Comparison of Traffic Signal Control Algorithms in a Large-Scale Traffic Simulation Environment

Anton Agafonov, Evgeniya Efimenko

САМАРСКИЙ УНИВЕРСИТЕТ
SAMARA UNIVERSITY

PROBLEM FORMULATION:

A traffic signal control problem as a reinforcement learning problem

- Each intersection in transportation network is controlled by an RL agent.
- The aim is to learn a policy for each agent that optimizes traffic situations for each intersection
- The traffic signal control problem - a Markov decision process – $\langle S, A, P, R, \gamma \rangle$ (S - set of states, A - set of actions, P - transition function, R - reward function, $\gamma \in [0, 1]$ - discount factor)
- The goal of an RL-agent is to find an optimal policy $\pi^*: S \rightarrow A$ maximizes the expected cumulative reward $E[R_t|s, \pi]$ for each state s , R_t - cumulative future discounted reward

$$R_t = \sum_{i=0}^{T-1} \gamma^i r_{t+i}$$

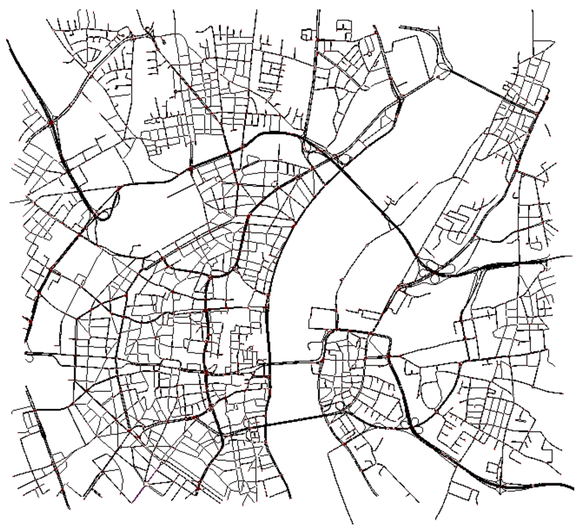
r_t - immediate reward at a time step t, T - total number of time steps

RL ALGORITHMS FOR TRAFFIC SIGNAL CONTROL:

- MaxPressure
- Gap-based actuated traffic control (SOTL)
- Independent DQN (IDQN), each intersection is controlled by a DQN-agent
- Independent Proximal Policy Optimization (IPPO)
- Advantage Actor-Critic (A2C)
- Soft Actor-Critic (SAC)

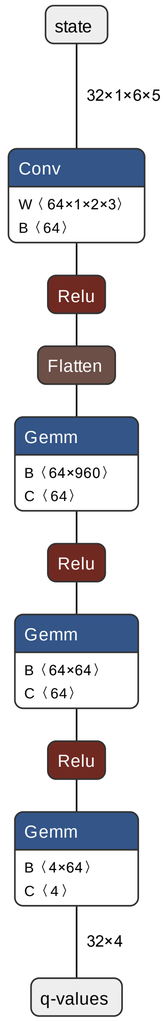
EXPERIMENTAL CONDITIONS:

- SUMO vehicle movement simulation system
- SUMO "TAPAS Cologne" movement scenario
- 2928 junctions of various configurations
- 316 signalized intersections
- Vehicle traffic from 7 am to 9 am



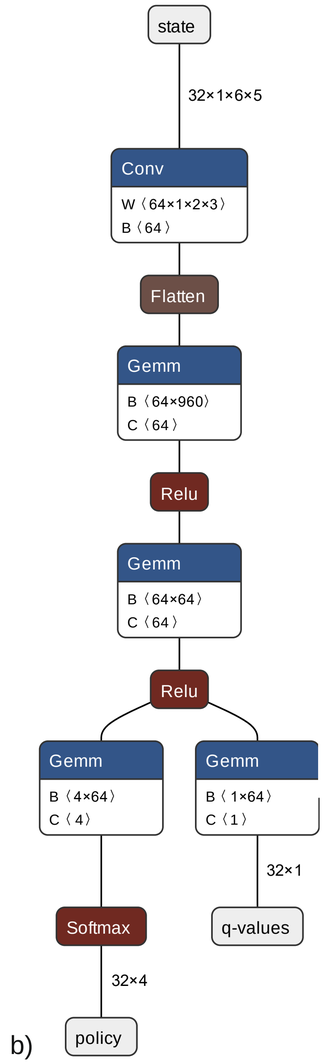
THE ARCHITECTURE OF DEEP RL MODELS:

a) DQN



a)

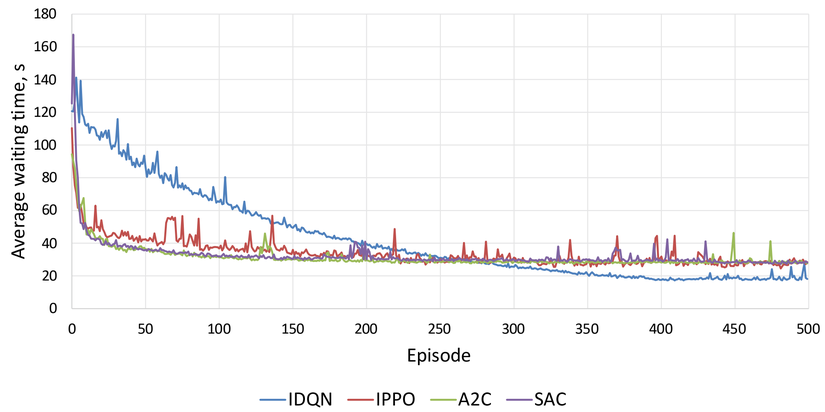
b) Actor-Critic



b)

COMPARISON OF TRAFFIC SIGNAL CONTROL ALGORITHMS

	Average waiting time	Average travel time
MaxPressure	27.35 ± 1.15	333.84 ± 2.3
SOTL	40.7 ± 1.25	357.9 ± 3.88
IDQN	**17.97 ± 0.24**	**317.92 ± 0.56**
IPPO	26.49 ± 1.07	329.36 ± 3.36
A2C	28.17 ± 0.25	336.38 ± 0.39
SAC	29.1 ± 0.28	336.33 ± 0.5



CONCLUSION:

Adaptive deterministic algorithms and algorithms, which are founded on reinforcement learning for solving the problem of controlling traffic signals, are compared (value-based (Q-learning) and policy-based (PPO, Actor-Critic) reinforcement learning algorithms and classical (SOTL, MaxPressure) approaches). The results show that the IDQN algorithm showed the best results by the selected criteria. It also should be noted that the deterministic MaxPressure algorithm showed similar results with other RL algorithms.

Acknowledgements:

The work was supported by the Russian Science Foundation grant No. 21-11-00321, https://rscf.ru/en/project/21-11-00321/