Adaptive Traffic Light System

CS659A: Autonomous Cyber-Physical Systems
5th May, 2021

Presenters :
Aakrati Jain (19111001)
Abir Mukherjee (19111005)
Vivek Agrawal (17807808)

Mentor:
Nirav

Previous works

- **Signal retiming** which involves optimizing traffic flow by gathering field data and minimizing the delay based cost function [3]. Optimizing traffic flow for fully-actuated intersections has been a challenge.

- **RL approaches** towards defining states as grid [4] and sensor based representations [5,6,7]. However, these works are limited by scalability.

- **Convolutional neural networks (CNN)** to capture state representations of the network via visual top-down images [8], discretized matrix representation [9,10,11,12] etc.

Deep Reinforcement Learning

- Deep reinforcement learning (DRL) is an extension of classical reinforcement learning with deep neural networks as function approximators.

- Value-based Reinforcement Learning method which is generally known as Q-learning.

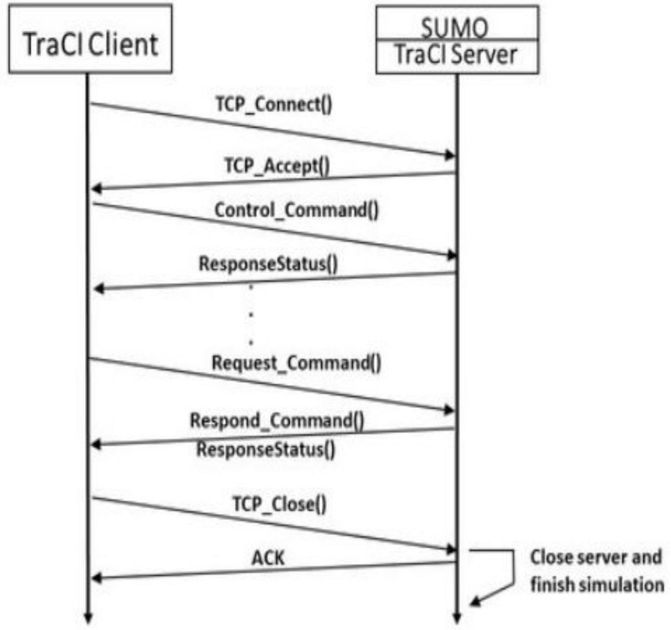
- Showing capabilities of Deep Neural Networks coupled with Reinforcement Learning.

- Main Paper: *Deep Reinforcement Learning for adaptive Traffic Signal Control* [1]
  - Reward function
  - Simulation tool, VISSIM

SUMO (Simulation of Urban MObility) Simulator

- SUMO is an open-source road simulation package that simulates traffic behavior, allows for road network construction, traffic light policy implementation, and traffic data collection.

- SUMO TraCI (Traffic Control Interface) extension which allows for dynamic control of the traffic lights at runtime like changing state of traffic lights while simulation.

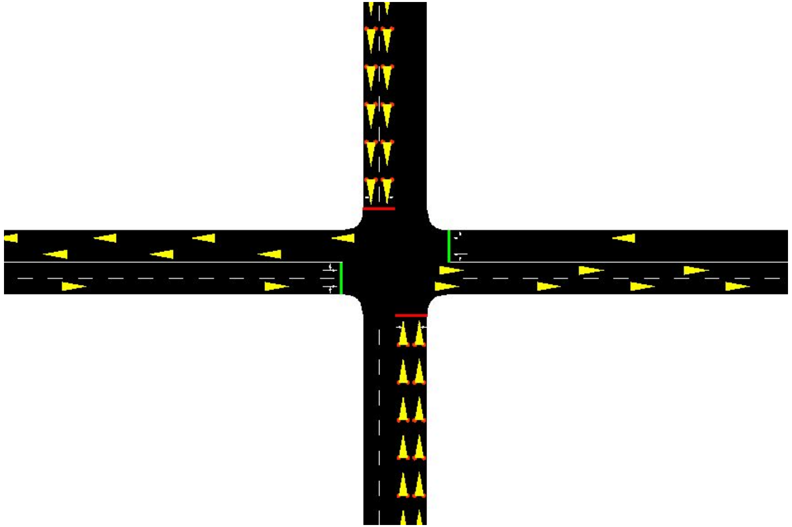


Pre-timed Traffic Light System

- Pre-timed traffic signal control uses a predefined set of red, yellow, and green time duration.

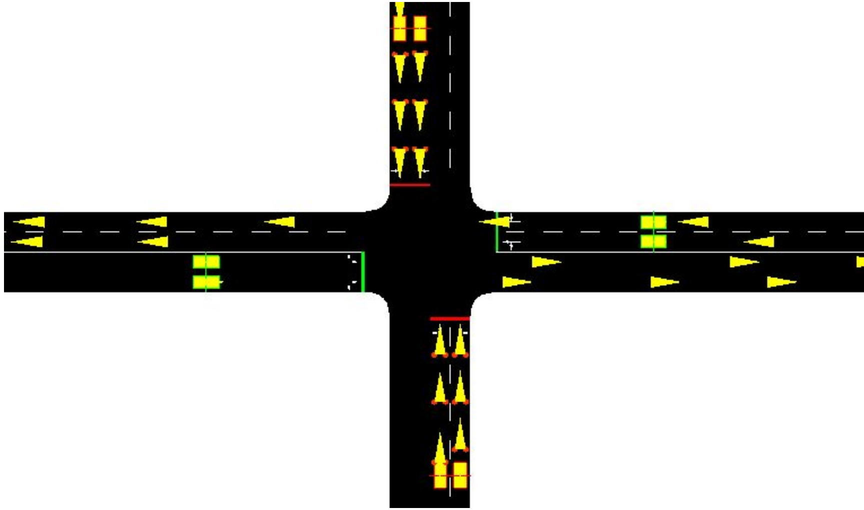
- Signals that adopt this traffic signal mode are cheaper as they do not require any kind of detection equipment near the intersection.

- This mode suffers from poor performance when the input volume towards the intersection fluctuates randomly.

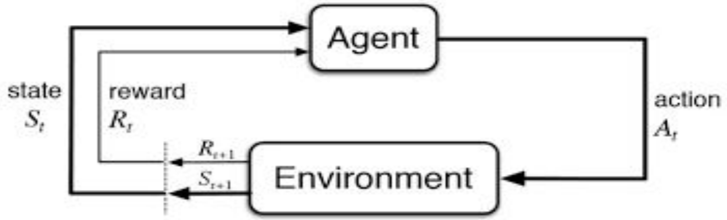


Actuated Traffic Light System

- Both directions of the intersection are equipped with traffic detection modules (induction loop detector placed at a distance).
- Green phase is variable. If vehicles continuously arriving at intersection then green phase is extended to maximum delay. Later when no vehicle passes through the induction loop detector, the green phase changes.
- This traffic signal mode is more effective than pre-timed traffic signal controls when both directions have high fluctuating volume throughout the day.



Deep RL based Traffic Light System



In reinforcement learning, an agent interacts with an environment with a given policy (that maps observable states to actions) and receives a reward signifying how well the agent performed.

A simple four-tuple describes a reinforcement learning model, < s_t, a_t, r_t, s_{t+1} >. This is an iterative interaction which goes on until T time steps, or until the agent arrives at a terminal state. In time sequence, the agent interacts with the environment as ... , s_t, a_t, r_t, s_{t+1}, a_{t+1} ...

The ultimate goal of the agent is to learn the optimal policy $\pi^*(s)$ that maximizes the state-action value function $Q^p(s,a)$ [11], which is defined as the expected sum of discounted future rewards.

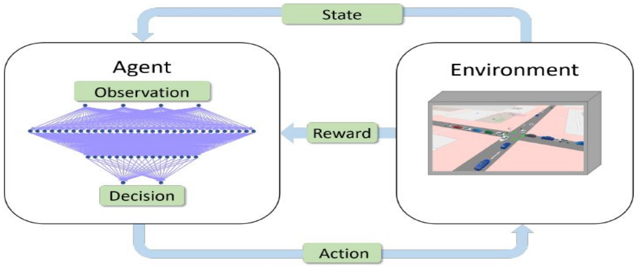
The optimal state-action function can be calculated by the Bellman equation:

$$Q^{\pi^*}(s_t, a_t) = \mathbb{E}_{s_{t+1}}\left[r_t + \gamma \max_{a_{t+1}} Q^{\pi^*}(s_{t+1}, a_{t+1}) | s_t, a_t\right]$$

where discount factor γ value ranges from [0,1).

Since we have apparently infinite states to work with, storing the Q-value for each and every state is infeasible. Hence we need an approximation technique for estimating the Q-value for the given state. We will be approximating the Bellman Equation by training the Deep Neural Network having the following structure.

Agent Design



- The design problem consists of three fundamental components of DRL: state, action and reward.
- STATE:
  - Upstream queue length, L for each lane
- ACTION:
  - Change current phase conditions
  - Keep current phase condition
- REWARD:

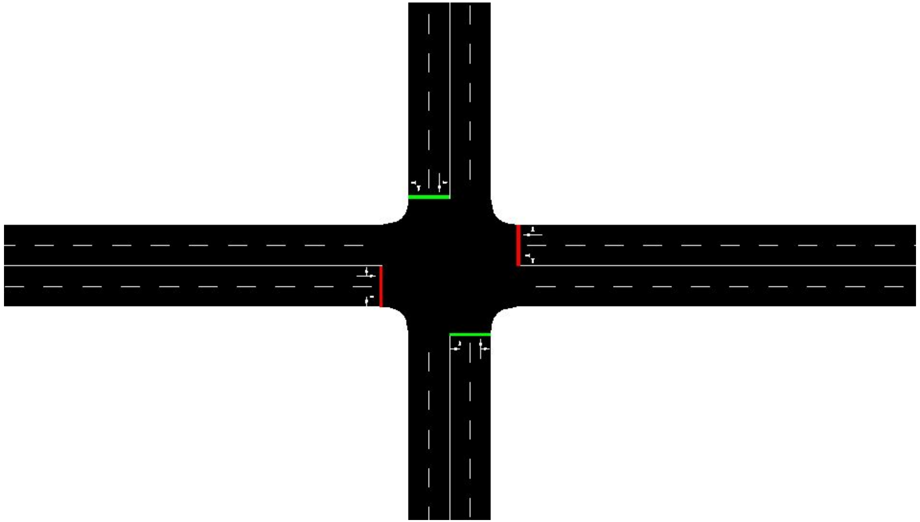
$$R = w_1 * \boldsymbol{\Sigma}\, IC + w_2 * \boldsymbol{\Sigma} ID$$

where IC = number of incoming vehicles within the detector distance that don't halt

ID = halting delay of incoming vehicles within the detector distance

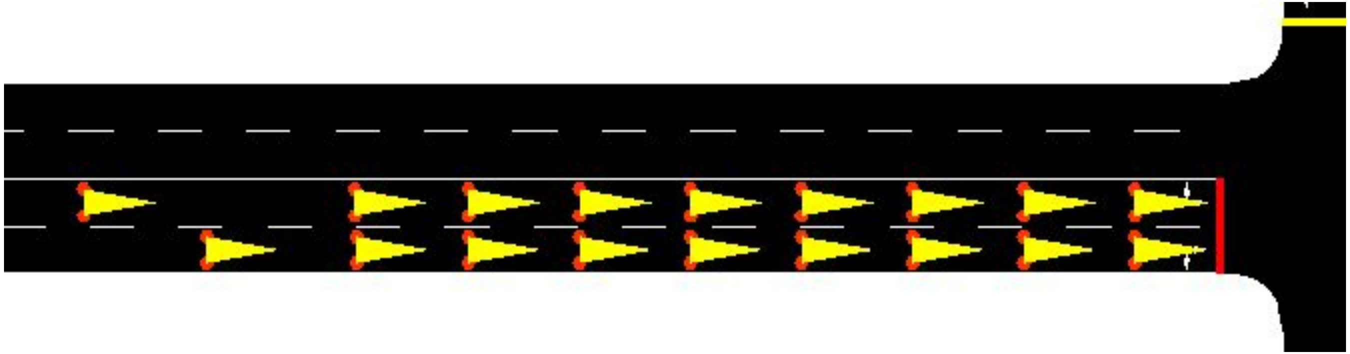
Intersection

- Modeled our problem as a 4-way intersection problem.
- All four directions, North, South, East, and West have four lanes each.
- Speed limit of each lane was selected to be 56 kmph (35 mph).
- TraCI was used to obtain traffic information of each lane similar to the actuated signal control.
- Minimum green time is selected to be 8 s.
- The length of the detector is 55 m.



Parameter: Queue Length

- It directly quantifies the total congestion for a particular intersection.
- Collected using TraCI.



Parameter: Max Out

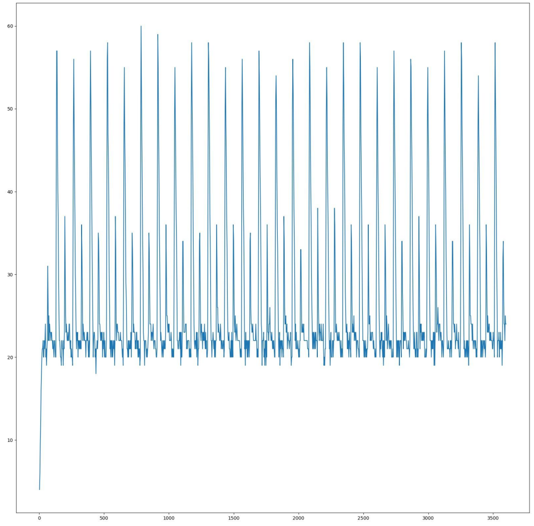
- Occurs for a given phase when the phase terminates due to reaching the designated maximum green time.
- The max out timer gets triggered when a vehicle is present in the opposing direction of traffic flow.
- To ensure that both the directions are served appropriately.
- Ensure that the cross traffic does not end up waiting too long.
- For both pre-timed and fully-actuated method, the max out was determined to be 50s for both directions of flow while we allowed our DRL agent to optimize max out duration based upon its policy.

Training Model

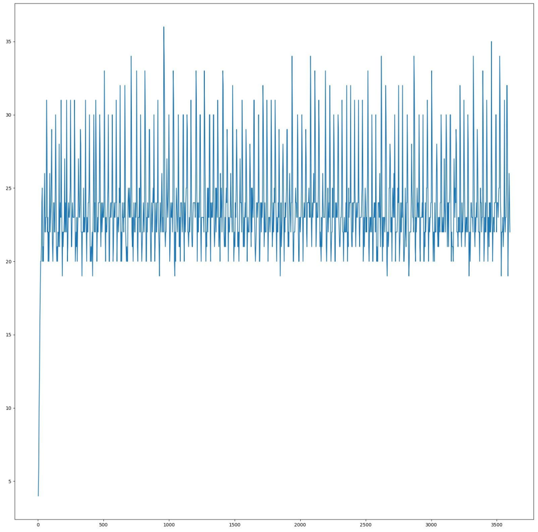
- 50 episodes, each episode simulating 15 mins of traffic simulation time
- Step duration in the environment is contingent upon which action the agent selects.
  - Keep the current signal phase then step duration is 3 simulation secs.
  - Change the current signal phase then step duration is 15 simulation secs.
- DRL agent is represented by a Multilayer Perceptron (MLP) with 2 fully connected layers, each layer consisting of 32 and 16 neurons respectively. Each fully connected layer is followed by a rectifier linear unit (ReLU) activation. The final outputs of the network are 2 possible Q-value representing 2 actions the agent can choose from.

Results (WE-EW-NS-SN: 500vh/hr, 500vh/hr, 1200vh/hr, 1200vh/hr) (NS heavy load)

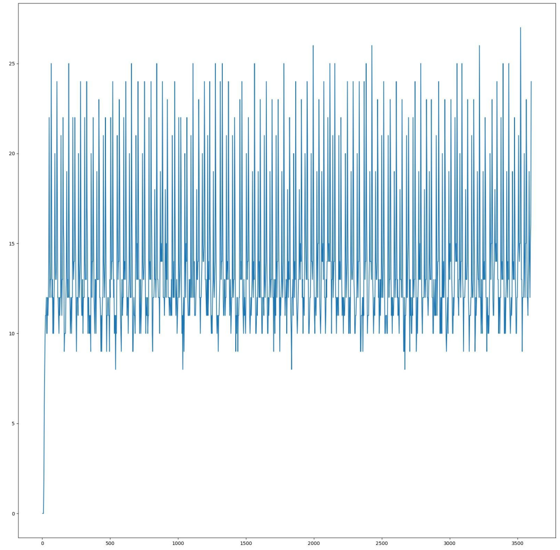
Pre-timed



Actuated

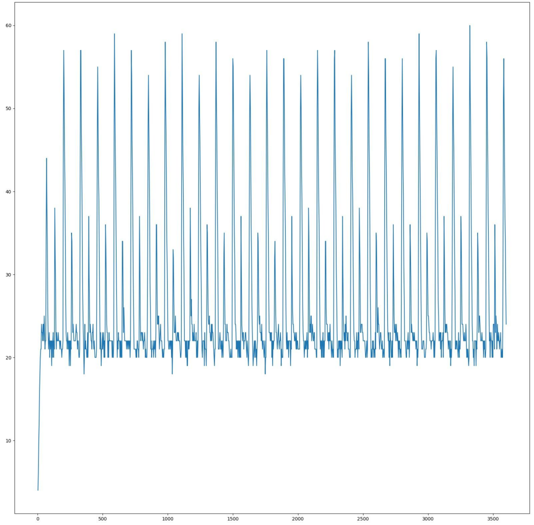


DQN

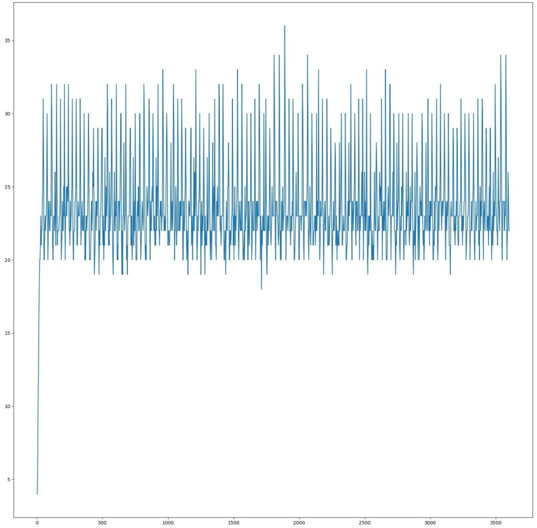


Results (WE-EW-NS-SN: 1200vh/hr, 1200vh/hr, 500vh/hr, 500vh/hr) (EW heavy load)

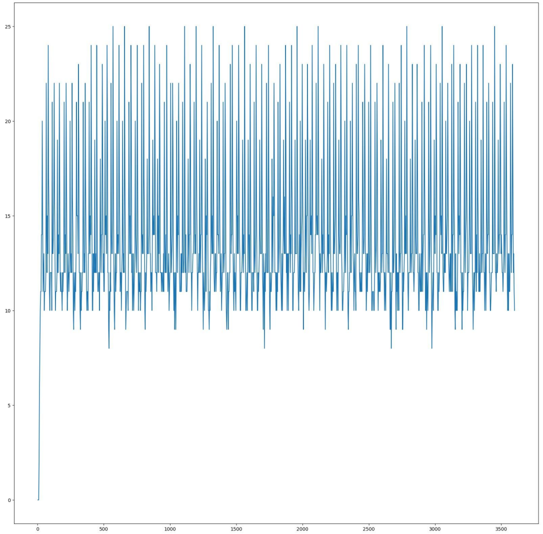
Pre-timed



Actuated

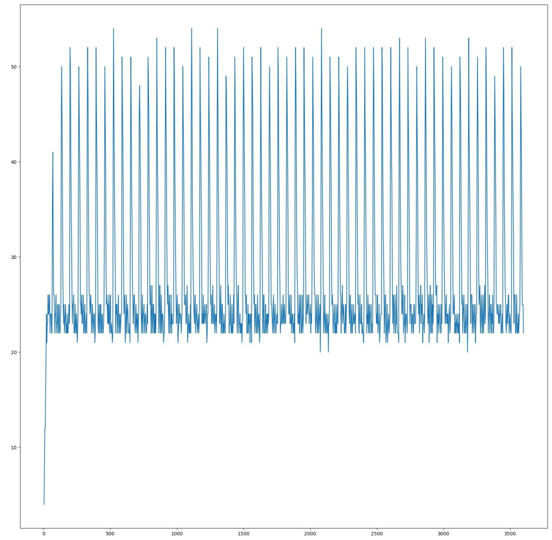


DQN

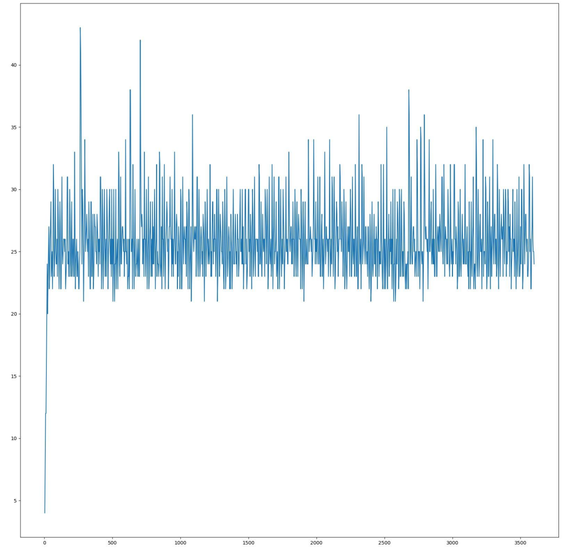


Results (WE-EW-NS-SN: 900vh/hr, 900vh/hr, 900vh/hr, 900vh/hr) (Moderate-Equal Load)

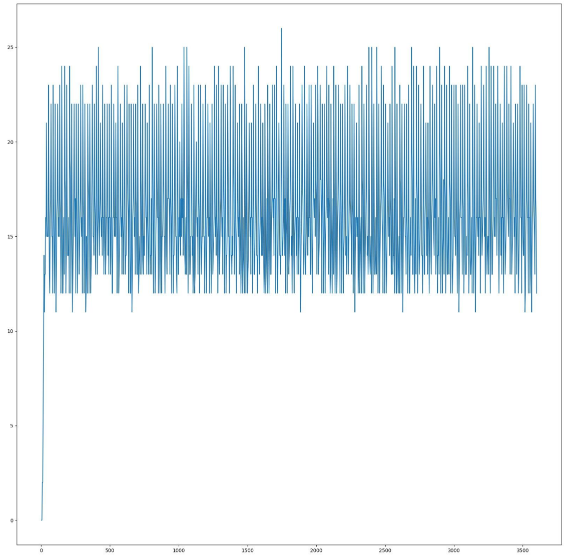
Pre-timed



Actuated



DQN



Conclusion

- Pre-timed signal control is usually used during nominal traffic flow, where signal timings are optimized for numerous iterations to ensure phase timings are sufficient to cope with normal traffic demands.

- Actuated signal control is used when the inverse situation from the pre-timed controller is predicted to happen such as a special event or a highly accident-prone location.

- The DQN algorithm showing increased efficiency over the pre-timed and actuated paradigm. Q-Learning has proven itself as a viable alternative to traditional traffic control policies making it possible to reduce traffic congestion on roads around the world.

Future Work

- Train the model on larger number of episodes for longer episode durations, to increase model robustness
- Extend our DRL framework towards intersections with left and right turns and arterial corridors.
- Ability to handle sudden changes in traffic patterns and anomalous scenarios such as accidents, construction, and other events.
- Multiple intersections coordinating to have an effective adaptive traffic light system.
- Testing our DRL agent's performance on real traffic data during high traffic demand periods.

References

[1] Tan, Kai Liang, et al. "Deep reinforcement learning for adaptive traffic signal control." *Dynamic Systems and Control Conference*. Vol. 59162. American Society of Mechanical Engineers, 2019.

[2] Wang, Yizhe, et al. "A review of the self-adaptive traffic signal control system based on future traffic environment." *Journal of Advanced Transportation* 2018 (2018).

[3] Gordon, R.L., 2010. *Traffic signal retiming practices in the United States* (Vol. 409). Transportation Research Board.

[4] Wiering, M.A., 2000. Multi-agent reinforcement learning for traffic light control. In *Machine Learning: Proceedings of the Seventeenth International Conference (ICML'2000)* (pp. 1151-1158).

[5] Medina, J.C., Hajbabaie, A. and Benekohal, R.F., 2010, September. Arterial traffic control using reinforcement learning agents and information from adjacent intersections in the state and reward structure. In *13th International IEEE Conference on Intelligent Transportation Systems* (pp. 525-530). IEEE.

[6] Arel, I., Liu, C., Urbanik, T. and Kohls, A.G., 2010. Reinforcement learning-based multi-agent system for network traffic signal control. *IET Intelligent Transport Systems*, *4*(2), pp.128-135.

[7] El-Tantawy, S., Abdulhai, B. and Abdelgawad, H., 2013. Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): methodology and large-scale application on downtown Toronto. *IEEE Transactions on Intelligent Transportation Systems*, *14*(3), pp.1140-1150.

[8] Wei, H., Zheng, G., Yao, H. and Li, Z., 2018, July. Intellilight: A reinforcement learning approach for intelligent traffic light control. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2496-2505).

[9] Gao, J., Shen, Y., Liu, J., Ito, M. and Shiratori, N., 2017. Adaptive traffic signal control: Deep reinforcement learning algorithm with experience replay and target network. *arXiv preprint arXiv:1705.02755*.

[10] Genders, W. and Razavi, S., 2016. Using a deep reinforcement learning agent for traffic signal control. *arXiv preprint arXiv:1611.01142*.

[11] Muresan, M., Fu, L. and Pan, G., 2019. Adaptive traffic signal control with deep reinforcement learning an exploratory investigation. *arXiv preprint arXiv:1901.00960*.

[12] Liu, M., Deng, J., Xu, M., Zhang, X. and Wang, W., 2017, August. Cooperative deep reinforcement learning for tra ic signal control. In *Proc. 23rd ACM SIGKDD Conf. Knowl. Discovery Data Mining (KDD)*.

Thank You.

Percentage contribution by each member

1. Aakrati Jain
2. Abir Mukherjee
3. Vivek Agrawal