Expression is enough: Improving traffic signal control with advanced traffic state representation

Liang Zhang1, Qiang Wu2, Jun Shen3, Linyuan Lui 2, Jianqing Wu2, Bo Du4

1 School of Life Sciences, Lanzhou University, Lanzhou 730000, China
2 Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China
3 School of Computing and Information Technology, University of Wollongong, Wollongong, Australia
4 SMART Infrastructure Facility, University of Wollongong, Wollongong, Australia

Abstract
Recently, finding fundamental properties for traffic state representation is more critical than complex algorithms for traffic signal control (TSC). In this paper, we (1) present a novel, flexible and straightforward method advanced max pressure (Advanced-MP), taking both running and queueing vehicles into consideration to decide whether to change current phase; (2) novelty design the traffic movement representation with the efficient pressure and effective running vehicles from Advanced-MP, namely advanced traffic state (ATS); (3) develop an RL-based algorithm template Advanced-XLight, by combining ATS with current RL approaches and generate two RL algorithms, “Advanced-MPLight” and “Advanced-CoLight”. Comprehensive experiments on multiple real-world datasets show that: (1) the Advanced-MP outperforms baseline methods, which is efficient and reliable for deployment; (2) Advanced-MPLight and Advanced-CoLight could achieve new state-of-the-art. Our code is released on Github.

Keywords: traffic signal control, reinforcement learning, advanced max pressure, advanced traffic state, Advanced-XLight.

1 Introduction
Traffic signal control (TSC) is essential for improving transportation efficiency and mitigating traffic congestion. Now, in many modern cities, FixedTime[12], SCOOT[11], and SCATS[16] are still the most common TSC systems, heavily relying on expert-designed traffic signal plans.

Max pressure (MP) control[19] is an optimization-based approach, which try to maximize the global throughput by balancing the queue length in the traffic network. The SOTL[3] is adaptive and flexible, using "request" to decide whether to change the current phase. These conventional methods can’t be easily adapted to complex dynamic traffic flows without expert prior knowledge. There is an urgent need to apply new approaches to improve transportation efficiency and mitigate traffic congestion.

Recently, reinforcement learning (RL) has drawn increasing interest, and some researchers start introducing RL techniques into TSC. The RL-based TSC methods[26][21] can directly learn from the complex conditions through trial-and-error, without requiring assumptions about the traffic model. Some RL-based methods have shown superior performance over traditional methods. In addition to the outstanding performance, some RL-based TSC approaches can be used to control large-scale traffic signals[2][21].

However, recent RL-based methods have a long training process, and their performance even can’t outperform traditional approach max pressure (MP) control[19] in a proper training settings[24]. Some complex RL-based approaches may not be the optimal solutions for TSC.

Both state design and neural network design play an essential role in RL-based methods for TSC. GCN[17] uses a graph convolutional neural networks to represent geometric features of multi-intersection. CoLight[21] uses a graph attention network to realize intersection cooperation for TSC. PressLight[20] and MPLight[2] introduce "pressure" into state and reward design and get a better performance. The success of MPLight[2] and Efficient-CoLight[24] has shown that combining and optimizing traffic state representation from traditional methods with RL-based models can bring significant improvements. The traffic movement pressure expressed by EP is an essential representation of traffic state. Although EP has strong representativeness, it ignores the running vehicles in the traffic network. Here, the "pressure" alone is not enough to represent the complex traffic state. In addition, with a greedy manner, Efficient-MP[24] is not flexible.

The limitations of current traffic state representation and TSC methods as flow:

- latest traditional TSC methods with a greedy manner are still not flexible (e.g., they can’t maintain current phases when there are multiple running vehicles while other competing phases have little queuing vehicles, as shown in Figure[1].
- current traffic state representation approaches neglect the running vehicles in the traffic network;
- latest RL-based TSC methods still have room for improvement under more effective traffic state representation.

To further address the challenge, we pay attention to combining the running vehicles and queuing vehicles for traffic

---

Qiang Wu is the corresponding author. Email: qiang.wu@uestc.edu.cn

1https://github.com/LiangZhang1996/Advanced_XLight

---
state design optimization and introduce the "request" idea from the actuated control method SOTL for TSC. In summary, the main contribution of this paper is as follows:

1. we present a novel, flexible and straightforward method advanced max pressure (Advanced-MP), taking both running and queuing vehicles into consideration to decide whether to change current phase;
2. we design advanced traffic state (ATS), which combines the pressure and effective running vehicles as the traffic state representation;
3. we develop an RL-based algorithm template through combining ATS with RL approaches called Advanced-XLight, and generate two RL algorithms: Advanced-MPLight and Advanced-CoLight.
4. we demonstrate our Advanced-MP method and two RL algorithms could achieve new SOTA.

2 Related Work

In order to mitigate traffic congestion and improve transportation efficiency, various approaches have been proposed. These methods can mainly be divided into two categories: traditional approach and RL-based method.

2.1 Traditional approach

Back in 1958, FixedTime method[12] gives a fixed cycle length and phase split for each traffic light. Subsequently, SCOOT[11], SCATS[16], MP control[19], SOTL[3], etc., try to set the signal according to the traffic conditions on the road, which are still widely used in modern cities.

The idea of MP control [18] is initially developed for scheduling packets in wireless communication networks. Varaiya et al.[19] and Gregoire et al.[9] gave the definition of max-pressure and proved its stability. There have been several works adapting the MP control for signalized intersections. Some studies on MP[23, 10, 9, 13] have demonstrated the efficiency of MP through simulation. Le et.al[14] and Levin et.al[15] proposed cyclical phase structure which has greater palatability for implementation. The Efficient-MP method[24] proposed a simple but efficient traffic state representation (one "lanes to lanes" way to calculate pressure) based on MP, and has achieved SOTA even among the latest RL-based approaches.

Despite the high performance of the MP-based control, it lacks flexibility and need an optimal hyper-parameter such as phase duration and cycle length. For the implementation of MP applications for TSC, some works set a fixed duration for each activated phase[23, 10, 9, 20, 22], while some studies use fixed cycle length and proportionally set phase split[13, 4, 15].

With a fixed phase duration or cycle length, the MP control can’t maintain a phase when a platoon is passing through the intersection, because the "pressure" only concentrate on the queueing vehicles and ignores the running vehicles. As shown in Figure 1 at case (a), there are no running vehicles at current phase, but many queueing vehicles at other phase, the phase signal should not be changed, MP control can’t work properly due to it only concentrates on the queueing vehicles. Thus, MP control can’t get its optimal performance, and we should take running vehicles into consideration to make MP control more adaptive.

Self-organizing traffic lights control (SOTL) [4, 1, 6, 8, 7, 5, 3] is one type of adaptive TSC method, which can autonomously adjust the phase signal according to the traffic state. It evaluates the conditions of the green phase and other competing phases, and adaptively decides whether to maintain or change the current signal phase. The main advantage of the SOTL is its flexibility to dynamic traffic flow. Besides, the SOTL pays attention to the running vehicles crossing through, which could also be essential for traffic optimization. We will introduce the advantages of SOTL into Efficient-MP[24] to make it more adaptive.

2.2 RL-based method

The RL-based methods have excellent performance mainly for mainly two reasons: 1) novel deep neural networks; 2) well-designed state and reward.

FRAP[26] proposes a novel network structure realized by phase competition and relation, which can deal with unbalanced traffic flow and have strong transferability. CoLight[21] uses graph attention networks to realize intersection level cooperation and is capable of handle large scale traffic signal control. PressLight[20] integrate "pressure" into state reward design, accomplish multi-intersection TSC. Integrating the concept of pressure into the RL-based methods could bring significant improvement. MPLight[2] integrates "pressure" into state and reward design and adopt FRAP as the base neural network to control city-level traffic signals. Efficient-CoLight[24] introduces "efficient pressure" (EP) and uses its traffic state representative ability to achieve the SOTA. The representation of traffic state will have a strong impact on the quality of the TSC models.

In this paper, our work focuses on: (1) how to improve Efficient-MP’s performance and make it more adaptive to dynamic traffic? (2) how to represent the traffic state more effectively? (3) how to further improve the performance of the RL-related methods without adding complexity?
3 Preliminaries

In this section, we summarize and organize current definitions of concepts for TSC.

Figure 2: An illustration of traffic network, traffic movements and traffic signal phase.

Table 1: Summary of notations.

| Notation | Meaning |
|----------|---------|
| \(\mathcal{L}_i\) | set of lanes of intersection \(i\) |
| \(\mathcal{L}_i^{\text{in}}\) | set of incoming lanes of intersection \(i\) |
| \(\mathcal{L}_i^{\text{out}}\) | set of outgoing lanes of intersection \(i\) |
| \(l', m', k'\) | lane |
| \(l, m, k\) | road which is the set of lanes |
| \(q(l')\) | queue length of lane \(l'\) |
| \(r(l')\) | running vehicle number at lane \(l'\) |
| \((l, m)\) | a traffic movement from road \(l\) to road \(m\) |
| \(p_e(l, m)\) | efficient pressure of road \(l\) to road \(m\) |
| \(S_i\) | phases of intersection \(i\) |
| \(p(s)\) | pressure of phase \(s\) |
| \(P_i\) | pressure of intersection \(i\) |
| \(t_{\text{duration}}\) | minimum phase duration |
| \(V_{\text{max}}\) | the maximum velocity of vehicles |
| \(L\) | The effective range within \(t_{\text{duration}}\) |
| \(r_s(l')\) | the effective running vehicle |
| \(d(s)\) | the demand of each phase |
| \(\text{ATS}(l, m)\) | advanced traffic state for \((l, m)\) |

**Definition 1** (Traffic network). The traffic network is described as a directed graph, in which each node represents the intersection, and each edge represents the road. Each intersection has incoming roads and outgoing roads, and each road consists of several lanes which determine how the vehicle pass through the intersection, such as turn left, go straight, and turn right. An incoming lane for an intersection is where the vehicle enters the intersection. An outgoing lane for an intersection is where the vehicle leaves the intersection. We denote the set of incoming lanes and outgoing lanes of intersection \(i\) as \(\mathcal{L}_i^{\text{in}}\) and \(\mathcal{L}_i^{\text{out}}\) respectively. We use \(l, m, k\) to denote the roads, and \(l', m', k'\) to denote the lanes. Fig 2(a) illustrates a four-way intersection.

**Definition 2** (Traffic movement). A traffic movement is defined as the traffic traveling across an intersection from one incoming road to one outgoing road. We denote a traffic movement from road \(l\) to road \(m\) as \((l, m)\), in which \((l, m) = \text{set}\{\{l', m'\}\}, l' \in l, m' \in m\). For an intersection with each road has three lanes, each traffic movement contains one entering lane and three exiting lanes. Fig 2(a) uses three green dash lines to describe the traffic movement from the south to west, and there are totally twelve traffic movements.

**Definition 3** (Signal phase). Each signal phase is a set of permissible traffic movements. We denote one phase with \(s\), in which \(s = \text{set}\{\{l, m\}\} \in S\). Figure 2(c) describes the mostly used four phases.

**Definition 4** (Efficient pressure). The efficient pressure of each traffic movement is the difference of average queue length between the upstream and downstream, denoted by

\[
p_e(l, m) = \frac{1}{M} \sum_{i=1}^{M} q(l'_i) - \frac{1}{N} \sum_{j=1}^{N} q(m'_j), l', m' \in m
\]

in which \(q(l')\) represents the queue length of lane \(l'\), and \(\frac{1}{M}\) \(\frac{1}{N}\) are the lane number of road \(l\) and \(m\) respectively.

**Definition 5** (Phase pressure). The pressure of each phase is the sum efficient pressure of the traffic movements that form the phase, denoted by

\[
p(s) = \sum e(l, m), (l, m) \in s, s \in S_i \tag{2}
\]

in which \(e(l, m)\) represent the traffic movement efficient pressure and calculated by equation (1).

**Definition 6** (Intersection pressure). The pressure of each intersection is defined as the difference of queue length between the upstream and downstream, denoted by

\[
P_i = \sum q(l') - \sum q(m'), l' \in \mathcal{L}_i^{\text{in}}, m' \in \mathcal{L}_i^{\text{out}} \tag{3}
\]

in which \(q(l')\) represents the queue length of lane \(l'\).

The summary of notations is listed in Table 1.

4 Method

In this section, we first review the Efficient-MP [24] and SOTL control [3] for TSC. Then we develop the Advanced-MP based on Efficient-MP and SOTL control. Next, we novelty design the traffic movement representation with pressure and effective running vehicles from Advanced-MP, namely advanced traffic state (ATS). Finally, we introduce AST into RL-base models and develop a template called Advanced-XLight.

4.1 Efficient Max Pressure

The Efficient-MP control is one of MP-based control approach. It maximizes the global throughput in the transportation area by reducing "pressure", which indicates the imbalance of queue length. The pressure of an intersection is the difference of the average queue length between upstream and downstream (Definition 6). At intersection \(i\), the phase pressure \(p_e(s)\) of each intersection is calculated by equation (2), and then the intersection pressure \(P_i\) is calculated by equation (3).
(2)), then activate the phase with maximum pressure every $t_{duration}$, denoted by

$$\hat{s} = \arg\max (p_c(s) | s \in S_i)$$

(4)

The Efficient-MP is formally summarized in Algorithm 1.

Algorithm 1: The Efficient-MP Control

Parameter: Current phase time $t$, minimum phase duration $t_{duration}$

for (time_step) do
  $t = t + 1$
  if $t = t_{duration}$ then
    For each intersection, get $p_c(s)$ by equation (2);
    Select the phase according to equation (4);
    $t = 0$;
  end if
end for

With a fixed phase duration, the Efficient-MP control is not optimal, $t_{duration}$ influence the performance of the Efficient-MP a lot (as Figure 3 shown). With prior knowledge to ensure a platoon of vehicles pass through the intersection, we should set a large $t_{duration}$ when the traffic is heavy and set a small $t_{duration}$ when the traffic is smooth.

4.2 Self-organizing traffic lights control

Self-organizing traffic lights control (SOTL) evaluates the traffic state of each phase and adaptively decides which phase to set. For the phase with a red signal, SOTL uses $c_t$ to count the number of cars approaching. When $c_t$ reaches a threshold $\theta$, then the green signal will be set to the phase. For the phase with a green signal, use $t_{step}$ to record the time steps, and the phase will not be changed until $t_{step}$ larger than the minimal phase duration $t_{duration}$. Before changing a green light to red, the controller checks if a platoon of running vehicles is crossing through in order not to break it. If there are more than $\mu$ cars approaching the green light within a distance $\omega$, the green signal will not be changed to red.

The method is formally summarized in Algorithm 2. Although the SOTL can be adaptive to changing traffic flow, the hyper-parameters of $\theta, t_{duration}, \mu, \omega$ should be manually designed with an expert prior knowledge.

Algorithm 2: SOTL

Input: minimal phase duration $t_{duration}$; threshold $\theta$ and $\mu$; distance $\omega$;

Parameters: vehicle counter $c_i^r$ and $c_i^g$; records of time step $t_{step}$

for (timestep) do
  if ($t_{step} \geq t_{duration}$) then
    $c_i^r$ = vehicles approaching red phase
    $c_i^g$ = vehicles approaching green phase within $\omega$
  if not ($0 < c_i^g < \mu$) then
    if ($c_i^r > \theta$) then
      switch light;
  end if
end if
end for

4.3 Advanced Max Pressure Control

To make Efficient-MP control more adaptive, inspired by SOTL, we propose advanced max pressure (Advanced-MP), which takes both pressure and running vehicles near the intersection into consideration.

New Definitions To depict the Advanced-MP, we firstly define new concepts based on basic definitions in the last section as follows:

Definition 7 (Effective range). The effective range is the maximum distance to the intersection that each vehicle can pass through within $t_{duration}$, denoted by

$$L = V_{max} \times t_{duration}$$

in which $V_{max}$ is the maximum velocity of vehicles. For example, if the vehicles’ maximum velocity is 11 m/s, then $L = 110m$ for $t_{duration} = 10s$, and $L = 165m$ for $t_{duration} = 15s$. The effective range is similar to $\omega$ in SOTL but with a more reasonable value.

Definition 8 (Effective running vehicle number). The effective running vehicle number is the number of running vehicles of the incoming lanes within the effective range to the intersection, denoted by

$$r_e(l, m) = \sum r(l'), l' \in l$$

(6)

in which $r(l')$ is the running vehicle number within $L$ of lane $l'$.

Definition 9 (Phase demand). The demand of each phase is the sum of the effective running vehicle number from the phase, denoted by

$$d(s) = \sum r(l, m), (l, m) \in s$$

(7)

in which $r(l, m)$ is calculated by equation (6). The phase demand precisely explains the demand of passing through the intersection. For the current phase, phase demand can represent the need to maintain the current phase for the next
they compete for a green signal. For the current phase, they may be congestion on the downstream.

Phase demand represents the need for running vehicles, and phase pressure expresses the need for queuing vehicles. They compete for a green signal. For the current phase, use $W$ to represent the importance weight of phase demand, and the "request" from the current phase is denoted by:

$$d(s) \times W_1$$

And the "request" from other competing phases is $\max (p_c(s))$. Compare the "request" values and activate the phase with maximum value.

### Algorithm 3: Advanced Max-Pressure control

**Parameter:** time $t$, minimum action duration $t_{duration}$, weights $W_1$, current phase $a_{cur}$

```text
for (timestep) do
  if $t/t_{duration} = 0$ then
    For each intersection, get $p(s)$ and $d(s)$.
    if $t == 0$ then
      $a_{cur} = \arg\max (p(s)), s \in S_i$
    else if $d(a_{cur}) \times W_1 > \max \{p(s)\}$ then
      Maintain the current phase
    else
      $a_{cur} = \arg\max (p(s)), s \in S_i$
  end if
  Set the phase as $a_{cur}$.
  $t = t + 1$
end for
```

The advanced MP method is formally summarized in Algorithm [3]. Through the competition of phase pressure and phase demand, Advanced-MP can realize adaptive control. With a fixed phase duration, we test several simple weights and finally find the best as the final results of this method.

### Discussion

One crucial question is whether Advanced-MP still stabilizes the queue length in the traffic? We will compare Advanced-MP with typical MP-based methods using proportional phase split. Advanced-MP can be regarded as one approximate realization of MP using proportional phase split under optimal cycle length. Advanced-MP does not change the properties of MP but a more adaptive realization. Imaging that MP-based methods could set an optimal phase duration with different traffic conditions, and this duration may vary with different pressure. To get the optimal duration, what if use small time slots, and evaluate whether to change the phase or not according to the traffic state, finally get the optimal duration through joint these slots. Obviously, that is what Advanced-MP has done.

Advanced-MP can be considered as set an optimal phase duration for each activated phase. Clearly, it could stabilize the traffic flows like MP-based methods and achieve superior performance.

### 4.4 Advanced Traffic State

Although the efficient pressure[24] has strong representativeness, it can not express the running vehicles in the traffic network. Now, "pressure" alone is not enough to represent the complex traffic state.

Based on Advanced-MP, we design advanced traffic state (ATS) with efficient traffic movement pressure and effective running vehicles for traffic state representation.

**Definition 10** (Advanced traffic state). The combination representation of the efficient pressure and effective running vehicles for traffic movement $(l, m)$, named advanced traffic state (ATS) and denoted by:

$$ATS(l, m) = \{(p_c(l, m); r_c(l, m))\}$$

is which $p_c(l, m)$ and $r_c(l, m)$ are computed by equation (1) and equation (6), respectively.

### 4.5 Advanced-XLight

We develop an advanced RL-based methods template applying ATS as traffic state, namely Advanced-XLight (as Algorithm 4 shows).

We will adopt MPLight[2] and CoLight[21] as the Q-network base architecture of our method, which have the advantage of high performance. We generated Advanced-MPLight and Advanced-CoLight through our Advanced-XLight algorithm template. It should be noted that the ideal of RL-based design is not limited to MPLight and CoLight, can also be integrated into other RL-based models.

**Algorithm 4: Advanced-XLight**

**Parameter:** Current phase time $t$, minimum action duration $t_{duration}$

```text
for (time step) do
  $t = t + 1$;
  if $t = t_{duration}$ then
    Set the phase by X RL model;
    $t = 0$;
  end if
end for
```

**State representation** Each agent observes the current phase and advanced traffic state, which consists of traffic movement efficient pressure (Definition 4) and effective running vehicles (Definition 8).

**Action** At time $t$, each agent chooses a phase $\hat{s}$ as its action $a_t$, and the traffic signal will be set to phase $\hat{s}$.

**Reward** For Advanced-MPLight model, the reward is the pressure of the intersection, denoted by $r_i = -|P_i|$. The agent try to stabilize the queues in the system by maximizing the reward. For the Advanced-CoLight model, the reward is the total queue length of the intersection, $r_i = -|\sum q(l')|, l' \in L_i'$. The Advanced-CoLight agent try to minimize the queue length of the system.

The Advanced-XLight is updated by the Bellman Equation:

$$Q(s_t, a_t) = R(s_r, a_t) + \gamma \max Q(s_{t+1}, a_{t+1})$$

(10)
5 Experiment

5.1 Experiment Settings
We conduct experiments on CityFlow\cite{25}, a simulator that supports large-scale traffic signal control and has faster speed than SUMO. The simulator provides the state to the agent and receives the phase signal settings. Each green signal is followed by a three-second yellow signal and two-second all red time to prepare the transition.

The traffic dataset consists of the road network data and traffic flow data. The road network data describes the traffic road links, traffic movements, and corresponding signal settings. In the traffic flow dataset, each vehicle is described as \((t, u)\) where \(t\) is time and \(u\) is the pre-planned route.

In multi-intersection TSC, the phase number and minimum phase duration are important hyper-parameters and should be set as the same before conducting baseline. We set the phase number as four and minimum phase duration as 15-seconds, the same as that from Efficient-MP\cite{24}.

5.2 Datasets

We use five real-world traffic datasets in experiments. Three from JiNan and two from HangZhou.

**JiNan datasets** The road network has 12 intersections \((3 \times 4)\). Each intersection is a four-way intersection, with two 400-meter(East-West) long road segments and two 800-meter(South-North) long road segments. There are three traffic flow datasets.

**HangZhou datasets** The road network has 16 intersections \((4 \times 4)\). Each intersection is a four-way intersection, with two 800-meter(East-West) long road segments and two 600-meter(South-North) long road segments. There are two traffic flow datasets.

Table 2: Average arrival rate of the two datasets

| Dataset        | Arrival rate(vehicles/s) |
|----------------|--------------------------|
| \(D_{JiNan_1}\) | 1.75                     |
| \(D_{JiNan_2}\) | 1.21                     |
| \(D_{JiNan_3}\) | 1.53                     |
| \(D_{HangZhou_1}\) | 0.83                     |
| \(D_{HangZhou_2}\) | 1.94                     |

These traffic datasets are not only different from traffic road network, but also arrival rate (as is shown in Table 2). Using these datasets is convincing enough.

5.3 Evaluation Metric

Following existing studies\cite{23} we use the average travel time to evaluate the performance of different models for traffic signal control. It calculates the average travel time of all the vehicles spent between entering and leaving the traffic network (in seconds), which is the most frequently used measure of performance to control traffic signals in the transportation field\cite{26, 21, 22}.

5.4 Compared Methods

We compare our methods with the following baseline methods, including traditional transportation methods and RL-based methods. All the RL methods are trained with the same hyper-parameters (learning rate, replay buffer size, sample size). Each episode is a 60-minutes simulation, and we adopt results as the average of the last ten testing episodes. The final report result is the average of three independent results.

**Transportation Methods:**
- **Fixed-Time[12]**: a policy gives a fixed cycle length with a predefined phase split among all the phases.
- **Max-Pressure[19]**: the max pressure control selects the phase that has the maximum pressure.
- **Efficient-MP[24]**: select the phase with the maximum efficient pressure. It is a SOTA method that has superior performance than CoLight\cite{21} and MPLight\cite{2}.

**RL Methods:**
- **MPLight[2]**: uses FRAP as the base model, and introduce pressure into the state and reward design. MPLight can realize city-level traffic signal control.
- **CoLight[21]**: uses graph attention network to realize intersection cooperation. CoLight can realize large scale traffic signal control.
- **Efficient-PressLight[24]**: one LIT\cite{27} based model, use current phase and efficient traffic movement pressure as observations, intersection pressure as reward. It has a significant better performance than PressLight.
- **Efficient-MPLight**: one FRAP\cite{26} based model, use current phase and efficient traffic movement pressure as observation, intersection pressure as reward. It has a significant better performance than MPLight.
- **Efficient-CoLight**: one Colight\cite{21} based model, use current phase and efficient traffic movement pressure as observation, intersection queue length as reward. It is the SOTA RL method. It has a significant better performance than CoLight.

**Our Proposed Methods:**
- **Advanced-MP**: introduce effective running vehicles and try to find the trade-off with efficient pressure.
- **Advanced-MPLight**: FRAP\cite{26} based model, use current phase, traffic movement efficient pressure, precise running vehicles as observations, intersection pressure as reward.
- **Advanced-Colight**: Colight\cite{21} based model, use current phase, efficient traffic movement pressure, precise running vehicles as observations, intersection queue length as reward.

5.5 Results

**Weight setting** Various testings are needed to get an appropriate \(W_1\) under a particular phase duration for Advanced-MP, and Figure\cite{24} demonstrate how advanced-MP performs under different \(W_1\). From our experience, the demand will need a higher weight with a smaller phase duration for mainly two reasons: (1) the time waste of phase
The performance of our RL-based method integrating AST with the RL-based approaches brings excellent improvements. The efficiency and adaptiveness of transportation methods. Furthermore, compared to RL-based methods, Advanced-MP is easier to deploy because Advanced-MP only requires phase duration and weight $W_1$, and do not need intensive training. (2) Advanced-CoLight outperforms all other methods. Integrating AST with the RL-based approaches brings excellent improvements. The performance of our RL-based methods have consistently significant improvement over Efficient-MPLight, and Efficient-CoLight. Further experiment is conducted to reveal the key exponent that brings the significant improvement.

Table 3: Performance (the average travel time in seconds) comparison of different methods evaluated on JiNan and HangZhou real-world datasets (the smaller the better).

| Method          | JiNan 1 | JiNan 2 | JiNan 3 | JiNan 4 | HangZhou 1 | HangZhou 2 |
|-----------------|---------|---------|---------|---------|------------|------------|
| FixedTime       | 428.11(58.64%) | 368.77(53.81%) | 383.01(59.57%) | 495.57(74.23%) | 406.65(24.21%) |
| MaxPressure     | 273.96(1.52%)   | 245.38(3.25%)   | 245.81(2.41%)   | 288.54(1.44%)  | 348.98(6.52%)  |
| Efficient-MP    | 269.87     | 239.75     | 240.03     | 284.44     | 327.62     |
| MPLight         | 297.46(+10.22%) | 270.05(+12.04%) | 276.15(+15.05%) | 314.60(+10.60%) | 357.61(+9.15%) |
| CoLight         | 272.06(+0.81%)  | 252.44(+5.29%)  | 249.56(+3.97%)  | 297.02(+4.42%)  | 347.27(+6.00%)  |
| Efficient-PressLight | 278.96(+3.37%)  | 253.10(+5.58%)  | 253.16(+5.47%)  | 312.32(+9.80%)  | 414.32(+26.46%) |
| Efficient-MPLight| 261.81(-2.99%)  | 241.35(0.67%)   | 238.80(-0.51%)  | 284.49(0.02%)   | 321.08(-2.00%)  |
| Efficient-CoLight| 256.84(-4.83%)  | 239.58(-0.07%)  | 236.72(-1.38%)  | 282.07(-0.83%)  | 324.27(-1.02%)  |
| Advanced-MP     | 253.61(-6.03%)  | 238.62(-0.47%)  | 235.21(-2.01%)  | 279.47(-1.75%)  | 318.67(-2.73%)  |
| Advanced-MPLight| 251.29(-6.88%)  | 234.78(-2.07%)  | 231.76(-3.45%)  | 273.26(-3.93%)  | 312.68(-4.56%)  |
| Advanced-CoLight| 245.73(-8.95%)  | 232.63(-2.97%)  | 229.01(-4.59%)  | 270.45(-4.92%)  | 310.74(-5.15%)  |

Overall performance Table 3 reports our experimental results under JiNan and HangZhou real-world datasets with respect to average travel time. We have the following findings: (1) The Advanced-MP consistently outperforms all other previous methods. With a manually designed weight $W_1$, Advanced-MP can outperform the SOTA RL method: Efficient-CoLight. Advanced-MP demonstrates the efficiency and adaptiveness of transportation methods. Furthermore, compared to RL-based methods, Advanced-MP is easier to deploy because Advanced-MP only requires phase duration and weight $W_1$, and do not need intensive training. (2) Advanced-CoLight outperforms all other methods. Integrating AST with the RL-based approaches brings excellent improvements. The performance of our RL-based methods have consistently significant improvement over Efficient-MPLight, and Efficient-CoLight. Further experiment is conducted to reveal the key exponent that brings the significant improvement.

Variation with phase duration Minimum phase duration(minimum action time, $\tau_{duration}$) is an essential hyper-parameters, and it influences the control performance of both transportation and RL-based methods. To prove our proposed methods have a better performance under different phase duration, experiments under different phase duration are conducted.

Figure 5 reports the model performance on JiNan and HangZhou datasets under different phase duration. We find that: (1) Phase duration has a great influence on the model performance. All the methods have different model performances under different phase duration. There is a general trend that models could get a better performance with smaller phase duration. (2) Our proposed methods outperform other methods over most the datasets under different phase duration(except that under HangZhou $d_{duration} =$ 30s).

Effective observation Does the more complex observation that makes the better performance of the Advanced-XLight and what is the key factor? To answer this question, an additional experiment is conducted with different state transition (yellow time and all-red time). The wasted time is fixed. Smaller phase duration will potentially cause more frequent transition and more wasted time. (2) The smaller phase duration will use a shorter efficient range (Definition 7), leading to a smaller value of traffic demand. Therefore the weight should set a higher value. Therefore, we should give a higher weight for demand with a small phase duration and a lower weight with a large phase duration.

Figure 4: The weight $W_1$ that Advanced-MP used to get an optimal performance.

Figure 5: Model performance under different phase duration, the smaller the better.
The configurations are as follows:

- **Config1.** The RL-models use effective running vehicles, traffic movement pressure, and current phase as the state representation, which is exactly Advanced-XLight.
- **Config2.** The RL-models use total running vehicles of incoming lanes, traffic movement pressure, and current phase as the state representation.
- **Config3.** The RL-models use traffic movement pressure and current phase as the state representation, which is exactly Efficient-XLight.

Figure 6 reports the performance under different observations. Model can’t get better performance with more complex observations. With additional observation of total running vehicles on the incoming lanes, the MPLight based model works worse, and the CoLight based model has no apparent changes. The model performance is significantly improved with the effective observation of running vehicles, both the MPLight based model and the CoLight based model. The results demonstrate the importance of effective observation of running vehicles.

Therefore, the effective observation of running vehicles is the key for model improvement. And the excellent performance of Advanced-CoLight and Advanced-MPLight should attribute to effective observation rather than more complex observation.

### 6 Conclusion

This paper proposes a novel method called Advanced-MP based on MP and SOTL and designs advanced traffic state (ATS) for the traffic movement representation with pressure and precise running vehicles. Experiments results show that our ATS expresses more information for traffic state and boosts the TSC methods. In the future, we will analyze more traffic factors and provide a more precise traffic state representation to further optimize the TSC methods.

### 7 Acknowledgments

### References

[1] Burguillo-Rial, J. C.; Pedro, S.; Montenegro, E. C.; and Gil, F. 2009. History-based self-organizing traffic lights. *Computing and Informatics*, 28(2): 157–168.

[2] Chen, C.; Wei, H.; Xu, N.; Zheng, G.; Yang, M.; Xiong, Y.; Xu, K.; and Li, Z. 2020. Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 3414–3421.

[3] Cools, S.-B.; Gershenson, C.; and D’Hooghe, B. 2013. Self-organizing traffic lights: A realistic simulation. In *Advances in applied self-organizing systems*, 45–55. Springer.

[4] Gershenson, C. 2004. Self-organizing traffic lights. *arXiv preprint nlin/0411066*.

[5] Gershenson, C. 2012. Self-organizing urban transportation systems. In *Complexity Theories of Cities Have Come of Age*, 269–279. Springer.

[6] Gershenson, C.; and Rosenblueth, D. A. 2009. Modeling self-organizing traffic lights with elementary cellular automata. *arXiv preprint arXiv:0907.1925*.

[7] Gershenson, C.; and Rosenblueth, D. A. 2012. Adaptive self-organization vs static optimization: A qualitative comparison in traffic light coordination. *Kybernetes*.

[8] Gershenson, C.; and Rosenblueth, D. A. 2012. Self-organizing traffic lights at multiple-street intersections. *Complexity*, 17(4): 23–39.

[9] Gregoire, J.; Frazzoli, E.; de La Fortelle, A.; and Wongpiromsarn, T. 2014. Back-pressure traffic signal control with unknown routing rates. *IFAC Proceedings Volumes*, 47(3): 11332–11337.

[10] Gregoire, J.; Qian, X.; Frazzoli, E.; De La Fortelle, A.; and Wongpiromsarn, T. 2014. Capacity-aware back-pressure traffic signal control. *IEEE Transactions on Control of Network Systems*, 2(2): 164–173.

[11] Hunt, P.; Robertson, D.; Bretherton, R.; and Royle, M. C. 1982. The SCOOT on-line traffic signal optimisation technique. *Traffic Engineering & Control*, 23(4).

[12] Koonce, P.; and Rodegerdts, L. 2008. Traffic signal timing manual. Technical report, United States. Federal Highway Administration.

[13] Kouvelas, A.; Lioris, J.; Fayazi, S. A.; and Varaiya, P. 2014. Maximum pressure controller for stabilizing queues in signalized arterial networks. *Transportation Research Record*, 2421(1): 133–141.

[14] Le, T.; Kovács, P.; Walton, N.; Vu, H. L.; Andrew, L. L.; and Hoogendoorn, S. S. 2015. Decentralized signal control for urban road networks. *Transportation Research Part C: Emerging Technologies*, 58: 431–450.

[15] Levin, M. W.; Hu, J.; and Odell, M. 2020. Max-pressure signal control with cyclical phase structure. *Transportation Research Part C: Emerging Technologies*, 120: 102828.
[16] Lowrie, P. 1992. SCATS: A traffic responsive method of controlling urban traffic control. Roads and Traffic Authority.

[17] Nishi, T.; Otaki, K.; Hayakawa, K.; and Yoshimura, T. 2018. Traffic signal control based on reinforcement learning with graph convolutional neural nets. In 2018 21st International conference on intelligent transportation systems (ITSC), 877–883. IEEE.

[18] Tassiulas, L.; and Ephremides, A. 1990. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. In 29th IEEE Conference on Decision and Control, 2130–2132. IEEE.

[19] Varaiya, P. 2013. Max pressure control of a network of signalized intersections. Transportation Research Part C: Emerging Technologies, 36: 177–195.

[20] Wei, H.; Chen, C.; Zheng, G.; Wu, K.; Gayah, V.; Xu, K.; and Li, Z. 2019. Presslight: Learning max pressure control to coordinate traffic signals in arterial network. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 1290–1298.

[21] Wei, H.; Xu, N.; Zhang, H.; Zheng, G.; Zang, X.; Chen, C.; Zhang, W.; Zhu, Y.; Xu, K.; and Li, Z. 2019. Co-light: Learning network-level cooperation for traffic signal control. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 1913–1922.

[22] Wei, H.; Zheng, G.; Gayah, V.; and Li, Z. 2019. A Survey on Traffic Signal Control Methods. arXiv preprint arXiv:1904.08117.

[23] Wongpiromsarn, T.; Uthaicharoenpong, T.; Wang, Y.; Frazzoli, E.; and Wang, D. 2012. Distributed traffic signal control for maximum network throughput. In 2012 15th international IEEE conference on intelligent transportation systems, 588–595. IEEE.

[24] Wu, Q.; Zhang, L.; Shen, J.; Lü, L.; Du, B.; and Wu, J. 2021. Efficient Pressure: Improving efficiency for signalized intersections. arXiv:2112.02336.

[25] Zhang, H.; Feng, S.; Liu, C.; Ding, Y.; Zhu, Y.; Zhou, Z.; Zhang, W.; Yu, Y.; Jin, H.; and Li, Z. 2019. Cityflow: A multi-agent reinforcement learning environment for large scale city traffic scenario. In The World Wide Web Conference, 3620–3624.

[26] Zheng, G.; Xiong, Y.; Zang, X.; Feng, J.; Wei, H.; Zhang, H.; Li, Y.; Xu, K.; and Li, Z. 2019. Learning phase competition for traffic signal control. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 1963–1972.

[27] Zheng, G.; Zang, X.; Xu, N.; Wei, H.; Yu, Z.; Gayah, V.; Xu, K.; and Li, Z. 2019. Diagnosing reinforcement learning for traffic signal control. arXiv preprint arXiv:1905.04716.