Self-Driving Network and Service Coordination Using Deep Reinforcement Learning

Stefan Schneider, Adnan Manzoor, Haydar Qarawlus, Rafael Schellenberg, Holger Karl, Ramin Khalili, Artur Hecker

Paderborn University, Germany
Huawei Munich, Germany

March 8, 2021
IETF 110 NMRG Meeting
Scenario & Motivation
Scenario & Motivation
## Limitations of Existing Work

| Existing work:                           | Proposed approach:                              |
|-----------------------------------------|------------------------------------------------|
| · Mid-/Long-term planning per deployment request | · Fast online coordination of rapidly arriving user flows |
| · Rigid models tailored to specific scenarios | · Self-adapt to new scenarios and objectives |
| · Global, a prior knowledge              | · Partial, delayed observations                |

› Self-learning coordination with model-free DRL
Approach: Overview

Schneider et al.: Self-Driving Network and Service Coordination Using Deep Reinforcement Learning
Approach: Joint Scaling, Placement, Scheduling

- Scheduling: Where to process incoming flows?
- Derive scaling and placement automatically

| $S$ | $M$ | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ |
|-----|-----|-------|-------|-------|-------|-------|
| $s_1$ | $c_1$ | 10\% | 40\% | 50\% | 0\% | 0\% |
| $s_1$ | $c_2$ | 0\% | 0\% | 0\% | 0\% | 100\% |
| $s_2$ | $c_1$ | 50\% | 50\% | 0\% | 0\% | 0\% |
Approach: Partially Observable MDP

- Observations:
  - Avg. incoming data rate per ingress node and service
  - Max. resource utilization per node

- Actions:
  - Scheduling probability per node, service, component
  - Probability distribution over all possible target nodes

- Reward:
  - Fraction of successful vs. dropped flows
  - Negative end-to-end delay
Approach: DRL Framework

- Deep Deterministic Policy Gradient (DDPG)
- Offline training → focus on exploration
- Online inference → fast exploitation

Schneider et al.: Self-Driving Network and Service Coordination Using Deep Reinforcement Learning
Evaluation: Setup

- 4 real-world network topologies
- Service:
- Traffic: Varying stochastic arrival patterns

Baseline Algorithms:
- BSP: State-of-the-art heuristic
- SP: Shortest path-based heuristic
- LB: Equal load balancing
Evaluation: Maximizing Successful Flows

- Abilene topology with increasing load
- Different flow arrival patterns:
  - Fixed, Poisson, MMPP, real-world traces

Our approach:
- Self-adapts to varying traffic load & traffic patterns
- Processes more flows successfully than all baselines
- Generalizes to unseen traffic patterns
- Supports optimizing multiple objectives
- Scales to large networks
Conclusion: Challenges in AI for Network Management

- Solved challenges:
  - Lack of data. Selecting a suitable AI approach → RL
  - Selecting a suitable RL algorithm → DDPG
  - Difficult debugging until first working version
  - Careful definition of MDP, particularly, reward function

Towards truly driverless networks in practice

- Open challenges:
  - Standard benchmarks (cf. Atari, Mujoco)
  - sim2real gap, Safe & Explainable AI, Robustness
  - Generalization, Sample-efficient online learning
  - Combine expert knowledge and AI

Still many open challenges