Contribution ID: **53** Type: **not specified**

Learning LP-indices in Average-Reward Restless Multi-Armed Bandits

Wednesday, June 19, 2024 1:30 PM (30 minutes)

Restless Multi-Armed Bandits (RMABs) are extensively used in scheduling, resource allocation, marketing and clinical trials, just to name a few application areas. RMABs are Markov Decision Processes with two actions (active and passive modes) for each arm and with a constraint on the number of active arms per time slot. Since in general RMABs are PSPACE-complete, several heuristics such as Whittle index and LP index have been proposed. In this talk, I present a reinforcement learning scheme for LP indices with almost sure convergence guarantee in the tabular setting and an empirically efficient Deep Q-learning variant. Several examples, including scheduling in queueing systems, will be presented. This is a joint work V.S. Borkar and P. Shah from IIT Bombay.

Primary author: Dr AVRACHENKOV, Konstantin (INRIA Sophia Antipolis)

Co-authors: Prof. BORKAR, Vivek (IITB); Mr SHAH, Pratik (IITB)

Presenter: Dr AVRACHENKOV, Konstantin (INRIA Sophia Antipolis)

Session Classification: Parallel session: Some applications of reinforcement learning to networks