A COMPARATIVE STUDY OF ALGORITHMS FOR INTELLIGENT TRAFFIC SIGNAL CONTROL

Hrishit Chaudhuri, Vibha Masti, Vishruth Veerendranath, Dr. S Natarajan

Department of Computer Science, PES University, RR Campus, Bengaluru

SUMMARY OF CONTENTS

1 The Problem And Introduction

2 Experimental Setup

3 Algorithms Compared

4 Results

5 Conclusions

THE PROBLEM

- Huge increase in the number of personal vehicles
- Lot of time **(149% at peak-hour)** and money **($22 billion)** are wasted in Traffic congestion

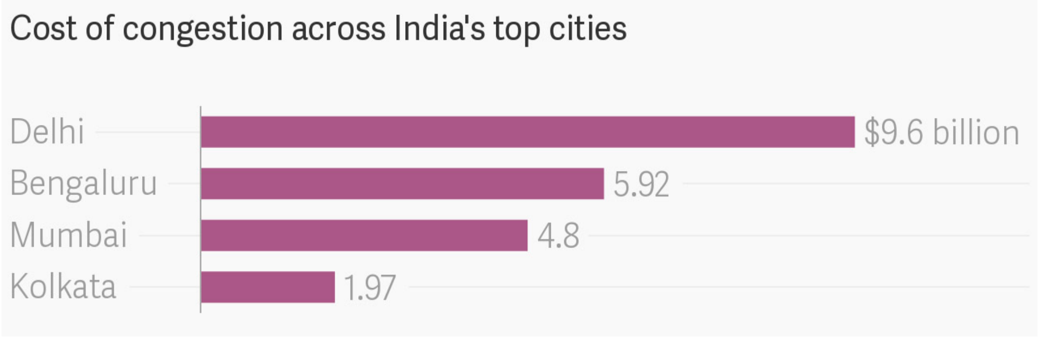


Fig. 1. Source: BCG-"Unlocking Cities, the Impact of ridership across India"

- Prevalence of simple but inefficient and outdated traffic signal policies
- *What are the better methods and algorithms to optimize traffic signals and what are the advantages and disadvantages of each of them?*

INTRODUCTION AND OBJECTIVE

- Quantitative Comparative analysis between four traffic signal control policies
- The policies/algorithms quantitatively compared are:
  - Round Robin (RR)
  - Feedback Control Mechanism (named MONOPOLY)
  - Deep Q-Networks (DQN)
  - Advantage Actor-Critic (A2C)

- DQN and A2C – State-of-the-art *Reinforcement Learning* Techniques
- *Is the computational complexity of Reinforcement Learning worth it?*

RELATED WORK

Paper Referenced	Idea Inferred
Learning Phase Competition for Traffic Signal Control (G Zheng et al.)	Experimental setup and Deep-Q learning
Reinforcement Learning for Traffic Optimization (M. Stevens et al.)	Q-learning in traffic signals
IntelliLight: A Reinforcement Learning Approach for Intelligent Traffic Light Control (H. Wei et al.)	Comparison of DQN and baseline traffic signalling models
Asynchronous Methods for Deep Reinforcement Learning (V. Mnih et al.)	Asynchronous Actor-Critic RL
A deep reinforcement learning approach to adaptive traffic lights management (A. Vidali et al.)	Deep reinforcement learning for four-way intersection

EXPERIMENTAL SETUP

- Traffic Intersection modelled as Markov Decision Process (MDP)
- Four-way intersection, each arm consisting of two lanes (one incoming, one outgoing)
- *State of intersection* – state space of MDP (different for each algorithm)
- *Actions* – different for each algorithm; optimize either the order of green signals or the duration of green signals
- Traffic simulations using Simulation of Urban Mobility (SUMO), interfaced using TraCI

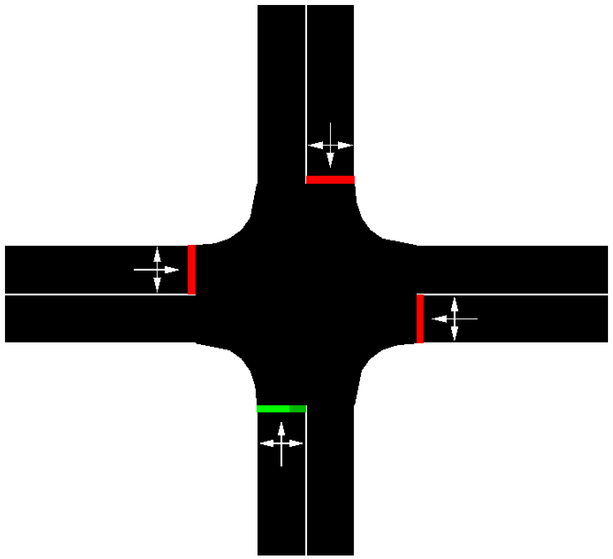


Fig. 2. Four-way single lane intersection (using Netedit)

ACTION SPACE

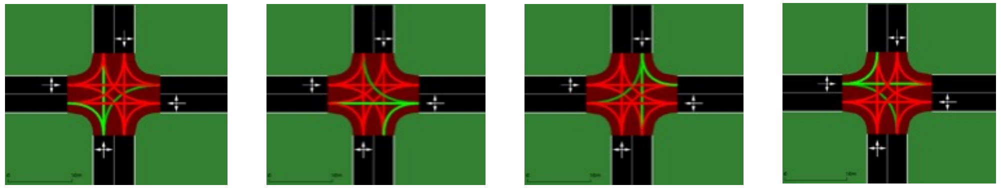


Fig. 3. AS 1: Four green-phase actions {S, E, N, W}

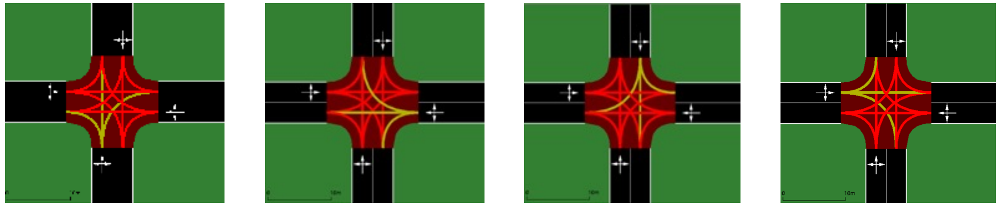


Fig. 4. Intermediate yellow phase actions associated with AS 1

- The action space is defined differently for different algorithms
- AS 1 – optimize **green signal order**
  AS 2 – optimize **green signal** timing
- AS 1: set of possible green phases
  - **`A = {S, E, N, W}`**
  - set of intermediate yellow phases also associated with AS 1
- AS 2: range of numbers from `min_time` to `max_time`
  - **`A = [min_time, max_time]`**

TRAFFIC SCENARIOS

- Traffic routes are created with traffic density generated according to a *Weibull distribution*
- Three traffic scenarios
  - Uniform (SCEN-1)
  - 90% NS (North-South) bias (SCEN-2)
  - 90% EW (East-West) bias (SCEN-3)
- Traffic probabilities for a single car as shown in Table 1

Direction	Source	Destination	Probability in SCEN-1	Probability in SCEN-2	Probability in SCEN-3
Straight	North	South	0.1875	0.3375	0.0375
	South	North	0.1875	0.3375	0.0375
	East	West	0.1875	0.0375	0.3375
	West	East	0.1875	0.0375	0.3375
Turn (Left or Right)	North	East	0.03125	0.05625	0.00625
	North	West	0.03125	0.05625	0.00625
	South	West	0.03125	0.05625	0.00625
	South	East	0.03125	0.05625	0.00625
	East	North	0.03125	0.00625	0.05625
	East	South	0.03125	0.00625	0.05625
	West	South	0.03125	0.00625	0.05625
	West	North	0.03125	0.00625	0.05625

Table 1. Traffic Probability for a single car

ROUND ROBIN SCHEDULING (RR)

- Circular Scheduling
- 1 quantum = 30 timesteps green & 3 timesteps yellow
- No regard for empty lanes - inefficient
- Simple, most prevalent
- No computational complexity
- Used as a benchmark to compare the other algorithms

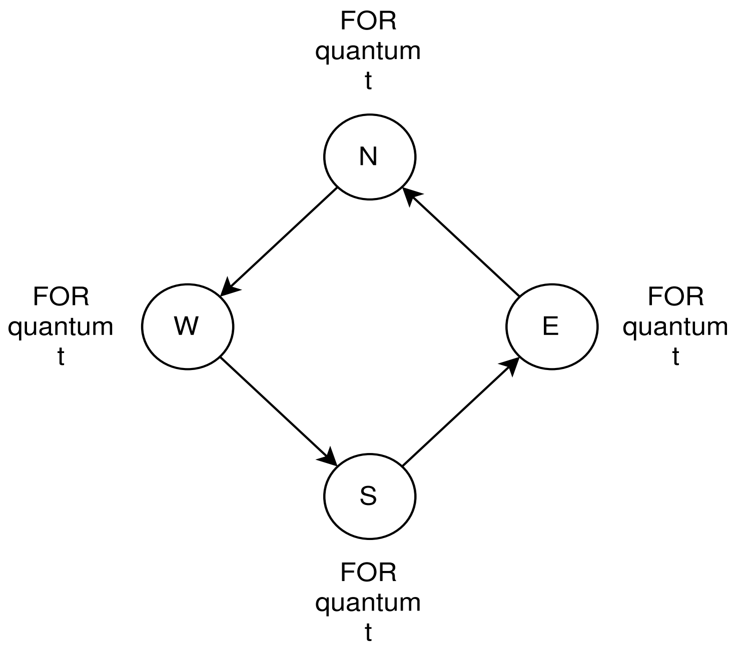


Fig. 5. Round Robin Scheduling

FEEDBACK CONTROL MECHANISM (MONOPOLY)

- Game-playing strategy
- Action space – AS 2

$$A = [min\_time,\ max\_time]$$

- Pick Max Reward over all Actions

$$R(s_t,\ a) = |\ s_t - v_t\ a\ |$$

s_t – state of corresponding lane at time t (queue length)

v_t – average speed of cars in the junction in the previous timestep

a – action currently being compared

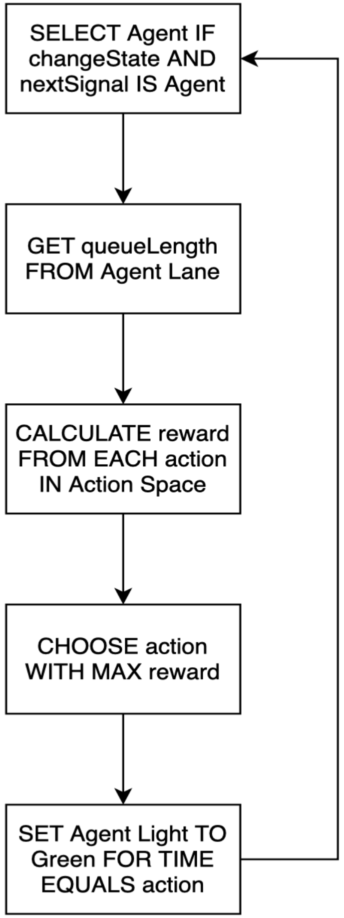


Fig. 6. Feedback Control Loop

DEEP Q-NETWORKS (DQN)

- Agent placed in state $\mathbf{s_t}$ and made to choose an action $\mathbf{a_t}$ from action space $\mathbf{AS}\ \ \mathbf{1}$ according to a **Markov Decision Process** to transition into next state $\mathbf{s_{t+1}}$, obtaining reward $\mathbf{r_{t+1}}$
- **Q-value**: Q(s,a) = Maximum Expected Rewards. Decides optimal policy.
- Deep neural network to estimate value of optimal $\mathbf{Q^*(s,\ a)}$ function, instead of a Q-table
- Q-function as proposed by Vidali et al. [7] (modification of Bellman Eq.)

$$Q(s_t,\ a_t) = r_{t+1} + \gamma \cdot \max_A Q'(s_{t+1},\ a_{t+1})$$

$\mathbf{s_t}$ – state of corresponding lane at time t (next slide)

$\max_A Q'(s_{t+1},\ a_{t+1})$ – max Q-value of taking action $\mathbf{a_{t+1}}$ in next state $\mathbf{s_{t+1}}$

γ – discount factor between 0 and 1

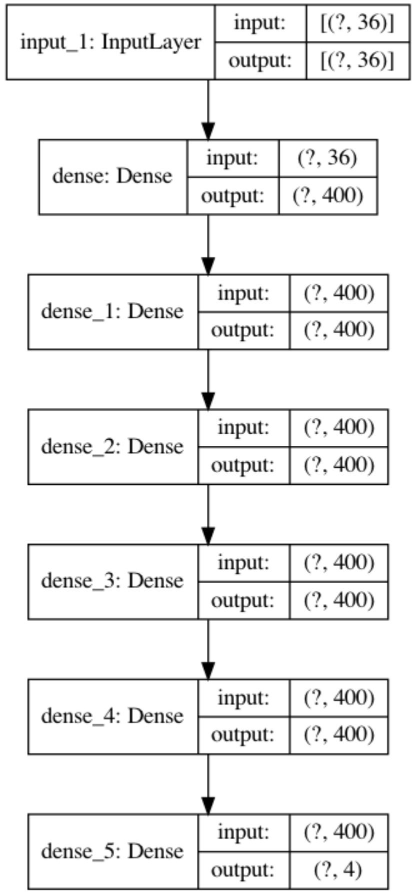


Fig. 7. Neural Network Architecture

STATE REPRESENTATION AND TRAINING IN DQN

- State $\mathbf{s_t}$ – Snapshot of the environment
- Each road divided into ***cells of varying size*** – cells closer to signal smaller for higher priority
- 9 cells per lane * 4 incoming lanes = 36 valued Boolean vector – 2^{36} total states

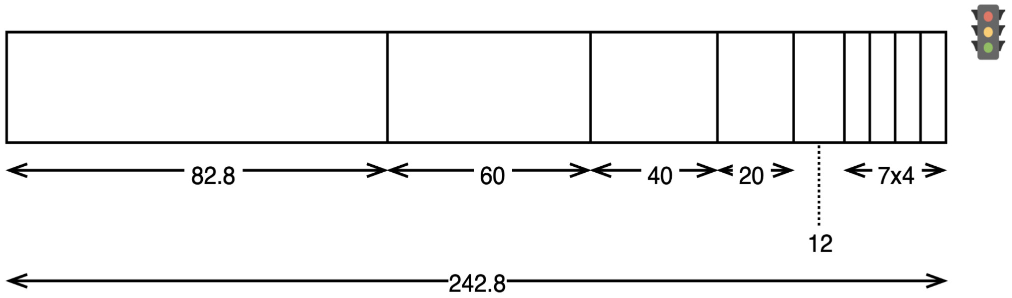


Fig. 8. State Respresentation

TRAINING

- ***Architecture*** – 1 input (state vector), 5 hidden, 1 output (4 valued vector/action space – which lane to make green)
- 100 episodes, Each episode – 800 times experience replay, $\gamma = 0.75$, learning rate = 0.001

ADVANTAGE ACTOR CRITIC (A2C)

- Improvement on DQN
- Actor-Critic Family of RL algorithms
- Underlying model – **Multilayer Perceptron (similar to DQN/ANN)**
- *Actor* – chooses particular action
- *Critic* – evaluates effect of action taken and sends feedback to actor
- *Advantage* – feedback, extra reward gained by taking this action compared to if other action had been taken
- *Parallelism* – multiple agents in independent environments **(multiple workers)**. Lesser Training time, CPU instead of GPU.

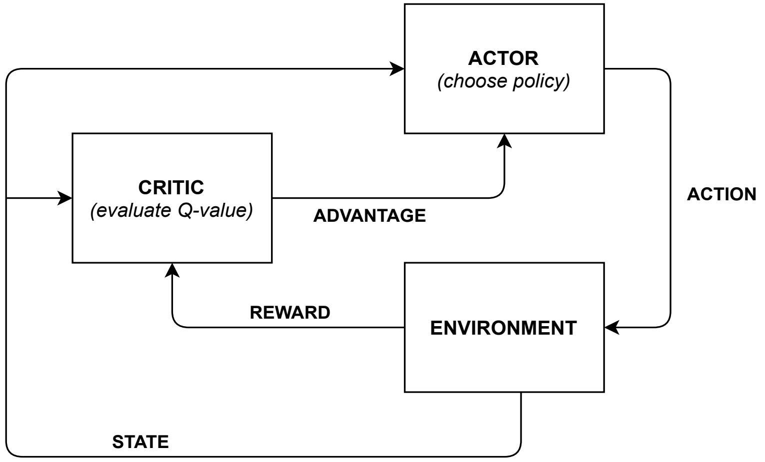
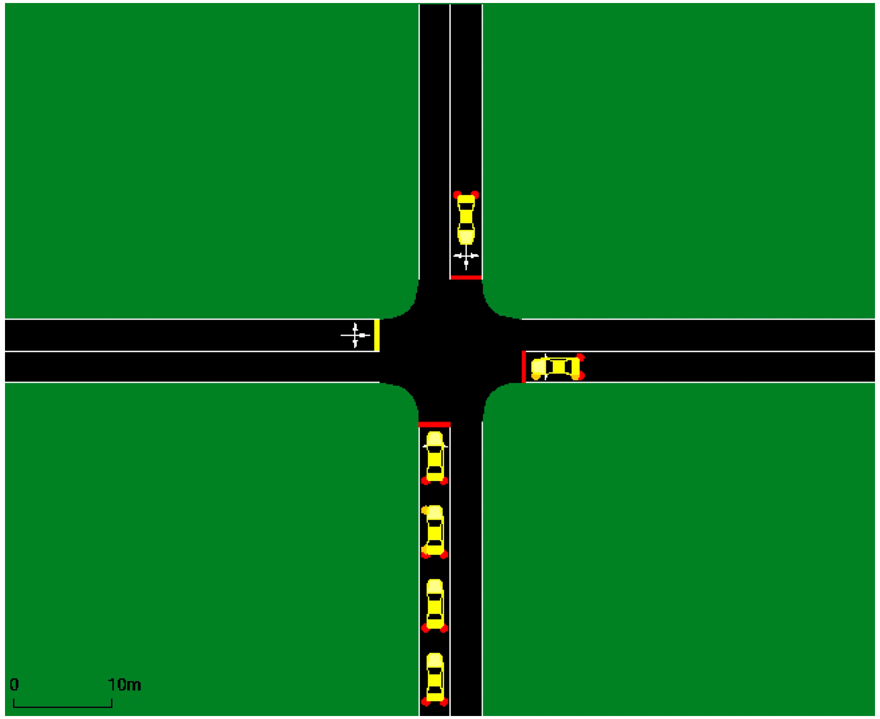
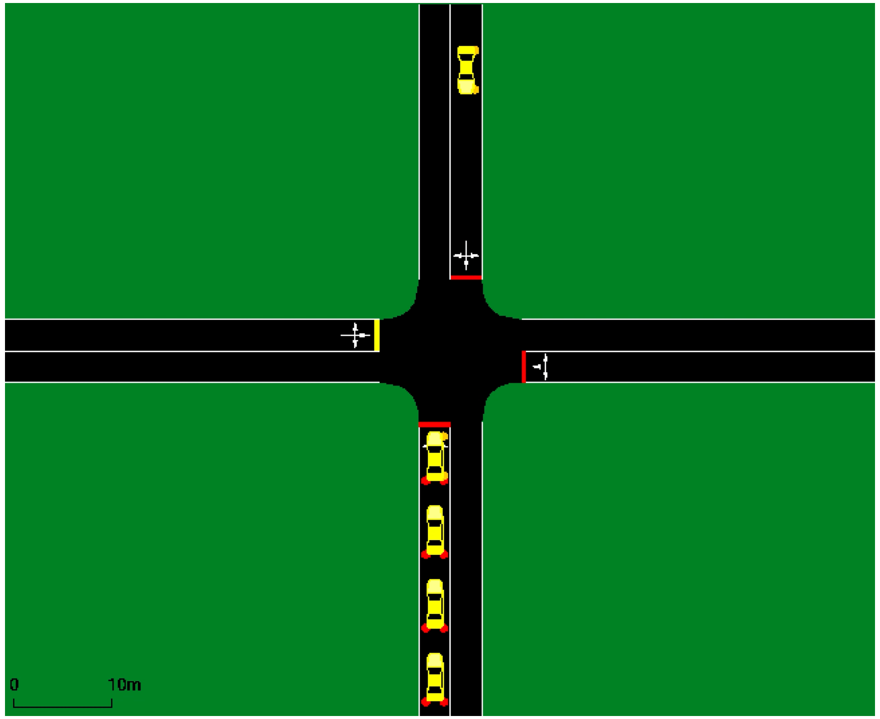


Fig. 9. A2C

SIMULATION IN SCEN-2

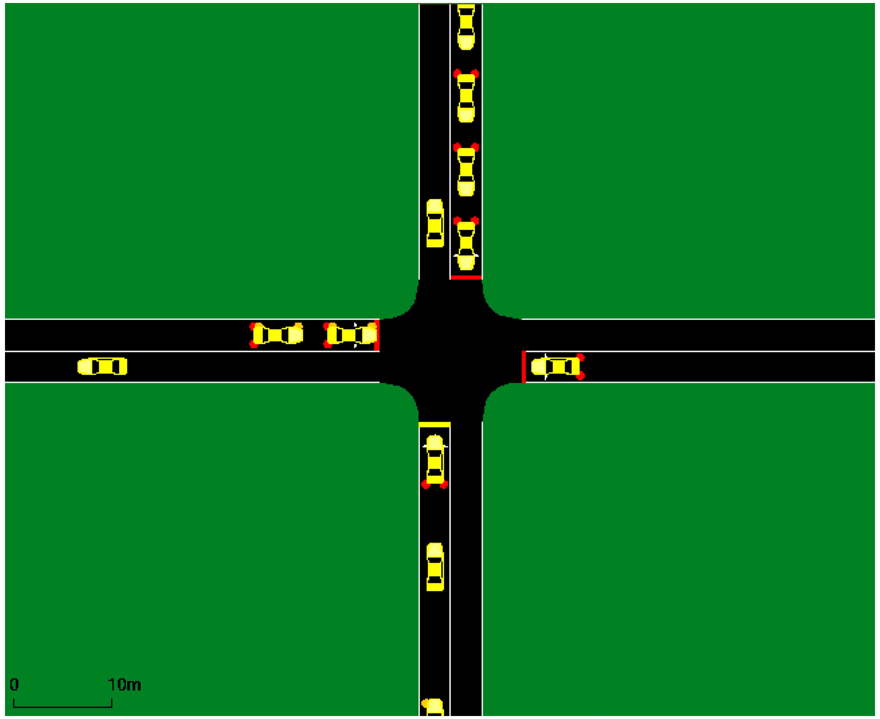


Vid. 1. 10 seconds of ***Round Robin*** in SCEN-2

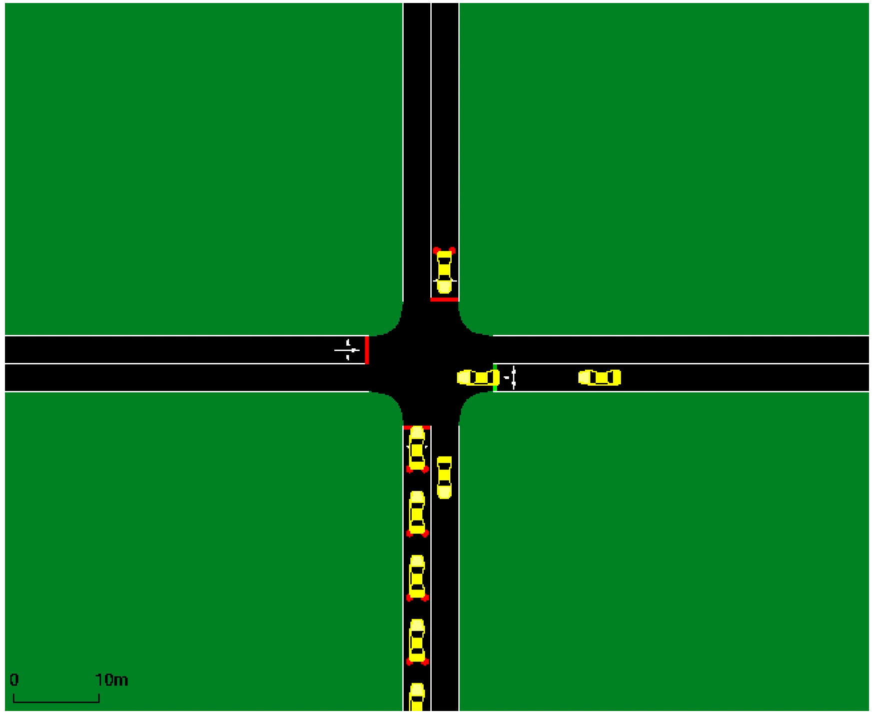


Vid. 2. 10 seconds of ***MONOPOLY*** in SCEN-2

SIMULATION IN SCEN-2 (CONTD.)



Vid. 3. 10 seconds of *DQN* in SCEN-2



Vid. 4. 10 seconds of **A2C** in SCEN-2

RESULTS (QUEUE LENGTH VS TIME)

Round Robin

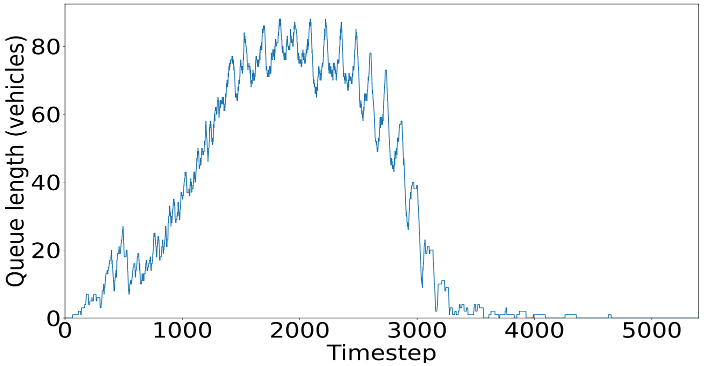


Fig. 10. (a) SCEN-1

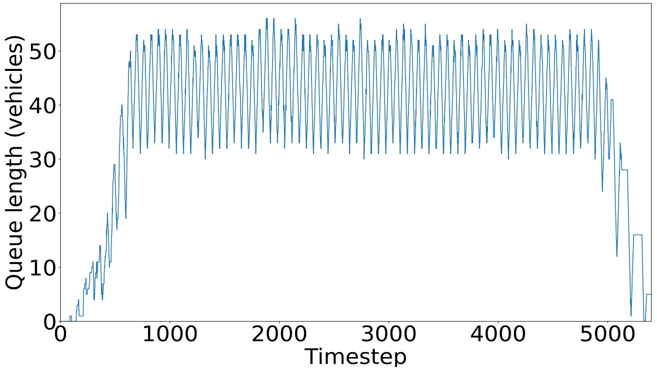


Fig. 10. (b) SCEN-2

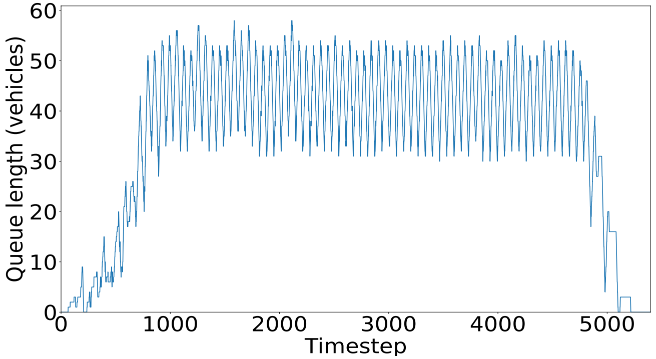


Fig. 10. (c) SCEN-3

Feedback Control Mechanism (MONOPOLY)

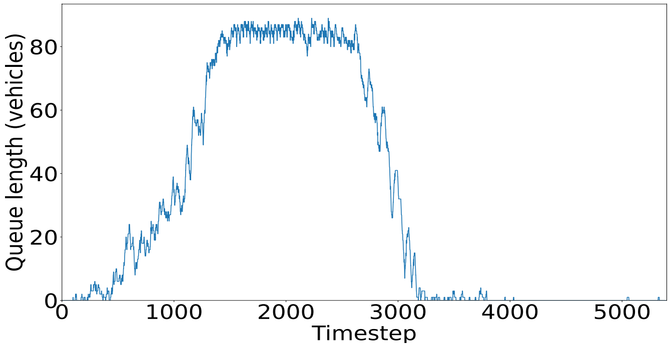


Fig. 11. (a) SCEN-1

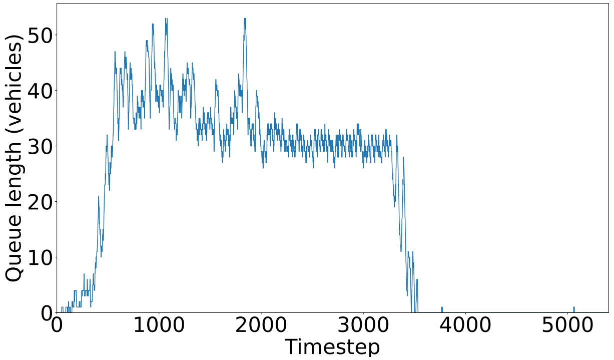


Fig. 11. (b) SCEN-2

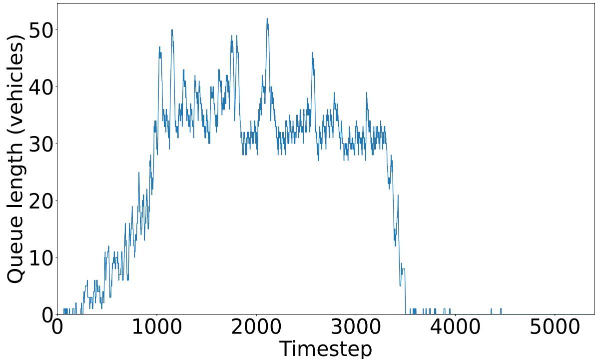


Fig. 11. (c) SCEN-3

RESULTS (QUEUE LENGTH VS TIME)

Deep Q-Networks (DQN)

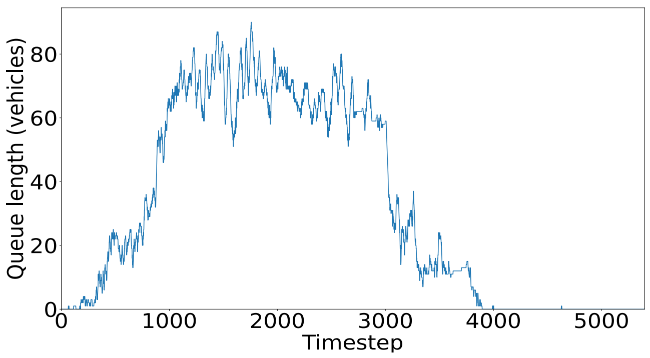


Fig. 12. (a) SCEN-1

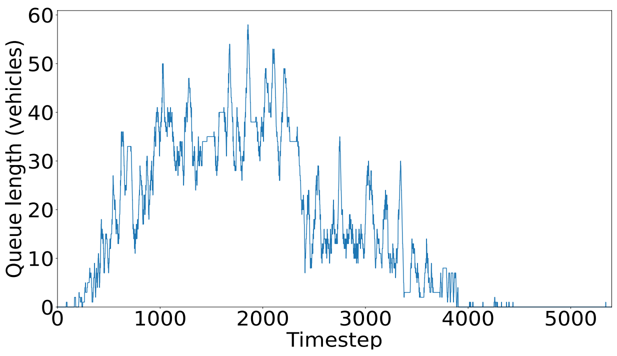


Fig. 12. (b) SCEN-2

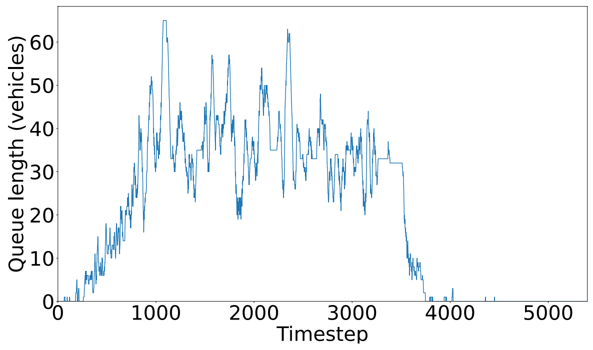


Fig. 12. (c) SCEN-3

Advantage Actor-Critic (A2C)

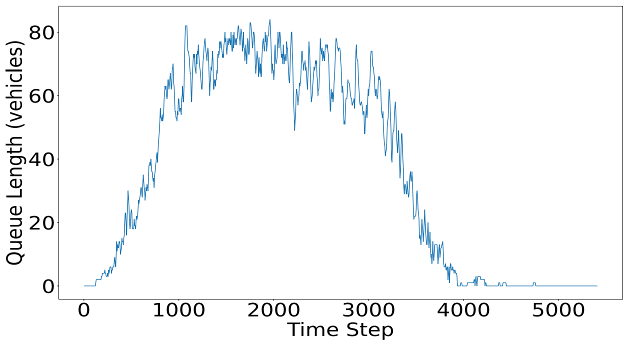


Fig. 13. (a) SCEN-1

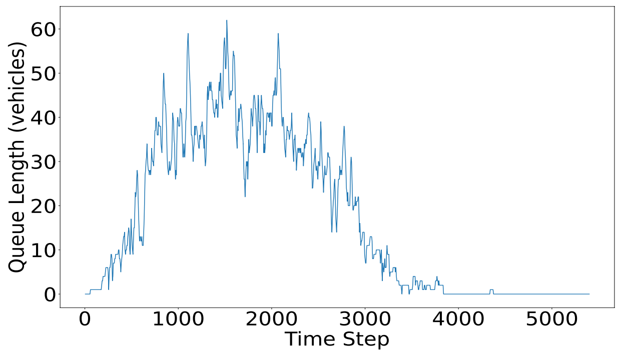


Fig. 13. (b) SCEN-2

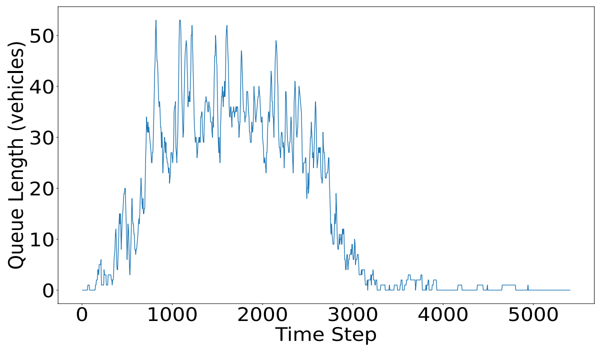
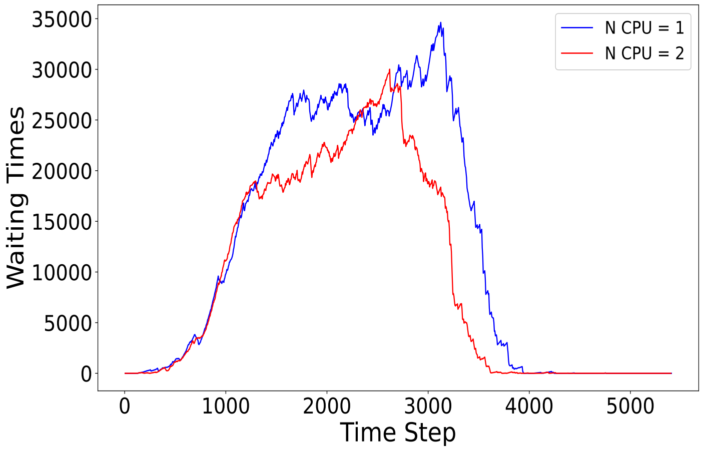


Fig. 13. (c) SCEN-3

RESULTS (PARALLELISM ADVANTAGES OF A2C)

N-CPU = 1, 2, 4 Comparison



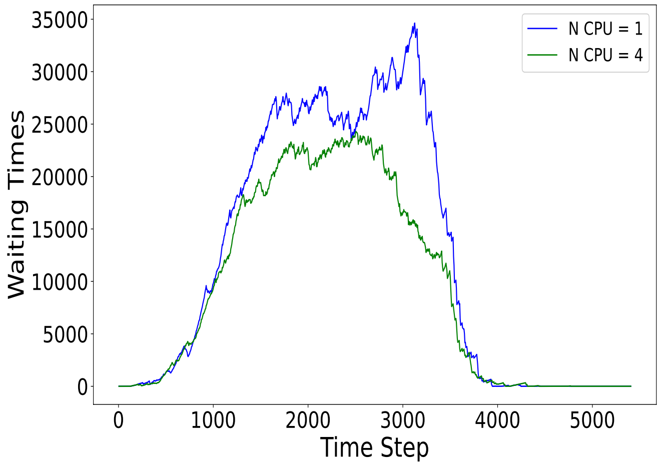


Fig. 14. N-CPU Comparisons

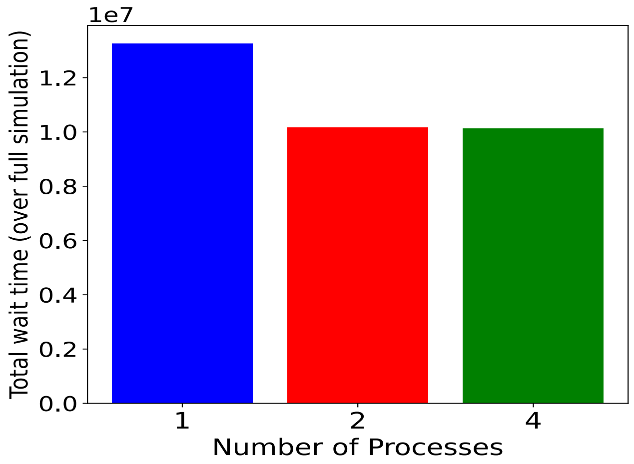


Fig. 15. Total Wait times Comparison

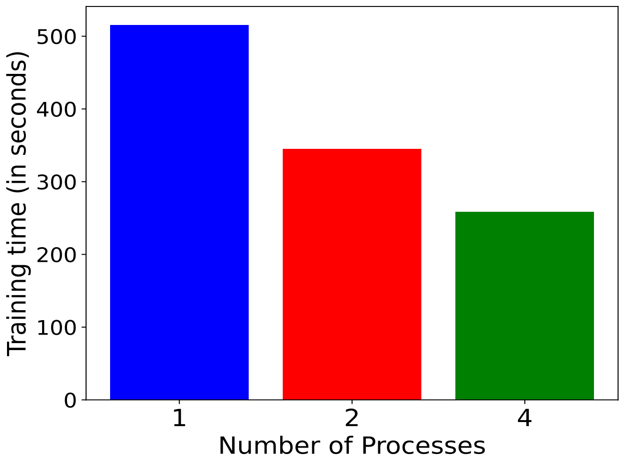


Fig. 16. Training Times Comparison

RESULTS (SUMMARIZED)

Algorithm	Peak Queue Length (no. of vehicles)			Fraction of time spent at above half queue length		
	SCEN-1	SCEN-2	SCEN-3	SCEN-1	SCEN-2	SCEN-3
RR	88	56	58	0.33	0.84	0.77
MONOPOLY	89	53	52	0.33	0.51	0.45
DQN	90	58	65	0.40	0.26	0.34
A2C	84	62	53	0.47	0.29	0.31

Table 2. Results of 4 algorithms

REAL WORLD INTERSECTION

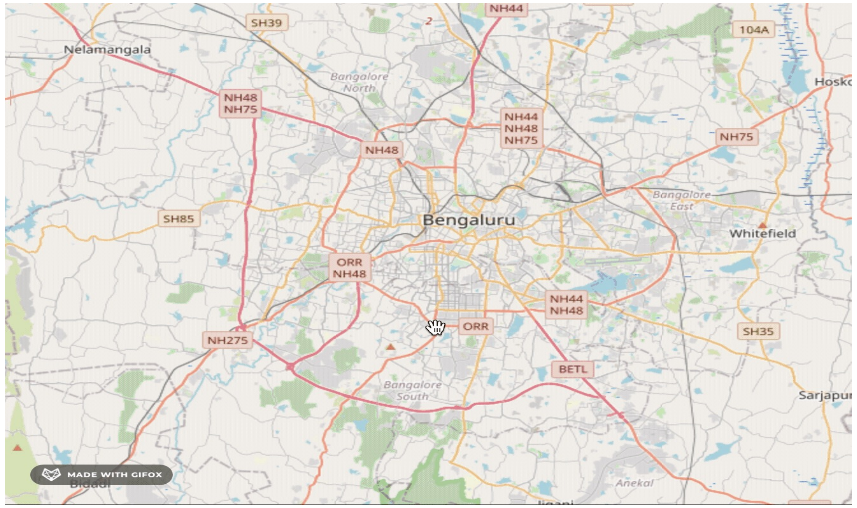
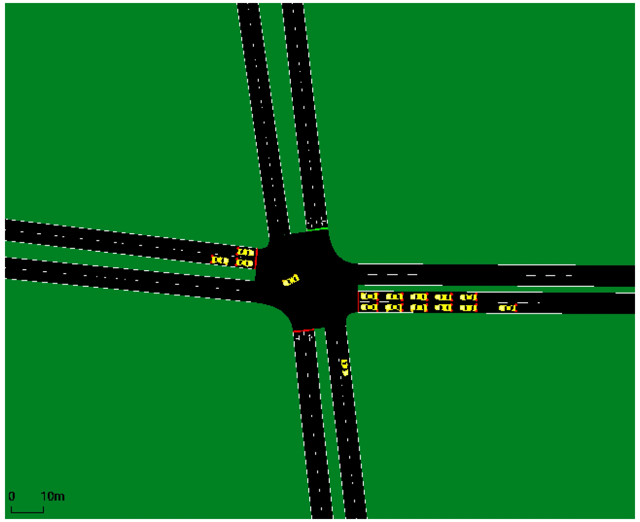


Fig. 17. Sarakki Intersection on OSM



Vid. 5. A2C on Sarakki Intersection in SCEN-1

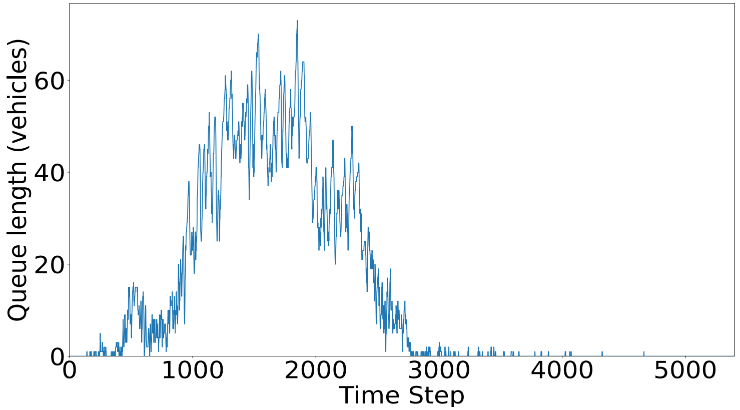


Fig. 18. Results of Round Robin on Sarakki in SCEN-1

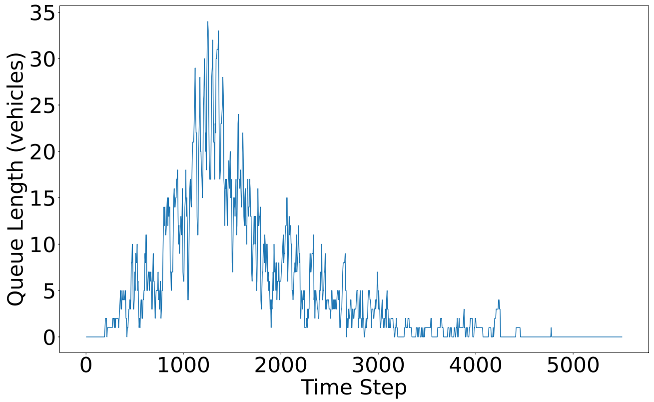


Fig. 19. Results of A2C on Sarakki in SCEN-1

CONCLUSION

- Round Robin
  - *Advantage:* Computationally simple
  - *Disadvantage:* Inefficient in clearing out traffic in non-uniform traffic scenarios
  - *Use Case:* Low-density, uniform traffic intersections
- Feedback Control
  - *Advantage:* Better than RR in Non-uniform, best in terms of ***performance-computation tradeoff***
  - *Disadvantage:* Fails if Traffic density is extremely high in all lanes (reverts back to RR)
  - *Use Case:* Moderate traffic density with unavailability of intensive computational resources
- Deep Q-Networks
  - *Advantage:* State of the art, best performance
  - *Disadvantage:* High Training times, availability of computational resources
  - *Use Case:* High Traffic density with availability of intensive computational resources
- Advantage Actor-Critic
  - *Advantage:* State of the art, best performance, parallelism, lower training times than DQN, best of both worlds
  - *Use Case:* High Traffic density

REFERENCES

1. Wey, W. M.: Model Formulation and Solution Algorithm of Traffic Signal Control in an Urban Network. Computers, Environment and Urban Systems. 24(4), 355-378 (2000)
2. Eriskin, E., Karahancer, S., Terzi, S., Saltan, M.: Optimization of Traffic Signal Timing at Oversaturated Intersections Using Elimination Pairing System. In: Procedia Engineering, vol. 187, pp. 295–300 (2017)
3. Abdoos, M., Mozayani, N., Bazzan, A.: Traffic Light Control in Non-Stationary Environments Based on Multi Agent Q-Learning. In: Intelligent Transportation Systems (ITSC), 14th International IEEE Conference (2011)
4. Brys, T., Pham, T., Taylor, M.: Distributed Learning and Multi-Objectivity in Traffic Light Control. Connection Science 26:1, 65-83 (2014)
5. Gregurić, M., Vujić, M., Alexopoulos, C., Miletić, M.: Application of deep reinforcement learning in traffic signal control: An overview and impact of open traffic data. Applied Sciences, 10(11), 4011 (2020)
6. Rasheed, F., Yau, K.L.A., Noor, R.M., Wu, C., Low, Y.C.: Deep Reinforcement Learning for Traffic Signal Control: A Review. IEEE Access (2020)
7. Vidali, A., Crociani, L., Vizzari, G., Bandini, S.: A Deep Reinforcement Learning Approach to Adaptive Traffic Lights Management. In: WOA, pp. 42-50 (2019)

REFERENCES (CONT.)

8. Mnih, V., Badia, A., Mirza, M., Graves, A., Lillicrap, T., Harley, T., Silver, D., Kavukcuoglu, K.: Asynchronous Methods for Deep Reinforcement Learning. International Conference on Machine Learning, PMLR (2016)
9. Krajzewicz, D., Hertkorn, G., Rössel, C., Wagner, P.: SUMO (Simulation of Urban MObility)-an open-source traffic simulation. In: Proceedings of the 4th Middle East Symposium on Simulation and Modelling, MESM20002 pp. 183-187 (2002)
10. Alegre, L.: SUMO-RL. GitHub, github.com/LucasAlegre/sumo-rl (2019)
11. Wei, H., Zheng, G., Yao, H., Li, Z.: IntelliLight: A Reinforcement Learning Approach for Intelligent Traffic Light Control. In: KDD (2018)
12. Muthupalaniappan, A., Nair, B. S., Rajan, R. A., Krishnan, R. K.: Dynamic Control of Traffic Signals using Traffic Data from Google Maps and Road Cameras. In: International Journal of Recent Technology and Engineering (2019)
13. Zheng, G., Xiong, Y., Zang, X., Feng, J., Wei, H., Zhang, H., Li, Y., Xu, K., Li, Z.: Learning Phase Completion for Traffic Signal Control. In: CKIM, International Conference on Information and Knowledge Management (2019)
14. Hilmani, A., Maizate, A., Hassouni, L.: Automated Real-Time Intelligent Traffic Signal Control System for Smart Cities Using Wireless Sensor Networks. In: Wireless Communications and Mobile Computing (2020)
15. Pandit, K., Ghosal, D., Zhang, H. M., Chuah, C.: Adaptive Traffic Signal Control With Vehicular Ad hoc Networks. In: IEEE Transactions on Vehicular Technology, , vol. 62, no. 4, pp. 1459-1471 (2013)

THANK YOU