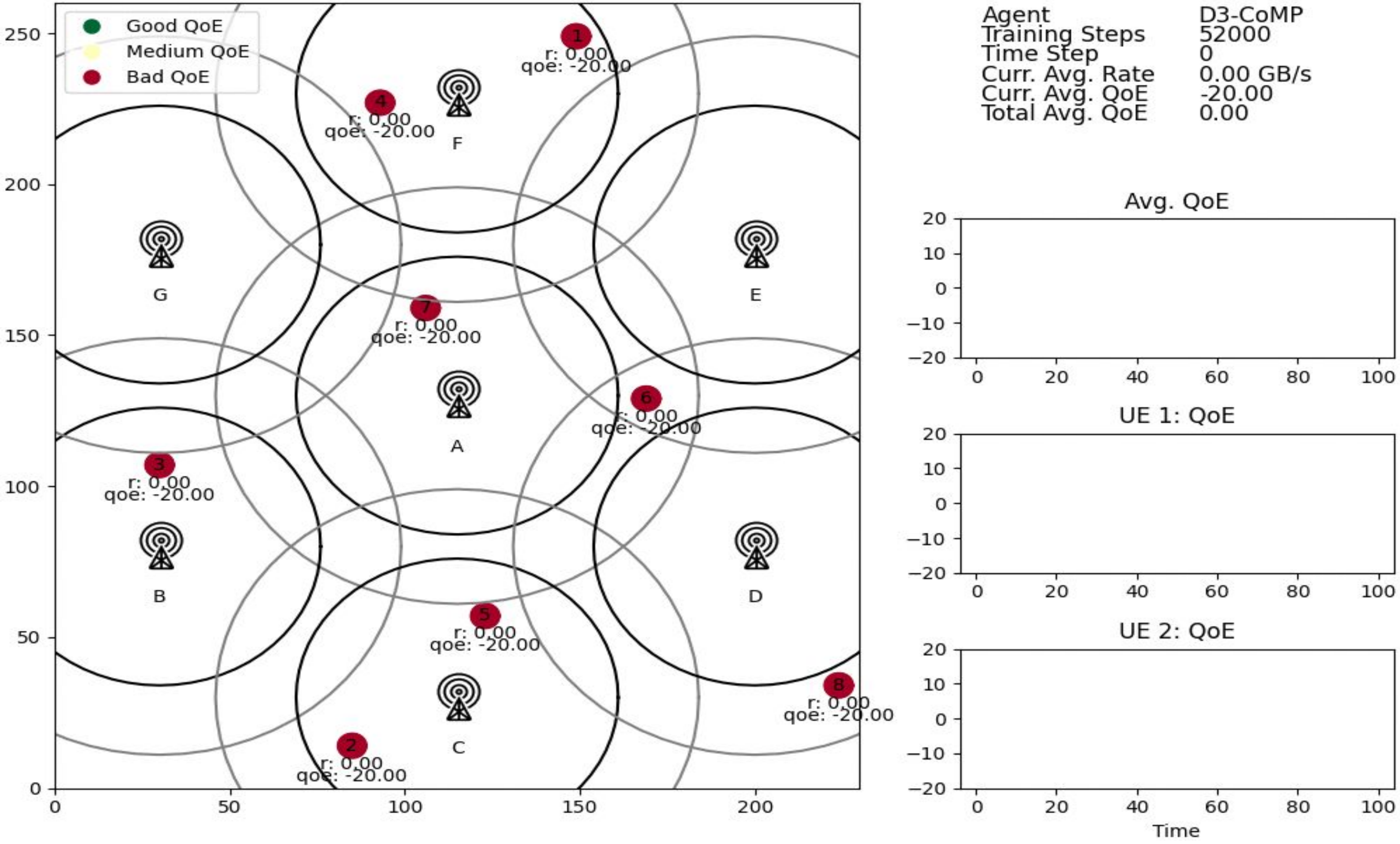
Name: DeepCoMP
Steps: 2,00,000
QoE: Good

Outputs



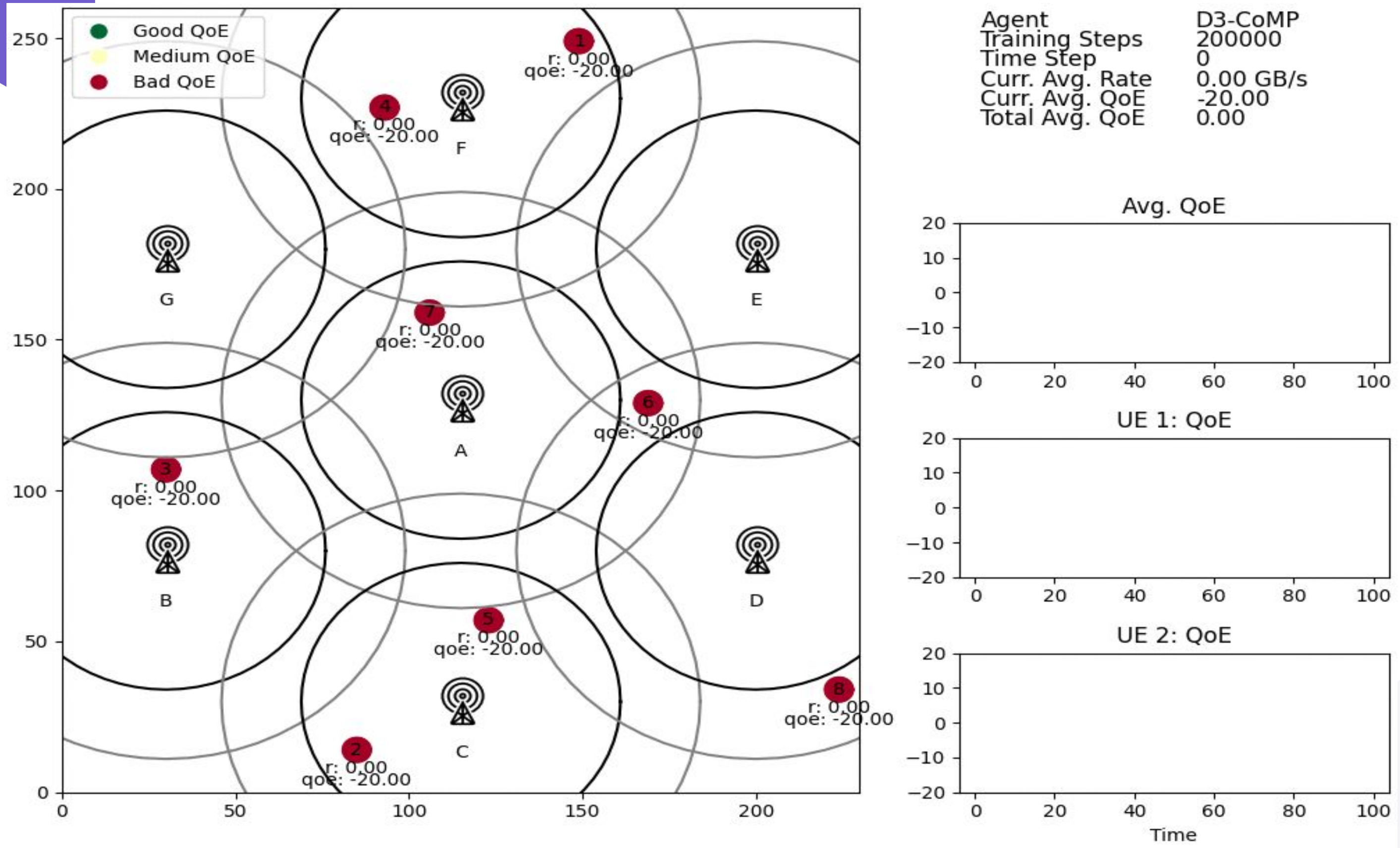
Name D3-CoMP:
Steps: 500
QoE: bad

Outputs



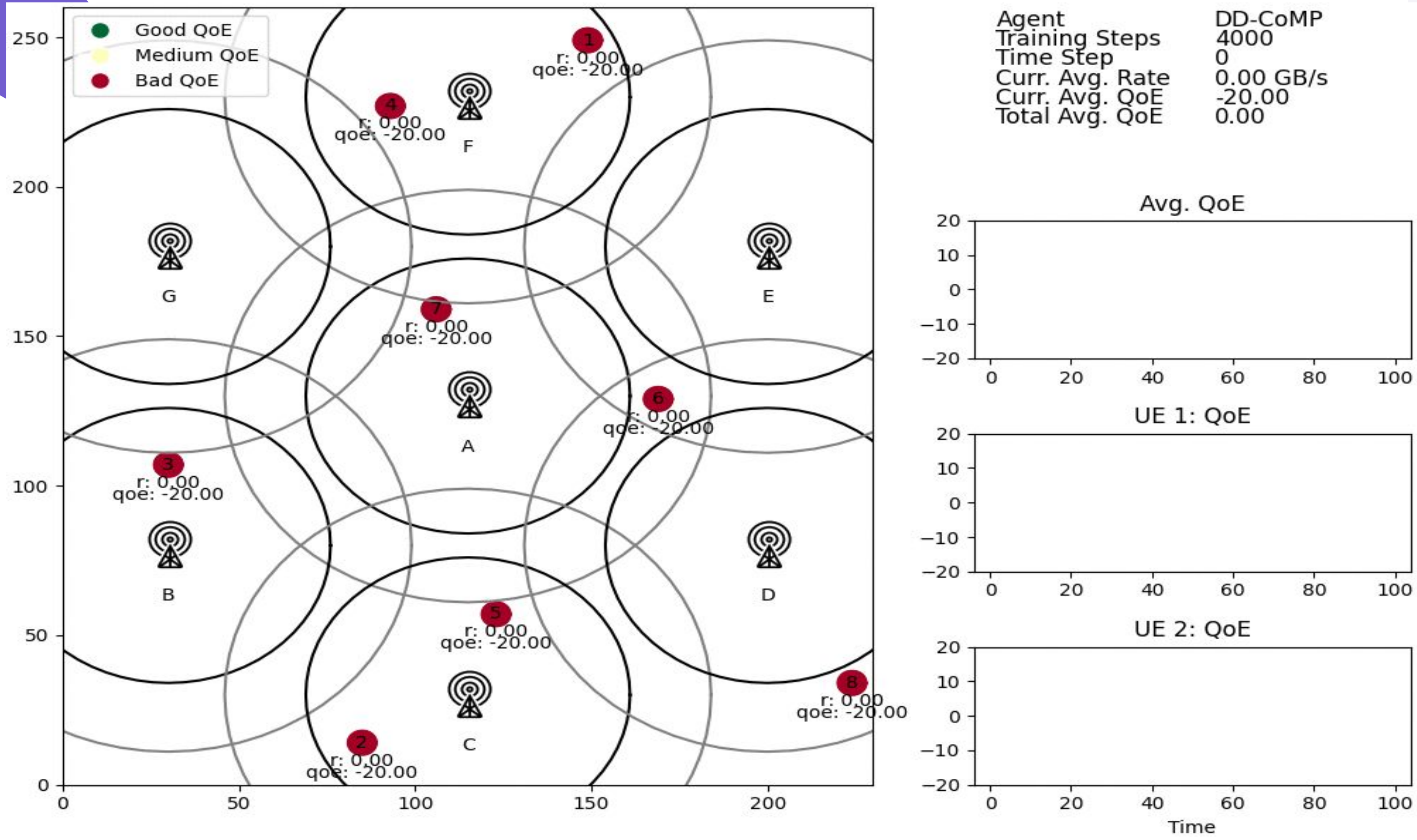
Name: D3-CoMP
Steps: 50,000
QoE: decent

Outputs



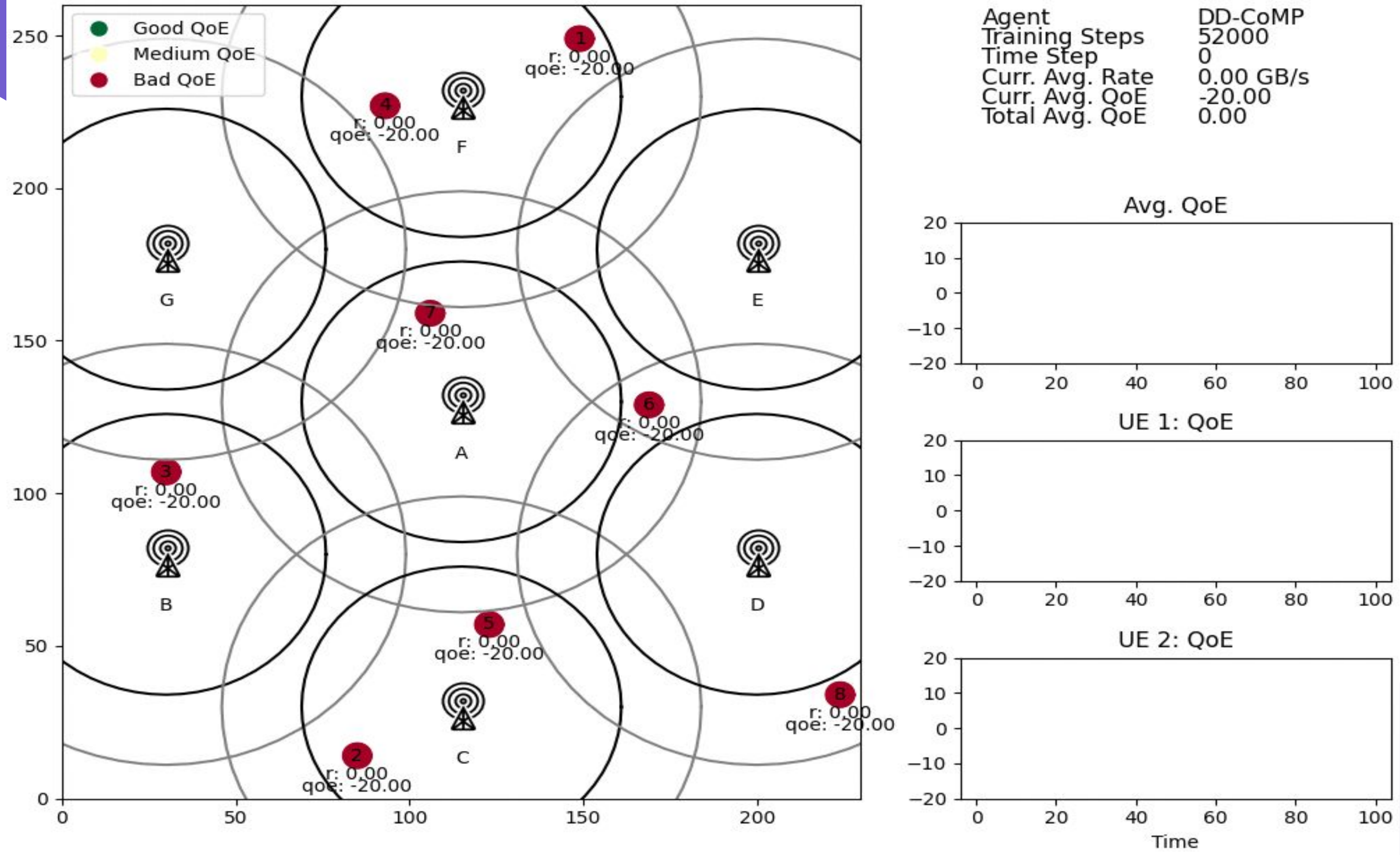
Name: D3-CoMP
Steps: 2,00,000
QoE: Good

Outputs



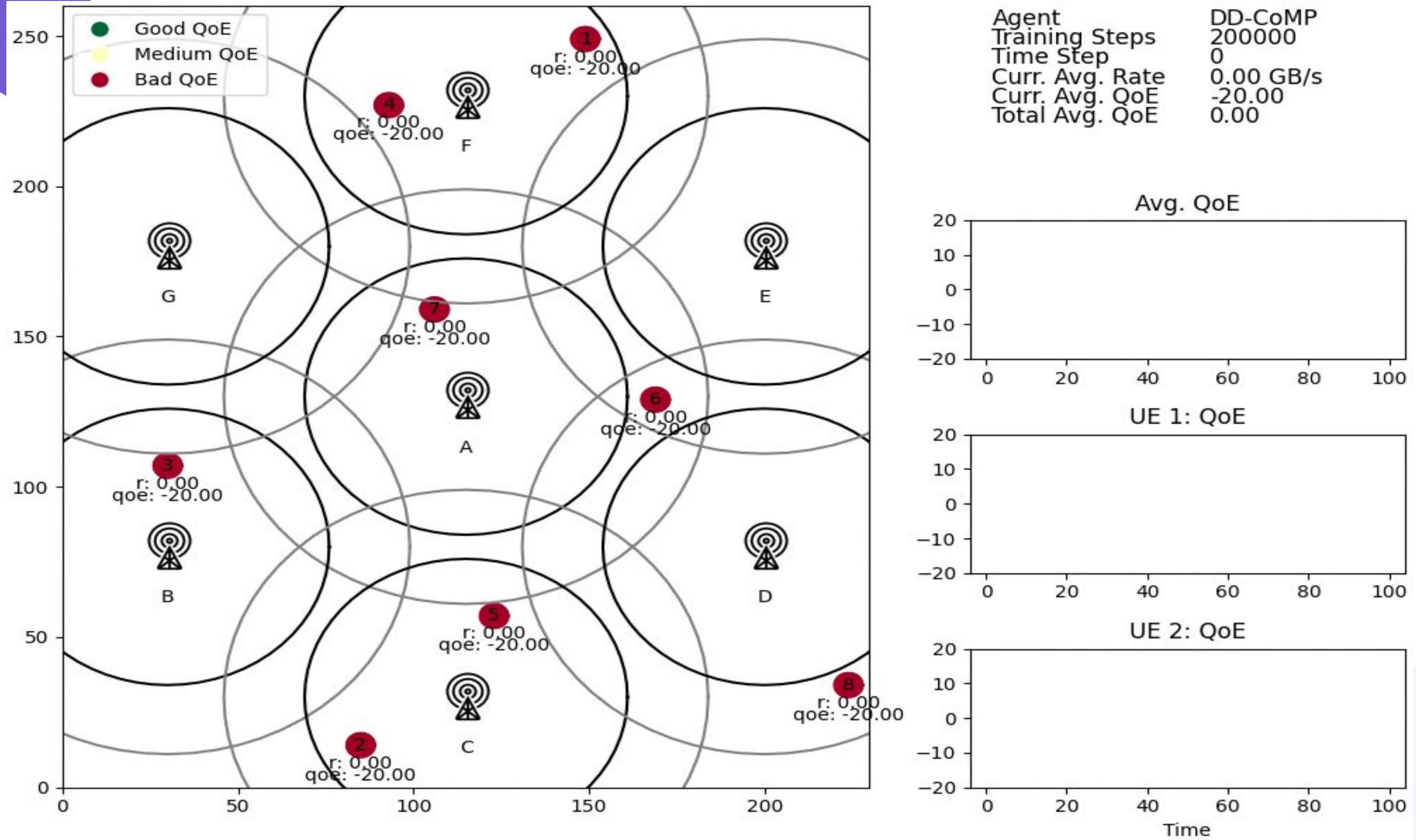
Name DD-CoMP:
Steps: 500
QoE: bad

Outputs



Name: DD-CoMP
Steps: 50,000
QoE: decent

Outputs



Name: DD-CoMP
Steps: 2,00,000
QoE: Good

Observation

- No need for human intervention or instructions
- DRL agents learn multi-cell selection effectively
- DRL agents self-adapt to each scenario
- DRL agents outperform existing approaches
- Distributed DRL learns good policy faster
- Central DRL ultimately learns better policy

Conclusion

- Three self-learning DRL approaches :
  - Central DeepCoMP: Slow but highly optimized multi-cell selection
  - Distributed DD-CoMP & D3-CoMP: Fast, local multi-cell selection
- Outperform existing approaches :
  - Work with minimal, realistically available information
  - Self-adapt to varying scenarios
  - Robust to sudden changes
  - Scale to large networks

THANK YOU!