Online a model-free adaptive traffic signal controller for an isolated intersection

D. Kurmankhojayev, N. Suleymenov, G. Tolebi

September 18, 2017

Statement of the problem

Develop a model-free adaptive traffic signal control system with one of the well researched methods.

Current solution: Reinforcement Learning (RL)

Agent-based RL Intersection Controller

- One-step Q-learning (Watkins, 1989) - off-policy TD control algorithm

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right) \quad (1)$$

Reinforcement learning (RL)

Markov Decision Processes (MDP):

- States: S
- Model: T(s,a,s') $\sim$ P(s'|s,a)
- Actions: A(s), A
- Reward: R(s), R(s,a), R(s,a,s')

Policy: $\pi(s) \rightarrow a$

Reinforcement learning (RL)

Markov Decision Processes (MDP):

- States: S
- ~~Model: T(s,a,s') $\sim$ P(s'|s,a)~~
- Actions: ~~A(s)~~, A
- Reward: ~~R(s)~~, R(s,s'), ~~R(s,a,s')~~

Policy: $\pi(s) \rightarrow a$

RL context

$$< s, a, r >^* \rightarrow \boxed{\text{RL algorithm}} \rightarrow \pi$$

- **Model-based**

  $< s, a, r >^* \rightarrow \boxed{\text{model learner (T/R)}} \rightarrow \boxed{\text{MDP solve } (Q^*)} \rightarrow \boxed{\text{argmax}} \rightarrow \pi$

- **Model-free (Value-function-based)**

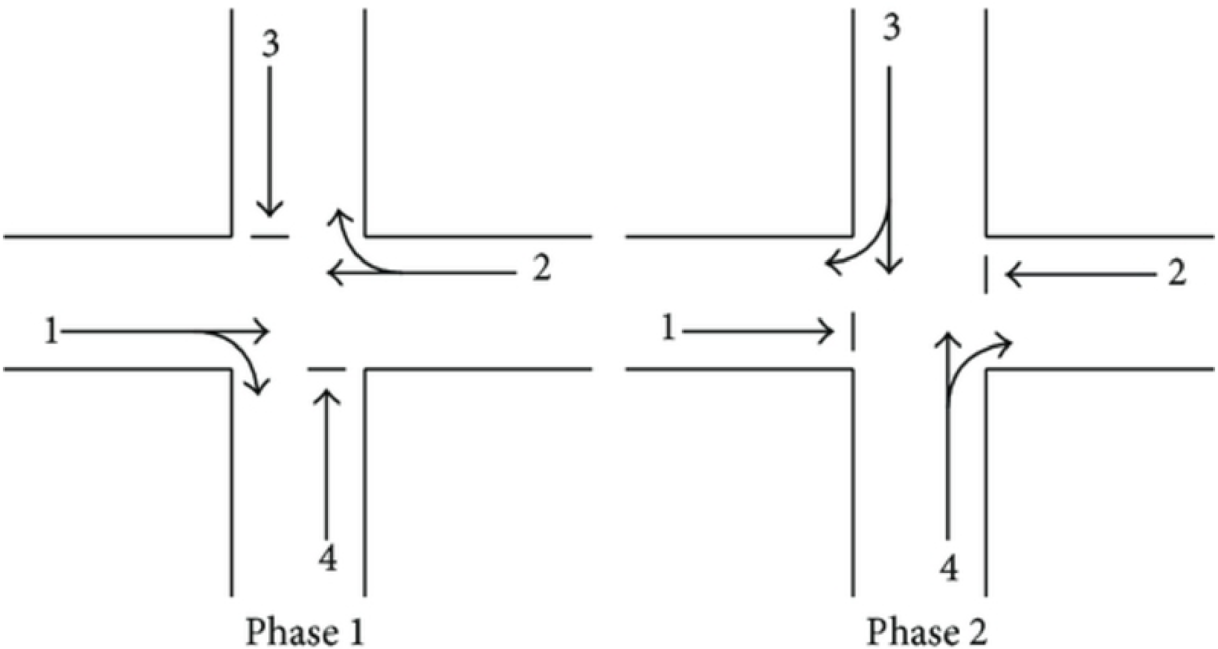
  $$< s, a, r >^* \rightarrow \boxed{\text{value update (Q)}} \rightarrow \boxed{\text{argmax}} \rightarrow \pi$$

- **Policy search**

  $$< s, a, r >^* \rightarrow \boxed{\text{policy update}} \rightarrow \pi$$

Parameters of a traffic signal control

- Cycle is a repetition of the basic series of signal combinations.
- Phase is a part of the signal cycle, during which one set of flows has right of way
- Split is an amount of green for each approach



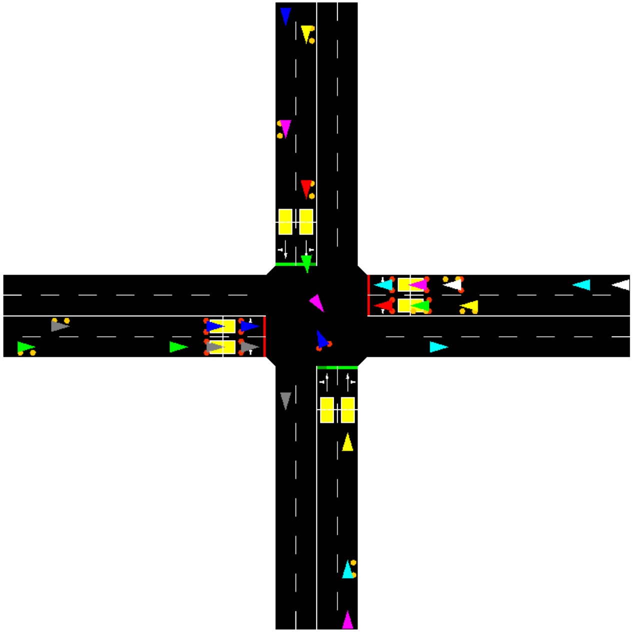
Test Intersection

General assumptions:

- Length of an edge is 100 meters
- An edge has 2 lanes
- Average vehicle's length is 5 meters
- Total number of halting vehicles is in the range $0 \leq N \leq 80$.

The problem was simplified by excluding the following:

- U-turns, public transport flows, pedestrian flows, traffic rules violation, traffic road accidents, and parking.
- We assume that the acquired data are free of noise and available on request.



Experimental environment SUMO

SUMO (Simulation of Urban MObility) [1] - is an open-source traffic simulation package, which has following features:

- Microscopic simulation of traffic flow
- Online control of simulation process (TraCI)
- Simulation of multimodal traffic (vehicles, public transport and pedestrians)
- No limitations in network size and number of simulated vehicles
- Supported import formats: OpenStreetMap, VISUM, VISSIM, NavTeq.

State space

Input is a number of halting vehicles in North-South, South-North and West-East, East-West directions.

We represent a state as a set of two components:

- number of halting vehicles approaching the intersection from North and South are summed up
- number of halting vehicles approaching the intersection from West and East are summed up

State component	Short	Description
Low	L	$0 \leq n \leq 30$
Medium	M	$30 < n \leq 60$
High	H	$60 < n \leq 90$

Action space

Given a fixed cycle length we extend/shrink duration of green phase in NS/WE directions.

Actions	Seconds
a_0	$(0, 0)$
a_1	$(+dt, -dt)$
a_2	$(-dt, +dt)$

Reward

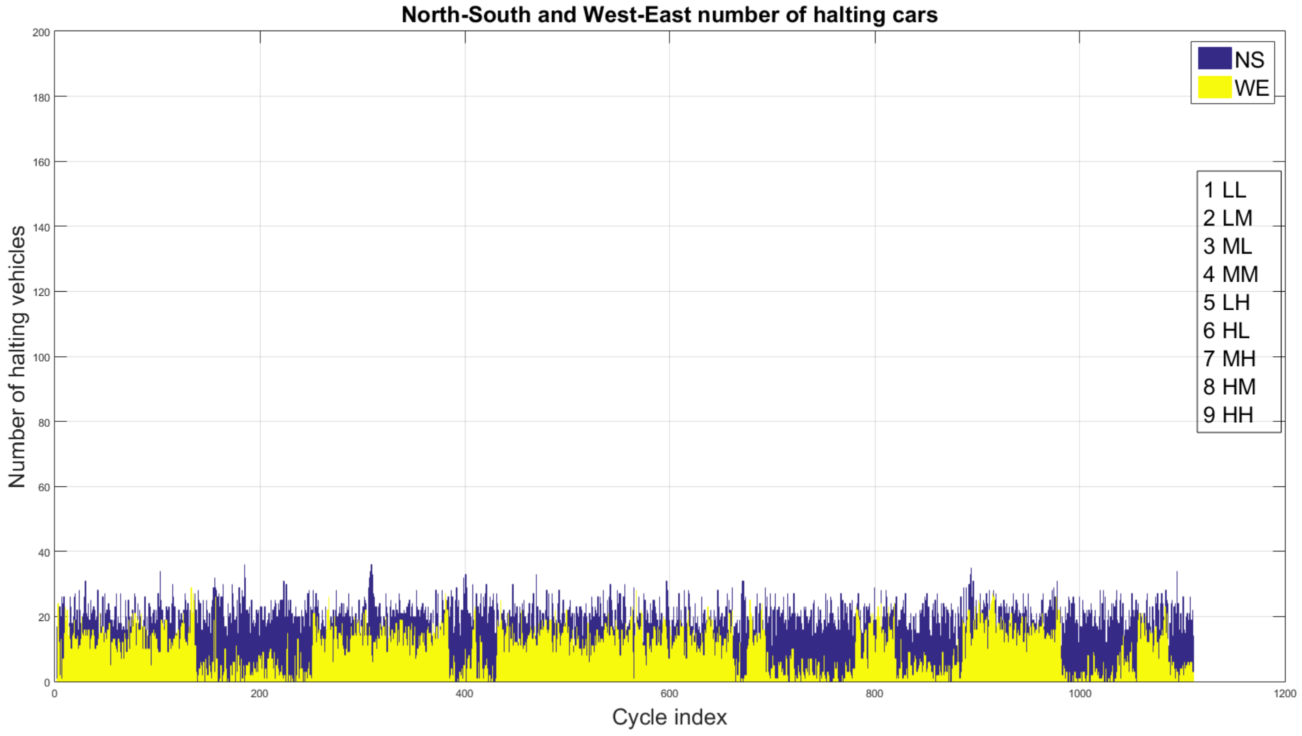
Reward function $R(s, s')$:

$$R = w_1 \overbrace{\left(|NS - WE| - |NS' - WE'|\right)}^{\text{equlibrium term}} + w_2 \underbrace{\left((NS + WE) - (NS' + WE')\right)}_{\text{queue reduction term}} \quad (2)$$

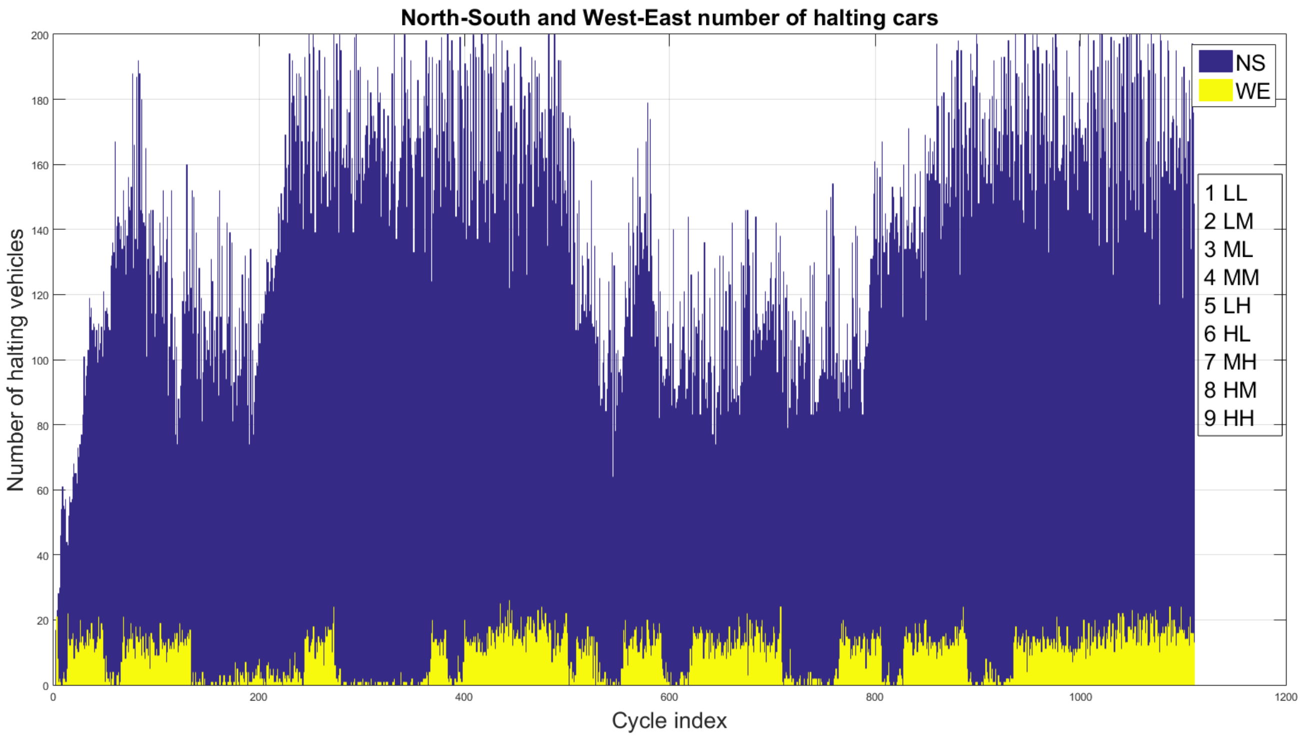
where

- w_1 and w_2 are control weights
- $s = [NS, WE]$ and $s' = [NS', WE']$ are previous and current states.

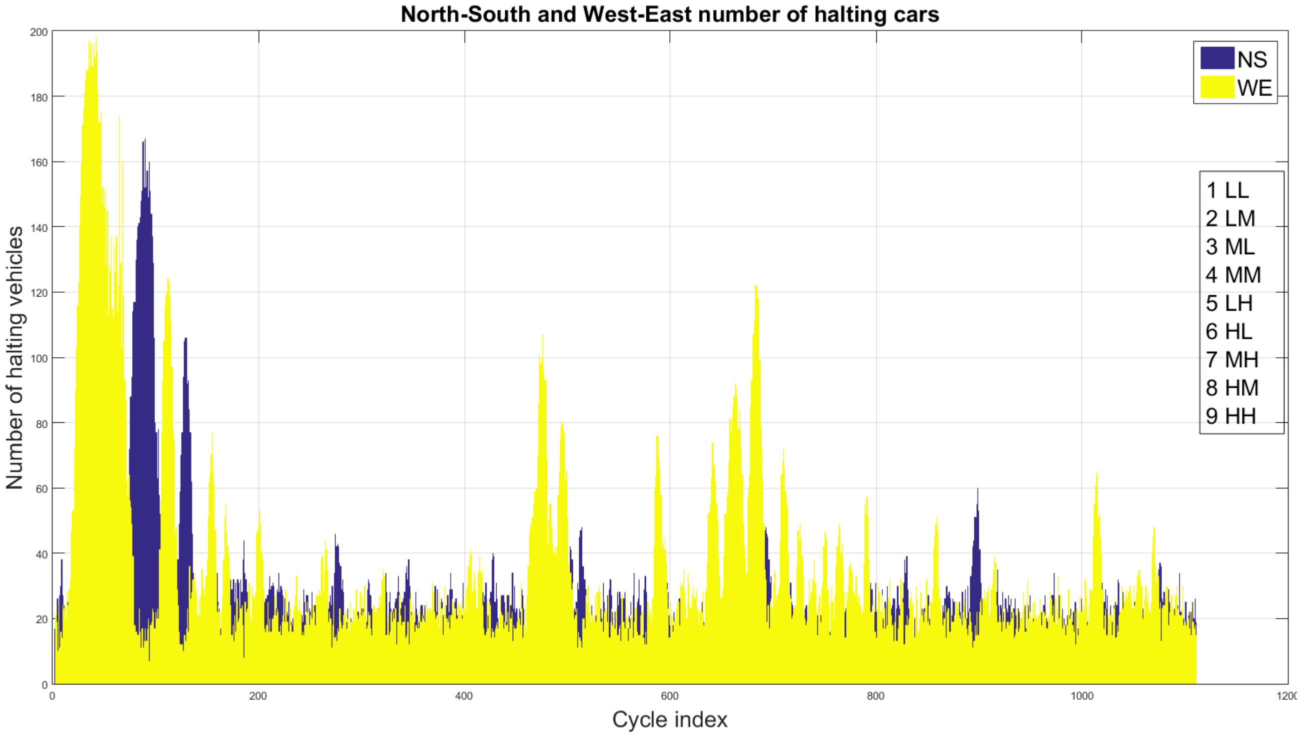
Results: Fixed symmetric signal plan



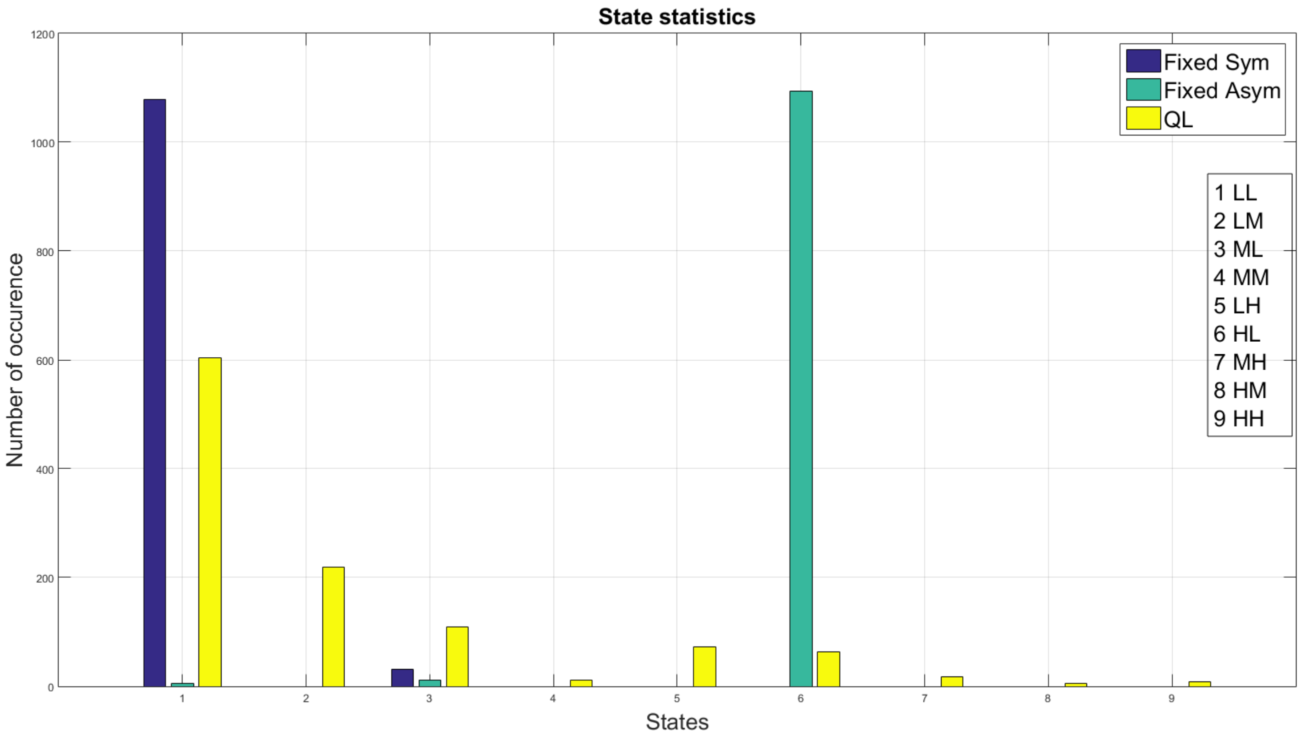
Results: Fixed asymmetric signal plan



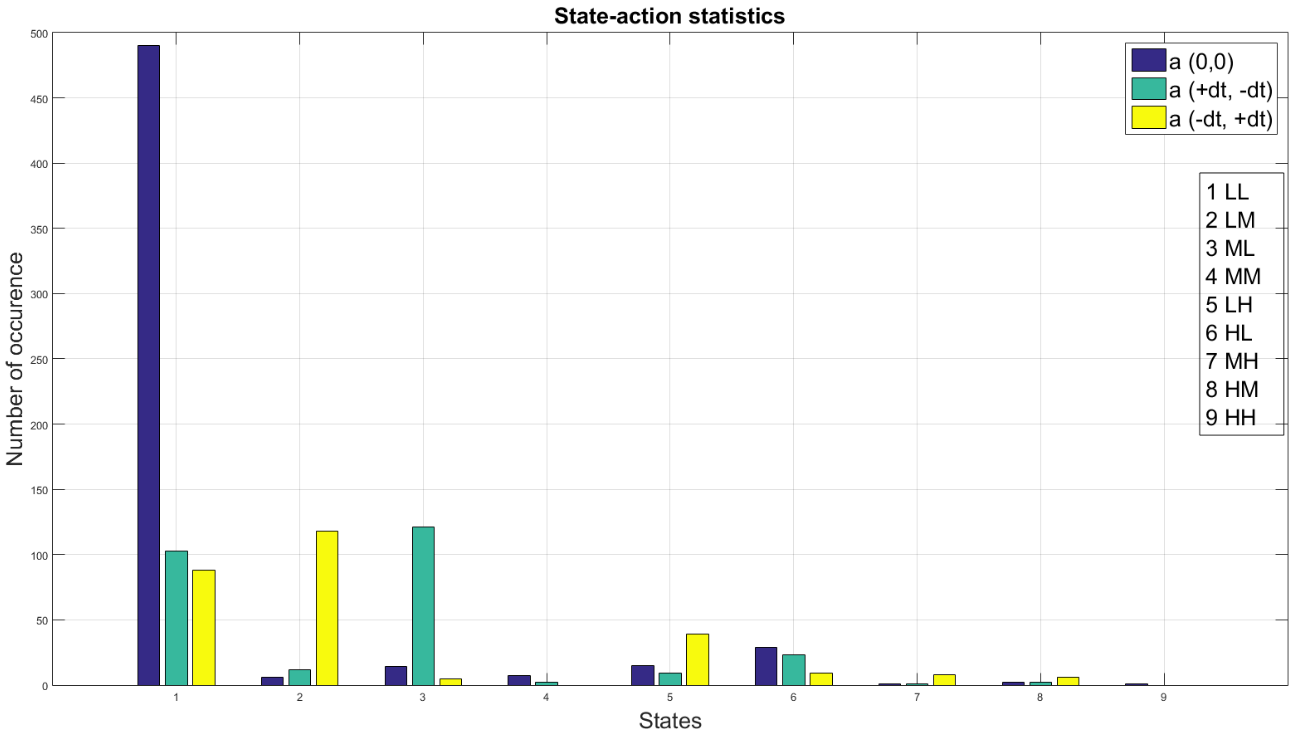
Results: Adaptive controller



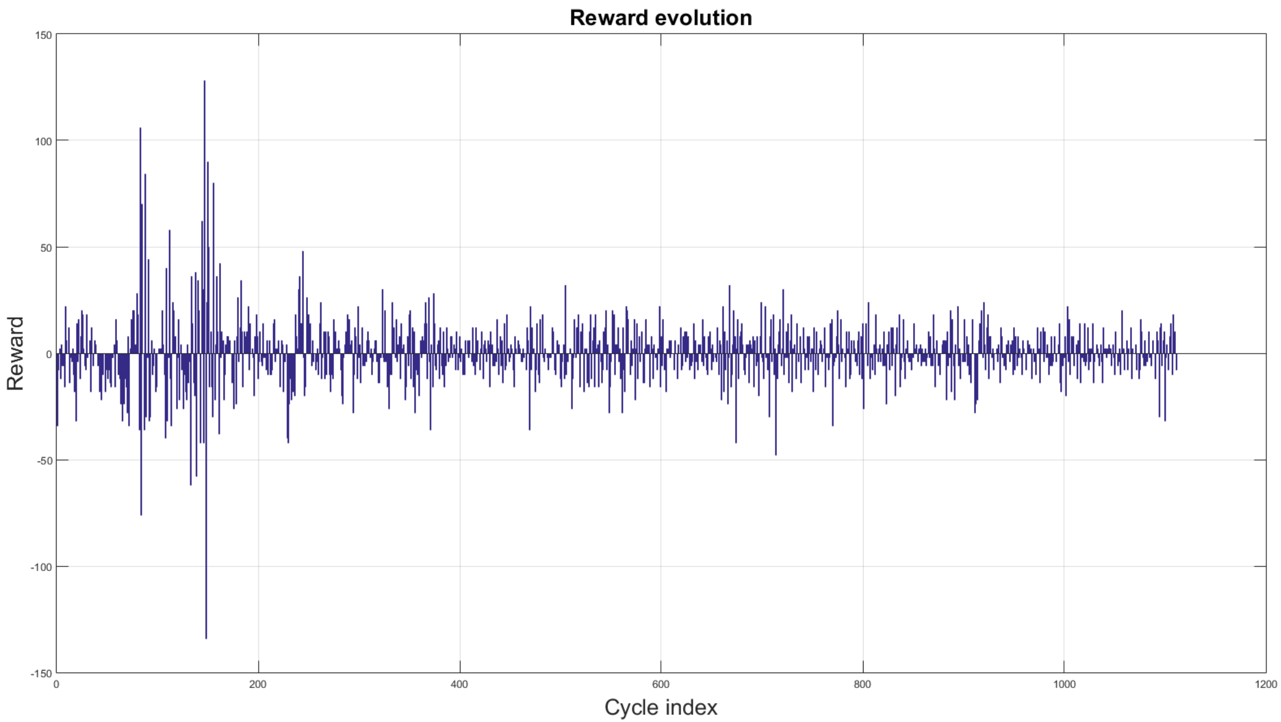
Results: State statistics



Results: State-action statistics for adaptive controller



Results: Reward evolution



Conclusion

Proposed model-free adaptive signal controller based on one-step Q-learning algorithm. The results of the experiments show that the method approached a near-optimal performance.

Future work:

- Replace the discrete state space by a continuous analog
- Apply proposed method on a multi-intersection network

References

SUMO – Simulation of Urban MObility. Institute of Transportation Systems.
http://sumo.dlr.de/wiki/SUMO.

Thank you for attention!