東南大學
SOUTHEAST UNIVERSITY

# MonitorLight: Reinforcement Learning-based Traffic Signal Control Using Mixed Pressure Monitoring

*Zekuan Fang, Fan Zhang, et al.*

*CIKM 2022*

# 提纲

研究背景

算法设计

实验

# 提纲

## 研究背景

- 背景介绍
- 研究现状

# 背景介绍

## 交通拥堵问题

- 城市交通拥堵情况越来越严重。交通拥堵会导致经济、环境、生活上的诸多问题。
- 如何缓解交通拥堵？

# 背景介绍

## 交通信号控制 Traffic Signal Control (TSC)

合理的交通信号控制策略可以最大化交通流，有效缓解道路拥堵。

# 研究现状

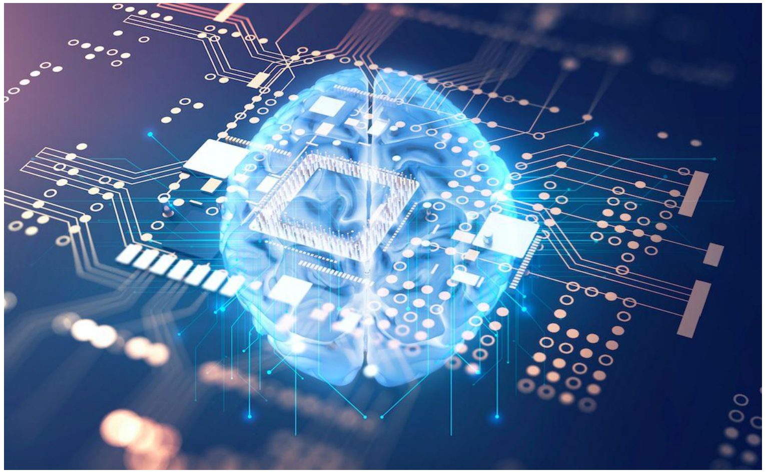
## Conventional TSC methods

【1】Seung-Bae Cools et al. . **Self-organizing traffic lights: A realistic simulation.**. Advances in applied self-organizing systems, 2013.

【2】Peter Koonce and Lee Rodegerdts. **Traffic signal timing manual**. Technical report, 2008.

- 在交通领域经典算法的基础上，设定固定的总周期和信号变化顺序，优化各相位的绿灯持续时间和交叉口之间的相位偏移。
- 不适用于复杂动态的交通环境。

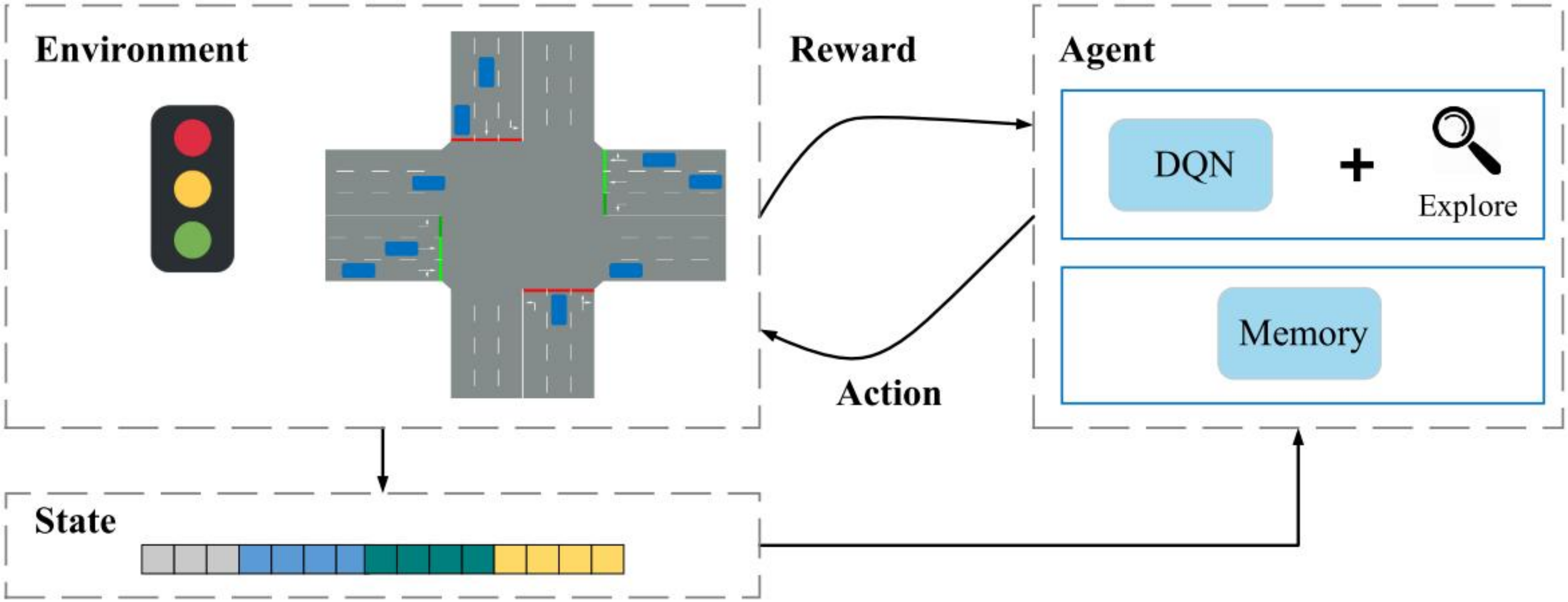
# 背景介绍

- Road Side Units (RSU)
- Artificial Intelligence (AI)
- Traffic-oriented Cyber-Physical Systems(CPS)

# 背景介绍

## RL-Based Traffic Signal Control (TSC)：




**Transition：** $< S , A , S' >$

# 研究现状

## RL-based TSC methods

【1】Hua Wei et al. **Colight: Learning network-level cooperation for traffic signal control.** CIKM 2019.

【2】 Chang Liu et al. **Generalight: Improving environment generalization of traffic signal control via meta reinforcement learning.** CIKM 2020.

【3】Yutong Ye et al. **Fedlight: Federated reinforcement learning for autonomous multi-intersection traffic signal control**. DAC 2021.

【4】Hua Wei et al. **Presslight: Learning max pressure control to coordinate traffic signals in arterial network**. SIGKDD 2019.

- 在复杂交通流条件下，与环境交互来自主优化信号策略。
- 只考虑相位的选择，忽略了相位的持续时间。可能会降低样本效率，导致收敛缓慢。

# 研究现状

## Dynamic duration RL methods

【1】Xiaorong Hu et al. **A traffic light dynamic control algorithm with deep reinforcement learning based on gnn prediction**. CoRR abs/2009.14627 (2020).

【2】Wupan Zhao et al. **IPDAlight: Intensity-and phase duration-aware traffic signal control based on reinforcement learning.** Journal of Systems Architecture 2022.

- 现有关于动态相位持续时间的RL方法，只分析车道上的车辆数量来计算相位的持续时间。

# 提纲

## 算法设计

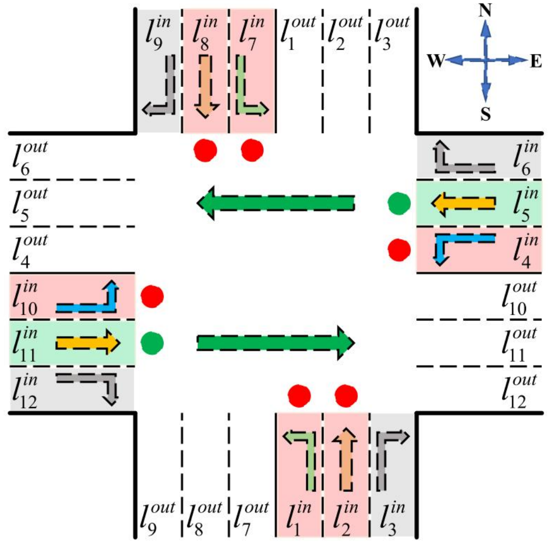
Preliminaries

Mixed pressure

Monitoring attribute

# Preliminaries


**Road Network**




**Intersection**
**(Signals、Lanes)**

| Phase | Movement | Action Lane | Non-action Lane |
|---|---|---|---|
| SN-straight | | $l_2$, $l_8$ | $l_1$, $l_4$, $l_5$, $l_7$, $l_{10}$, $l_{11}$ |
| WE-straight | | $l_5$, $l_{11}$ | $l_1$, $l_2$, $l_4$, $l_7$, $l_8$, $l_{10}$ |
| SN-left | | $l_1$, $l_7$ | $l_2$, $l_4$, $l_5$, $l_8$, $l_{10}$, $l_{11}$ |
| WE-left | | $l_4$, $l_{10}$ | $l_1$, $l_2$, $l_5$, $l_7$, $l_8$, $l_{11}$ |

**Traffic Movement**

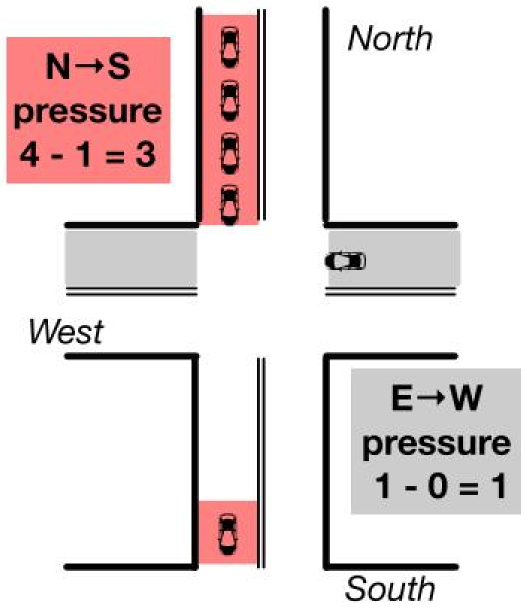
# Pressure

**Pressure of Lane[1]：**

车道上的车辆数

**Pressure of Movement：**

出车道与入车道上的车辆数差




**Mixed pressure**

**Static Pressure：** $P_s = \Sigma_{veh \in V_i^s}(1 + \omega \times t_s)$

**Dynamic Pressure：** $P_d = \Sigma_{veh \in V_i^d} \frac{1}{v \times \frac{L}{L_{max}} + 1}$

[1] Presslight: Learning max pressure control to coordinate traffic signals in arterial network. SIGKDD 2019.

# Agent设计

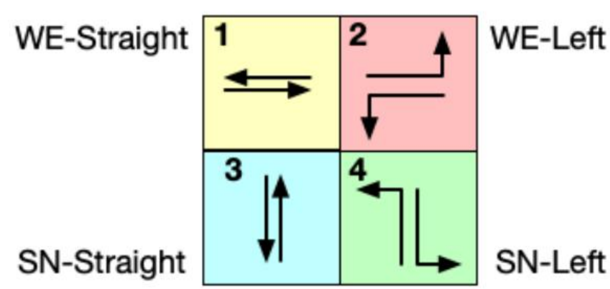
| Element | Design |
| --- | --- |
| **State** | Mixed pressure for each lane： $S_j = (P_m^{l_1^i}, P_m^{l_2^i}, \ldots, P_m^{l_{12}^i}, -P_m^{l_1^o}, -P_m^{l_2^o}, \ldots, -P_m^{l_{12}^o})$ |
| **Action** | Choose one of the four Phases |
| **Reward** | $R_j = -\sum_{l^i \in L^{in}} P_m^{l^i}$，进入车道混合压力值和的负数 |

$P_m$： 车道的混合压力

$L_{in}$： 进入交叉口的车道

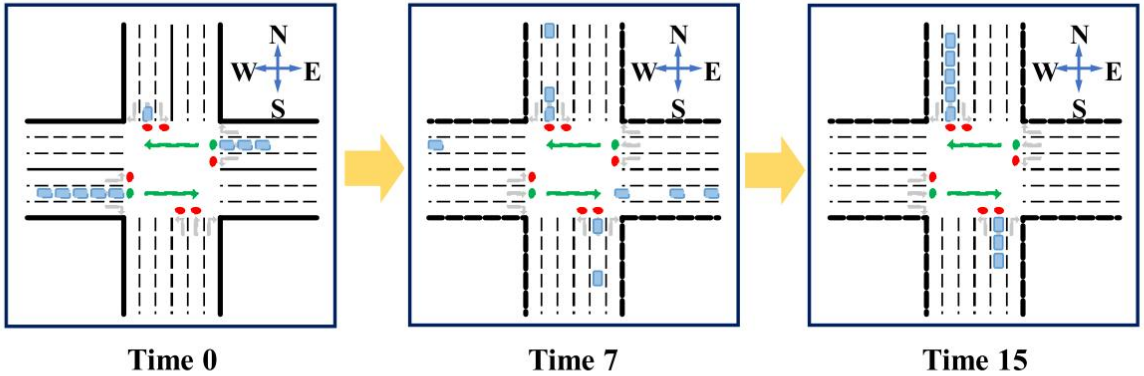
$L_{out}$： 离开交叉口的车道

Four **phases**

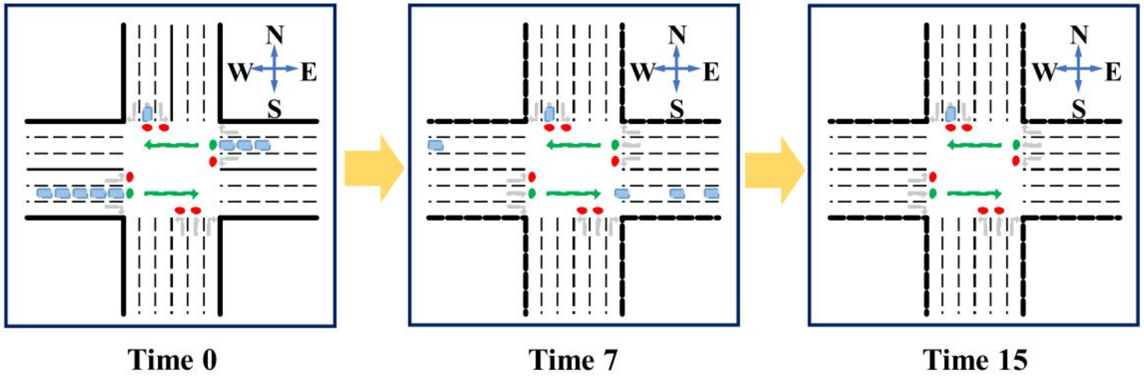


# Monitoring attribute

## 相位时间固定存在的问题?




(a) A WE-straight phase with vehicles entering the intersection.




(b) A WE-straight phase with no vehicle entering the intersection.

# Monitoring attribute

## 相位持续时间选择

### Monitoring Attribute：

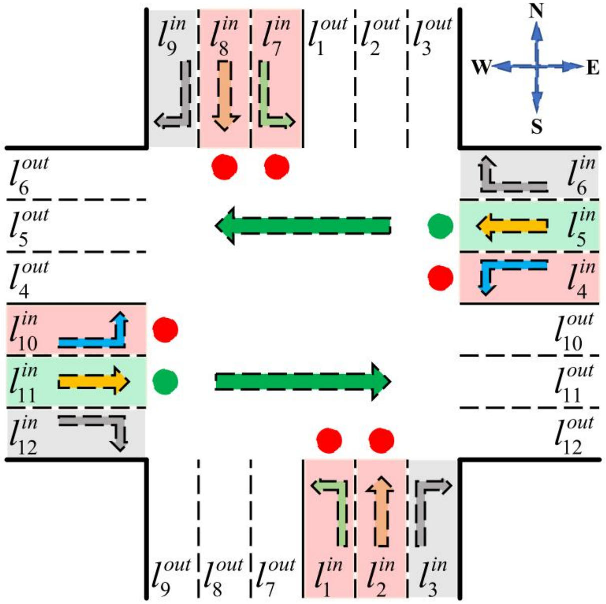
$$Monitoring_j = \frac{\sum_{l_i \in l_{act}} P_d}{max_{l_i \in l_{nac}} P_s + \epsilon}$$

$P_s$： 车道的静压

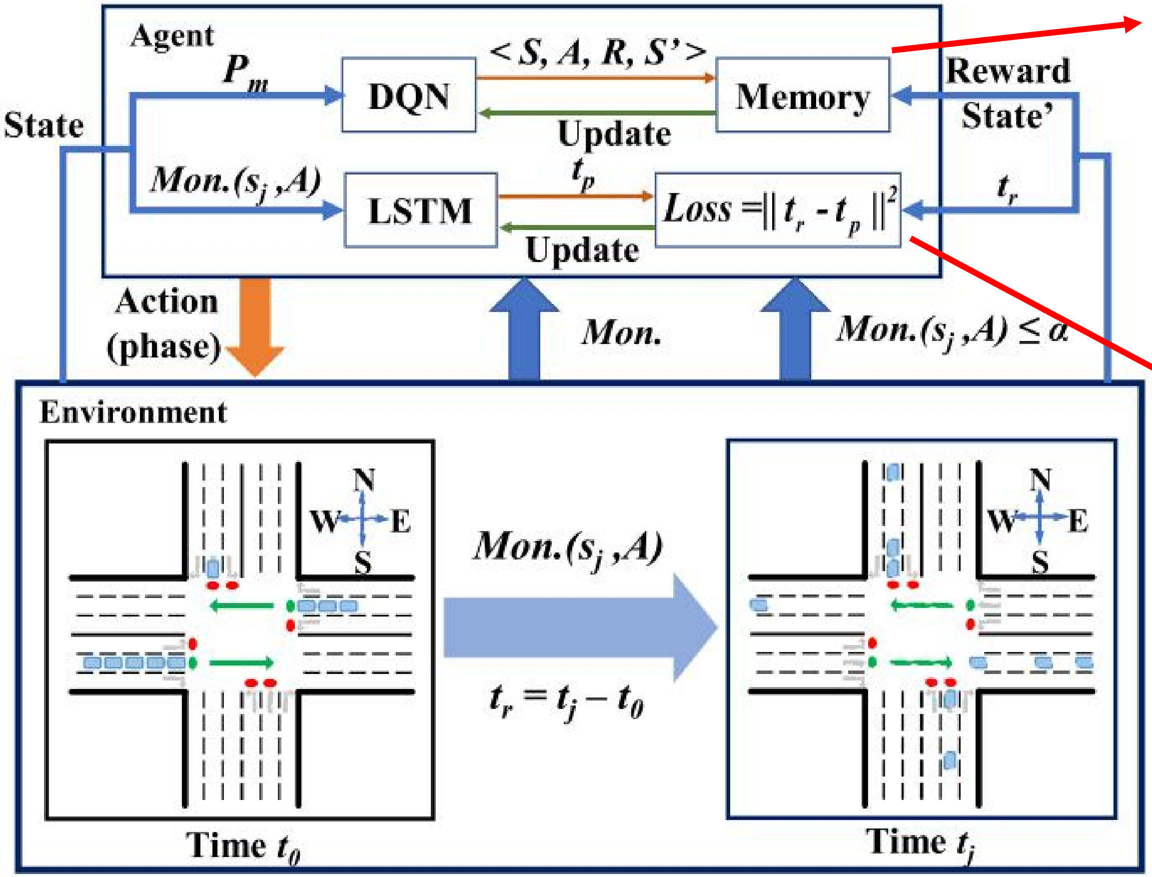
$P_d$： 车道的动压

$l_{act}$： 动作车道，允许车辆通行的驶入车道(右转车道除外)。

$l_{nac}$： 非行动车道，禁止车辆通行的车道。



# Training Framework




相位选择模型：

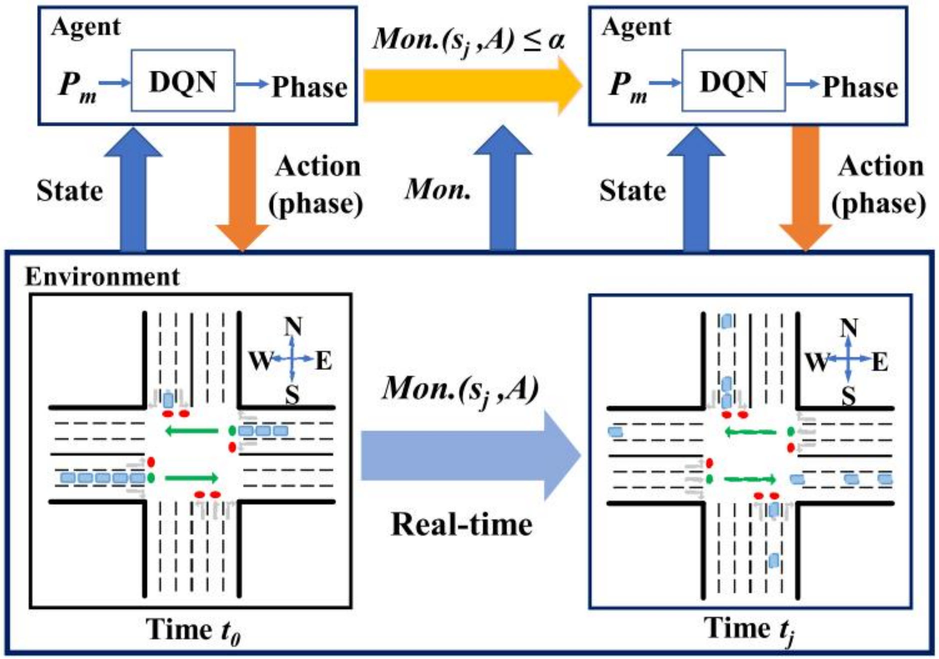
DQN；

experience replay

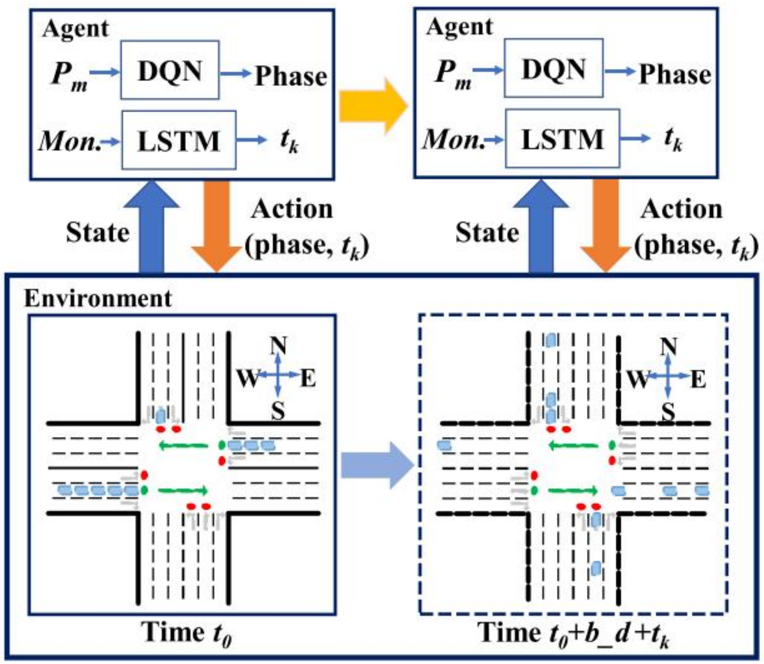
持续时间预测模型：

LSTM

# Test Framework




MonitorLight-r




MonitorLight-p

- MonitorLight-r： 相位切换基于实时Monitoring属性。
- MonitorLight-p： 无Monitoring属性，基于时间预测模型确定相位时间。

# 提纲

实验

实验设计

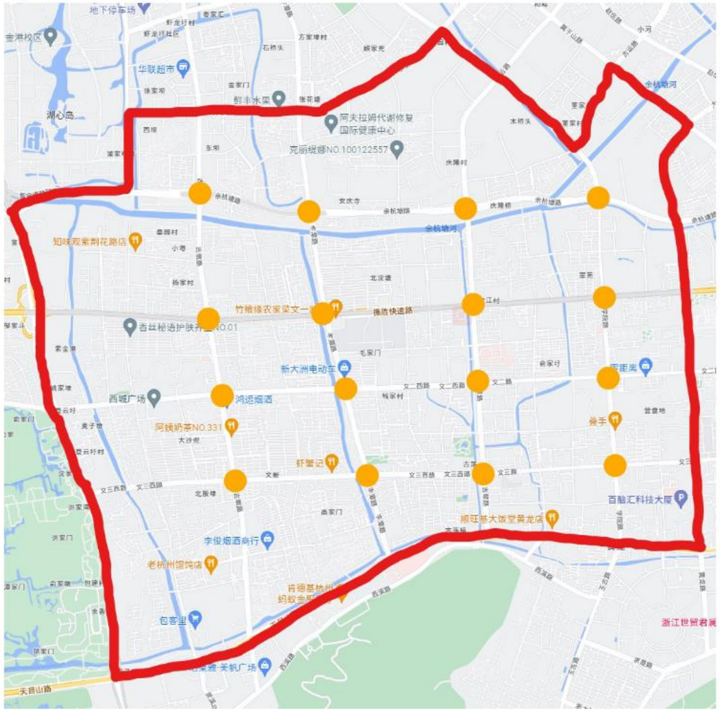
实验结果

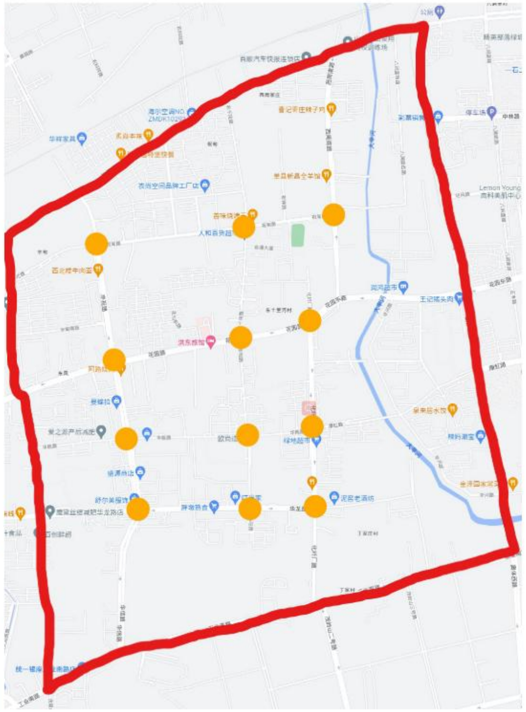
总结

# 实验设计

**交通模拟器：**

CityFlow





**数据集：**

- 杭州 4x4 network，16 intersections
- 济南 3x4 network，12 intersections
- 四个合成数据集 Syn_1x3，Syn_2x2，Syn_3x3，Syn_4x4

# 实验设计

**对比算法:**

1）传统TSC方法:

Fixedtime

SOTL

2）强化学习方法:

GRL

CoLight

PressLight

3）动态相位时间的RL方法:

IPDALight

**评价指标:**

- 平均行驶时间（ATT）：所有车辆从进入网络到离开网络的平均时间
- 收敛性能（SEC）：开始收敛时的Episode轮数

# 实验结果

| Type | | Method | Synthetic Datasets | | | | Real-World Datasets | |
|---|---|---|---|---|---|---|---|---|
| | | | Syn_1×3 | Syn_2×2 | Syn_3×3 | Syn_4×4 | Jinan | Hangzhou |
| Traditional | | Fixed-Time | 384.47 | 454.01 | 508.87 | 565.99 | 405.91 | 488.51 |
| | | SOTL | 247.07 | 331.64 | 424.66 | 474.32 | 410.65 | 505.53 |
| RL | Fixed | GRL | 208.21 | 239.13 | 431.43 | 523.01 | 562.91 | 598.17 |
| | | CoLight | 210.01 | 312.29 | 328.70 | 397.07 | 327.65 | 337.45 |
| | | PressLight | 98.74 | 123.90 | 166.28 | 215.32 | 285.65 | 341.99 |
| | Dynamic | IPDALight | 88.01 | 109.66 | 146.92 | 184.54 | 255.35 | 298.99 |
| Our | | MonitorLight-r | **84.03** | **103.56** | **136.90** | **173.56** | **246.21** | **292.71** |
| | | MonitorLight-p | 87.64 | 109.16 | 143.44 | 183.52 | 251.33 | 296.79 |
| Improvement over IPDALight | | | **4.52%** | **5.56%** | **6.82%** | **5.95%** | **3.58%** | **2.10%** |

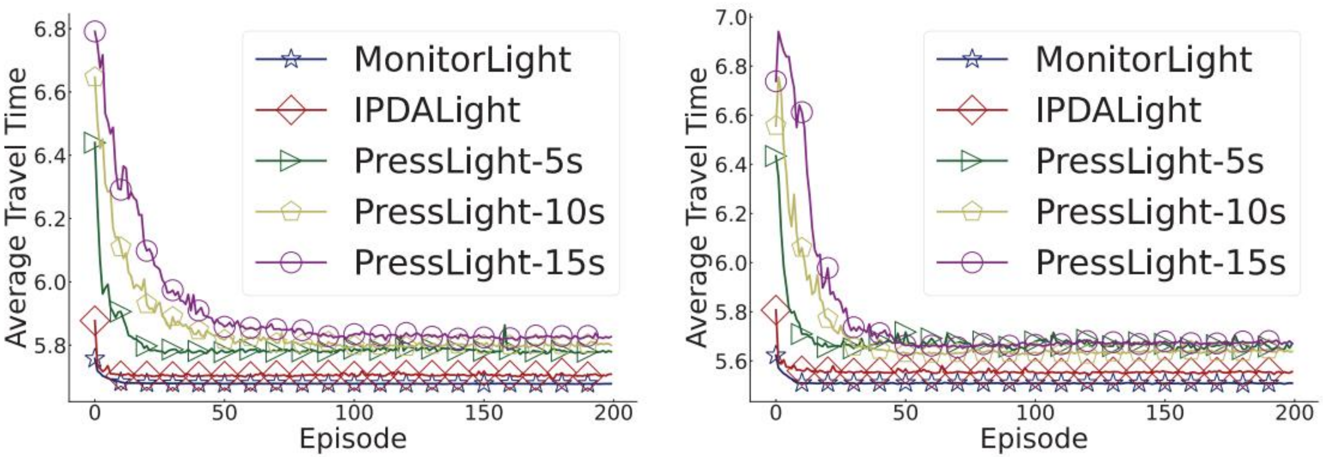
Performance comparison

- 基于RL的方法优于传统TSC方法
- MonitorLight-r 和 MonitorLight-p 的表现都优于其他Baseline

# 实验结果

| Method | Syn_2×2 | | | Syn_4×4 | | | Hangzhou | | |
|---|---|---|---|---|---|---|---|---|---|
| | SEC | JP (s) | APD (s) | SEC | JP (s) | APD (s) | SEC | JP (s) | APD (s) |
| **PressLight-5s** | 28 | 542.75 | 5 | 66 | 842.84 | 5 | 21 | 626.31 | 5 |
| **PressLight-10s** | 56 | 707.62 | 10 | 46 | 1068.62 | 10 | 54 | 770.15 | 10 |
| **PressLight-15s** | 58 | 837.35 | 15 | 71 | 1084.22 | 15 | 81 | 890.19 | 15 |
| **IPDALight** | 12 | 279.44 | 14.96 | 26 | 507.38 | 13.21 | 7 | 356.90 | 7.09 |
| **MonitorLight** | **10** | **160.48** | 10.99 | **9** | **254.30** | 11.09 | **5** | **316.58** | 14.35 |
| **Improvement** | **16.67%** | **42.57%** | - | **65.38%** | **49.88%** | - | **28.57%** | **11.30%** | - |

**Learning Convergence Rate Comparison**




(a) Hangzhou (b) Jinan

# 总结展望

## 总结

- 基于Mixed pressure的新Agent定义
- 对于相位持续时间的研究

# 请指正！

Thank you！