

Multi-Agent Reinforcement Learning for Resource Allocation in Wireless Network Communication

Tuesday, 31 March 2026 16:10 (20 minutes)

Multi-Agent Reinforcement Learning (MARL) is an important subfield of Reinforcement Learning, in which multiple agents learn in a shared environment. The simultaneous learning of several players naturally arises in domains like robotics, network communication and traffic control, where agents affect and influence one another. Thus, MARL can simulate real-world problems in a reliable way, and consequently, interest in MARL continues to grow.

In this work, we consider the real-world problem of resource allocation in wireless network communication. Due to the fast development of wireless network communication, data traffic is rising, and more devices are communicating like mobile users and machines in factories.

These devices affect their communication by trying to gain the same resources for guaranteeing reliable communication. Thus, avoiding the overlap of used frequency bands and controlling the wireless network communication becomes more complicated. We use MARL to solve the problem of overlapping frequency bands so that the trained algorithm should distribute frequency bands properly. This ensures reliable network communication.

Accordingly, all communicating devices are agents in a specific area where they communicate. As actions, they choose communication channels. Across different scenarios, the set of possible channel selections is variable.

To enable a reliable solution of the problem, each agent receives the following information in its state: the communication channel used in the previous step, the own Quality of Service (QoS) achieved by the last action, a vector of all neighbouring devices and the communication channels the neighbouring devices used in their last action.

After their selection of a communication channel, all agents a reward, chosen to be the sum of the achieved QoS of all agents, since a shared reward avoids adversarial behaviour and leads to cooperation between the agents.

We start the training by using a single-agent Q-learning algorithm. This leads to optimal training results for a small amount of agents. However, there occurs additional MARL problems like the non-stationarity of the environment, as well as scalability problems and non-unique learning goals. To attack these problems, we use different MARL algorithms, like a NashQ-algorithm and an IQL-algorithm. These algorithms give optimal training results for a small amount of agents and outperform the regular Q-Learning algorithm in training time. However, the scalability problem still persists, so we want to address this in future work by using a VDN-QMIX algorithm, which uses the global state during training, but goes back to a decentralized setting in the final execution. Thus, we hope to ensure a good scaled training with many agents as well as a good vision of the real-world problem.

Student

Yes

Primary authors: DORFINGER, Peter; KWITT, Roland; POCHABA, Sabrina (Salzburg Research); HIRLAENDER, Simon

Presenter: POCHABA, Sabrina (Salzburg Research)

Session Classification: Contributed talks