



Master thesis No. 1082

Designing and Implementing Exploration Strategies for RL Network Agents



Methods

Simulation
Machine Learning

Topics

Network control
Reinforcement Learning

Background

Modern communication networks operate in large, high-dimensional state spaces with complex dynamics and strict performance constraints. Reinforcement learning (RL) has emerged as a promising approach for adaptive network control, enabling policies that react to time-varying traffic demands and network conditions. However, RL in networking is fundamentally limited by sample efficiency and exploration: naive exploration strategies are too slow or unsafe for realistically sized networks. Recent advances in intrinsic-motivation-based exploration address this limitation by explicitly rewarding novelty. Methods such as Random Network Distillation (RND) quantify state novelty via prediction error, enabling targeted exploration of rarely visited network states. In parallel, Integer Linear Programming (ILP) models are widely used in networking to compute optimal or near-optimal solutions offline, providing structured, high-quality reference data. Combining ILP-derived information with RL exploration promises faster convergence and improved policy quality by guiding exploration toward relevant regions of the state space.

Problem Description

The objective of this student work is to study and implement exploration strategies for RL-based network control that leverage intrinsic motivation and optimization-based prior knowledge. The focus lies on Random Network Distillation and on exploiting existing ILP solution data to bias or guide the exploration process. The resulting RL agent should learn effective control policies significantly faster than with uninformed exploration. The work consists of the following tasks:

- Familiarization with the given network model, simulation environment, and ILP datasets
- Formalization of the network control problem as a Markov decision process
- Implementation of an RL agent with intrinsic exploration based on Random Network Distillation
- Integration of ILP-derived information to shape exploration or initialization
- Quantitative evaluation of learning speed, stability, and final performance against standard exploration baselines

Acquired Knowledge and Skills

The student will acquire in-depth knowledge of reinforcement learning for networked systems, with a strong focus on exploration mechanisms and intrinsic reward design. They will gain practical experience with Random Network Distillation, hybrid learning approaches combining optimization and RL, and the use of ILP data for learning guidance. Furthermore, the work strengthens skills in scientific programming, experimental evaluation of learning algorithms, and advanced networking concepts related to dynamic resource allocation and control.

Requirements

Programming Experience

Desirable knowledge

Communication Networks Architecture and Design

Contact

M.Sc. Nicolas Hornek
room 1.402 (ETI II), phone 685-67992, E-Mail nicolas.hornek@ikr.uni-stuttgart.de