PressLight: Learning Max Pressure Control to Coordinate Traffic Signals in Arterial Network

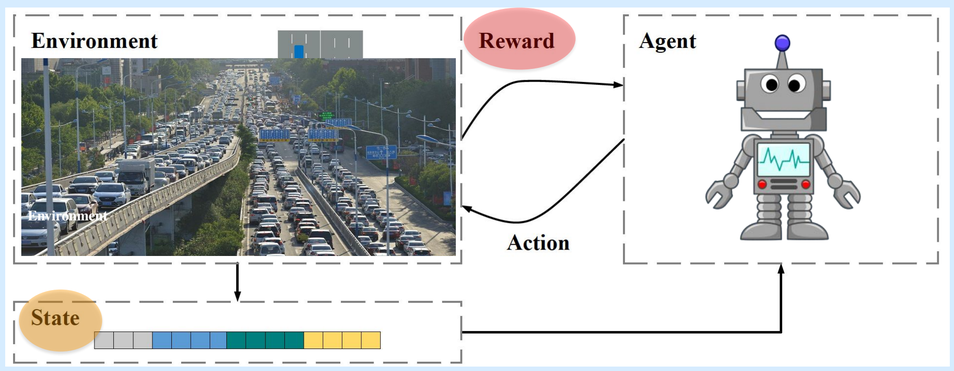
Hua Wei, Chacha Chen, Guanjie Zheng, Kan Wu, Vikash Gayah, Kai Xu, Zhenhui (Jessie) Li

Pennsylvania State University, Shanghai Jiao Tong University, Tianrang Inc.

{hzw77, gjz5038, jessieli}@ist.psu.edu, chacha1997@sjtu.edu.cn, kxw5389@psu.edu, gayah@engr.psu.edu, kai.xu@tianrang-inc.com

1. Introduction

Reinforcement Learning Framework for Traffic Signal Control



Goal: Minimize the travel time of all vehicles

Hard to optimize directly

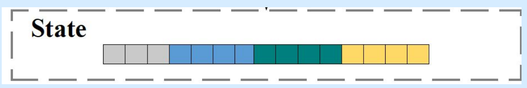
2. Two Key Questions in Reinforcement Learning (RL)

1. How to define **reward**?

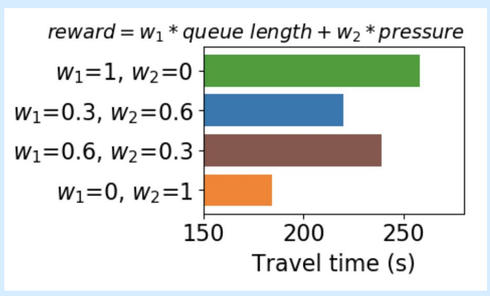
$$\text{reward} = w_1 * \text{queue length} + w_2 * \text{vehicle speed} + w_3 * \text{delay} + w_4 * \text{waiting time} + \ldots$$

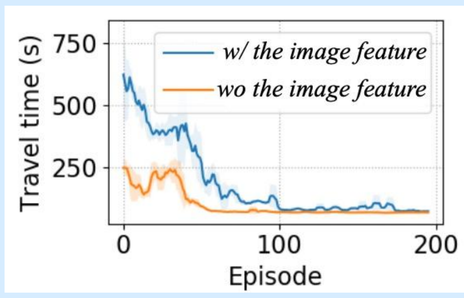
1	Queue length
2	Speed
3	Delay
4	Waiting time
5	Number of stops
6	Frequency of signal change
7	Accident avoidance
...	...

2. How much information is enough in **state**?



1	Queue length
2	Waiting time
3	Traffic volume
4	Speed
5	Phase duration
6	Position of vehicles
7	Phase
...	...

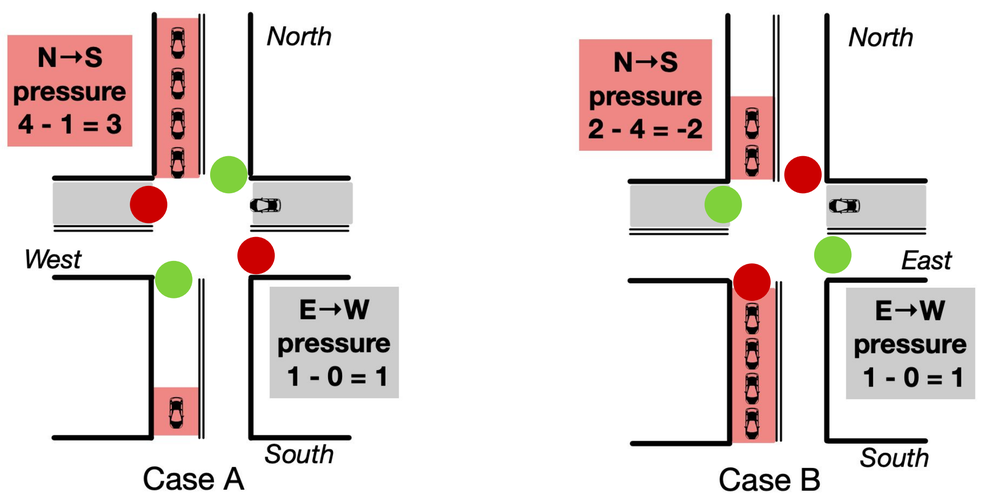




3. Support from Transportation Field: MaxPressure (MP)

Key Idea of MP

Goal: minimize the "pressure" of an intersection



Nice Properties of MP

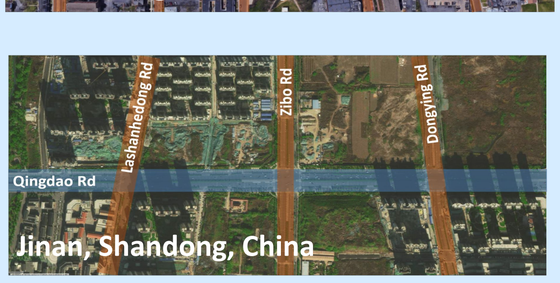
1. MP is proven to **maximize the system throughput** by minimizing the pressure of each intersections *under the assumption of no physical queue*.
2. MP derives **evolution equations** to formulate the state transition of the traffic as **a Markov chain**

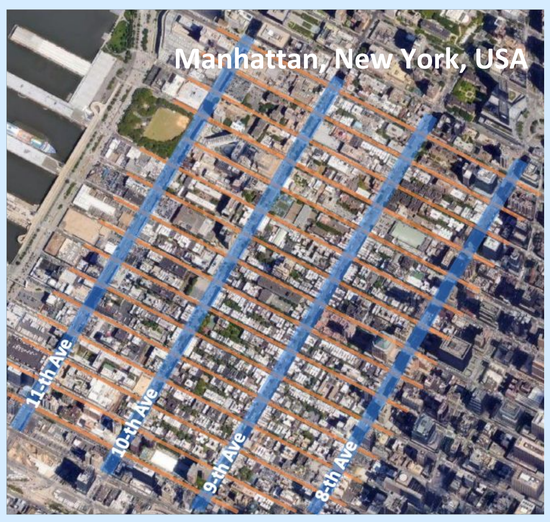
4. Method: Connecting RL with MP

	RL		MP
Objective	Maximize long-term reward	Define reward as derivation of pressure	Minimize the pressure
Solution	Trial-and-error search	Find optimal policy with Epsilon-Greedy	Greedy algorithm
System Dynamics	State transitions as Markov Decision Process	Include variables from the equation of MP into the state of RL	Formulation of evolution equations as a Markov chain

Reward	pressure of each intersection
State	current phase; number of vehicles on incoming lanes; number of vehicles on the segments of incoming lanes

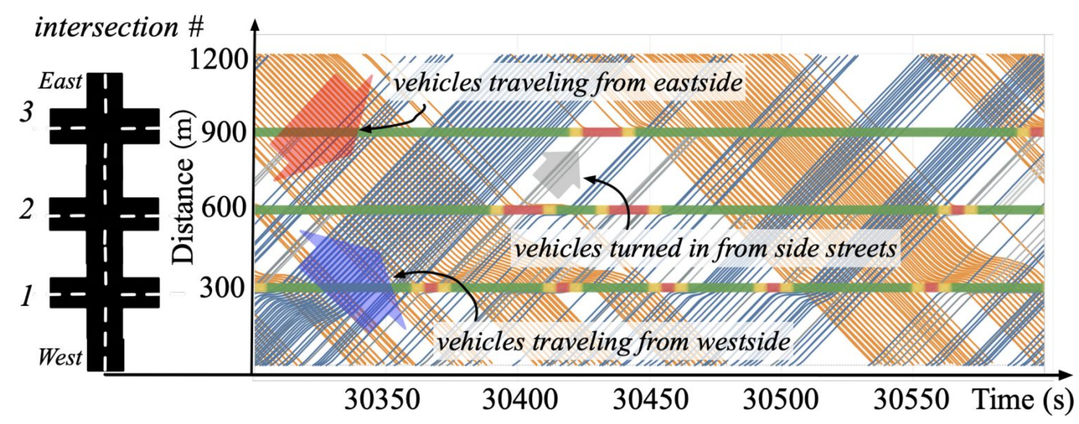
5. Experiments





	Real-world traffic					
	Qingdao Rd., Jinan	Beaver Ave., State College	8th Ave., NYC	9th Ave., NYC	10th Ave., NYC	11th Ave., NYC
FixedTime	317.40	336.29	432.60	469.54	347.05	368.84
GreenWave	370.30	332.06	451.98	502.30	317.02	314.08
MaxPressure	567.06	222.90	412.58	370.61	392.77	224.54
GRL	238.19	455.42	704.98	669.69	676.19	548.34
LIT	58.18	338.52	471.30	726.04	309.95	340.40
PressLight	**54.87**	**92.00**	**223.36**	**149.01**	**161.21**	**140.82**

Learnt coordination policy



Demos

MP vs. PressLight

GreenWave by RL

References

[1] Zheng, Guanjie, et al. "Diagnosing Reinforcement Learning for Traffic Signal Control." *arXiv preprint arXiv:1905.04716*(2019).

[2] Varaiya, Pravin. "The max-pressure controller for arbitrary networks of signalized intersections." Advances in Dynamic Network Modeling in Complex Transportation Systems. Springer, New York, NY, 2013. 27-66.

[3] Wei, Hua, et al. "A Survey on Traffic Signal Control Methods." arXiv preprint arXiv:1904.08117 (2019).

The work was supported in part by NSF awards #1652525, #1618448, and #1639150. The views and conclusions contained in this paper are those of the authors and should not be interpreted as representing any funding agencies.