Trade-off between Safety and Congestion: an Application in Traffic Reinforcement Learning

A Preprint

Chenyang Zhu, Wenqi Chen, Aihua Li, Kanghong Yuan, Yajie Zhang
Shanghai University of Finance and Economics
Shanghai, China
{chenyang.zhu, wenqichen, liaihua, yuankanghong, zhangyajie98}@163.sufe.edu.cn

December 3, 2019

Abstract

Injuries of students by distracted drivers have been reported around a prestigious university in Shanghai, where terrible traffic designs have long been accused for those tragedies. Thus, there has long been a debate on whether or not to set up traffic lights in the post-accident areas where students can then be safe from distracted or over-speeding drivers. Objectors claim that this will slow down the already stressful traffic system and lead to a significant increase of travelling time. Faced with the urgent and controversial dilemma, we promote a study using a high-frequency data set from Shanghai Qiangsheng Taxi, and by designing a stochastic model and a multi-agent reinforcement model, we show that the traffic can be improved by 50% from the current standard, even when a traffic light is set up to protect students' safety. The models we proposed in this paper are universal and adaptable for all similar scenarios, and we hope that this paper could raise awareness of government on adopting new technologies to improve the current inefficient traffic system and also systematically prevent people from being hurt by distracted driving and over-speeding.

Keywords Traffic Signal Duration Adaptation · Multi-agent System · Reinforcement Learning · Real-life Application

1 Introduction

There has long been a debate between adding a traffic light to ensure pedestrian safety and cancelling a traffic light to make sure that traffic flows faster. In areas with heavy traffics, city planners are less willing to add new traffic lights, because that always means that the traffic would be significantly slowed down. On the other hand, this makes people to have a hard time to find a break in the traffic to cross the road.

Another problem in modern traffic is over-speeding driving. People are long been reported to be killed accidentally and innocently by over-speeding, who commit murder to innocent people only because of their own mistakes. While local authorities find it hard to control over-speeding, some set up more traffic lights to at least mitigate some over speeding accidents. This again would create traffic congestion.

Our paper focuses on the balance between adding a traffic lights and traffic congestion. We use a high-frequency data set from Shanghai Qiangsheng Taxi for solving traffic congestion at an accidentally-dense area. We propose a stochastic model and a reinforcement model. Both models show that, with careful planning, the traffic congestion brought by adding a traffic light at the accidentally-dense area could be effectively controlled.

We conduct the optimization and experiment on a specific area, where a close friend of our research group and many other people were injured by over-speeding cars. The area involves a obtuse angle intersection, where cars are more likely to over-speed. This area also has a huge pedestrian crossing every day.

While our experiment focuses on a specific area, our model is universal and is adaptable for all similar cases. We hope that our study could provide a way to solve this long-lasting debate, and also prevent people from getting injured by both the poor traffic design and over-speeding drivers.

Above all, on the perspective of target setting, most of the existing research results focused on a single traffic light, but this paper has considered the linkage system of multiple traffic lights, trying to eliminate the limitations caused by single traffic light research.

Additionally, Focusing on the research method, compared with the Autoregressive Conditional Heteroskedasticity (a.k.a. ARCH) model and cellular automaton used in the existing research, this paper will combine the latest research theories in the academic field, such as Q-learning algorithm in Reinforcement Learning, to create more innovative and effective models.

More importantly, in the field of applying reinforcement learning methods to transportation, although there is practical application on regional intersections, it is still altered by controlling single intersection inside, and few studies research on controlling multiple intersections at the same time. We directly establish a large-scale optimization system for controlling multiple intersections, which has strong versatility. When the number of intersections to be studied increases, they can be directly added to the model.

And the most essential factor is that compared with the multi-agent system using DQN, this paper uses real data combined with probability model to offer a more convincing setting thanks to the development of data gathering. Instead of using grid data, it is more convenient to follow our process of data preprocessing and our model will be more efficient to be widely used in the large-scale traffic light control in reality. After all, the equipment requirement of DQN is demanding and its functioning speed will be relatively slow. Therefore, we have created a practical model and a process of data disposing, most available for the government to apply in the construction of large-scale traffic light control system.

2 Previous works

In the field of traffic congestion research, traditional literature focuses on analyzing traffic congestion features and setting up analyzing factors. In recent years, many are combining big data and Internet of Things (IoT). Others use ARCH model to solve data combination and path optimization. However, in general, most of these papers are based on traditional concepts in traffic system and changing the non-systemetic traffic lights. There are still limited application of big data and AI system in the research area of traffic.

With the development of data science, more and more algorithms in Machine Learning are applied to optimizing traffic congestion. The one aspect that we are interested in is reinforcement learning. El-Tantawy et al. [1] summarized related works from 1997 to 2010 that use Reinforcement Learning to optimize traffic controls. However, these works only use tabular Q-learning and discrete state spaces. Watkins et al. [2] proposed the Q learning method as an extension of the application of reinforcement learning. Q-learning could still find the best action strategy, even when the reward function and state transfer function is not available.

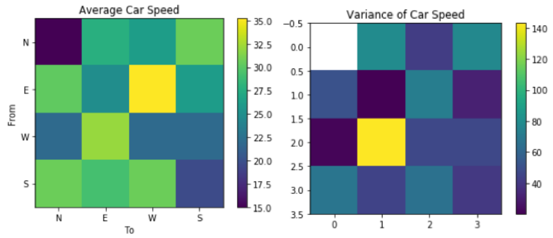
Most reinforcement learning papers in the field of optimizing traffic congestion focus on single crossroad traffic control. This is essentially true as multi-agent reinforcement learning would often lead to non-optimal solutions. Abdulhai et al.[3] use Q-learning to control a single crossroad traffic light, and shows that under fixed traffic flow, the reinforcement algorithm could reproduce the same traffic light rules as the signal set by human. But, under more complex and dynamic traffic flows, the reinforcement learning method is better. These algorithms have good performance in single small roads, but are not suitable for traffic networks in big cities.

Therefore, many researchers start to look at multi-crossroad reinforcement learning. Wiering [4] created reinforcement learning with models, and proved that in the scenario of large traffic flows, reinforcement learning is significantly better than traditional fixed-length traffic control. Based on Wiering's model, many researchers adopt this method in many other applications. Steingrover et al. [5] introduced information sharing between traffic lights, which takes the surrounding crossroads' congestion into consideration when adjusting traffic lights.

Furthermore, with the development of deep learning, Deep Q Network, which is a revised Q-learning algorithm based on deep learning, has been introduced into the research of traffic light control. Liang and Du [6] in 2018 compared different methods in DQN, such as Dueling and Double, and proposed a double dueling deep Q network, reducing over 20% of the average waiting time from the start training. The motivation of using DQN is that they cut the intersection into small grids so the data becomes too complex to be disposed by traditional methods. But the disadvantage of his model is that they only focused on singular intersection. Based on that, Chu et al. [7] in 2019 created a multi-agent traffic control system based on deep reinforcement learning. However, their method also relies on the grid data, which requires a difficult process of data processing.

Research in traffic congestion often needs simulations. Most papers use SUMO for simulation (see, e.g. [8], [9]) in simulating software to create simulated experiment. In our paper, we use the GPS signal from Shanghai Johnson & Johnson taxi. The data consists of one month's data of all the taxi's in Shanghai, including its location, speed and direction. With the real data, we can provide a closer-to-reality simulation than the software does. As the goal of this paper is to mitigate the traffic congestion around our school.

3 Experiment Setup



(a) The mean speed of taxis (b) The variance of speed.

Figure 1: Average taxi speed and variance of cars at the intersection of Zhengli Road. and Guoding Road. The Y-axis shows the direction that the car comes from, while the x-axis represents the direction the car is going to. The color palette shows the value of the speed or variance.

Based on the models established before, we apply it into an urging problem in reality, illustrating the application approach of our model and using the real case to compare the performances of two models.

For the surrounding intersections which we selected, we find that during rush hours, the number of cars increases significantly and causes congestion. The congestion could also be identified with the speed of taxis.

Meanwhile, there has long been a question under debate about whether it is reasonable to add a traffic light in front of Shanghai University of Finance and Economics. Our model can be used to clarify the problem and reduce the congestion in the same time.

There are 2 intersections near the school gate, one in north 260m and the other in south 440m. There is no traffic light in front of the gate and thus crossing the road is often unsafe. Hence, we apply the models proposed to prove the feasibility of adding a new traffic light.

Besides, the angle of the intersection from Zhengli Road to Guoding Road is obtuse, where cars are more likely to over-speed, compared with a 90-degree turn. This part looks into detail whether to add a new traffic light to the school gate using the proposed model.

After pre-investigation and visualizing the minute data of all taxi running situation from Shanghai Qiangsheng Taxi Company (including headlight status, car speed, car direction, etc.) on all 30 days in April, 2015, we focus on solving the traffic congestion problem in rush hours. As a result, this paper will try to solve the following three problems:

For a single intersection, we need to establish a dynamic optimization model for adjusting the length of traffic lights according to the traffic flow, in order to optimize the traffic light setting. We will take the Guoding Road - Zhengli Road intersection as example, do simulation with the established model, and show the optimization results.

For multiple intersections, we will use the Q-learning algorithm in reinforcement learning to create another dynamic optimization model for the linkage adjustment of the traffic lights phase, in order to achieve global optimization in all relevant intersections. We take Zhengli Road - Wuchuan Road intersection, Zhengli Road-Guoding Road intersection, Zhengli Road-Songhu Road intersection and Guoding Road-Zhengmin Road intersection as examples, doing simulation with the current model, and show the optimization plan and results.

In the view of whether it is necessary to set a traffic light at the Guoding Road Entrance of Shanghai University of Finance and Economics to prevent future traffic accidents, we try to use the established model to prove that adding traffic lights will not increase the waiting time too much, and the waiting time could be decreased due to the optimization of the traffic lights.

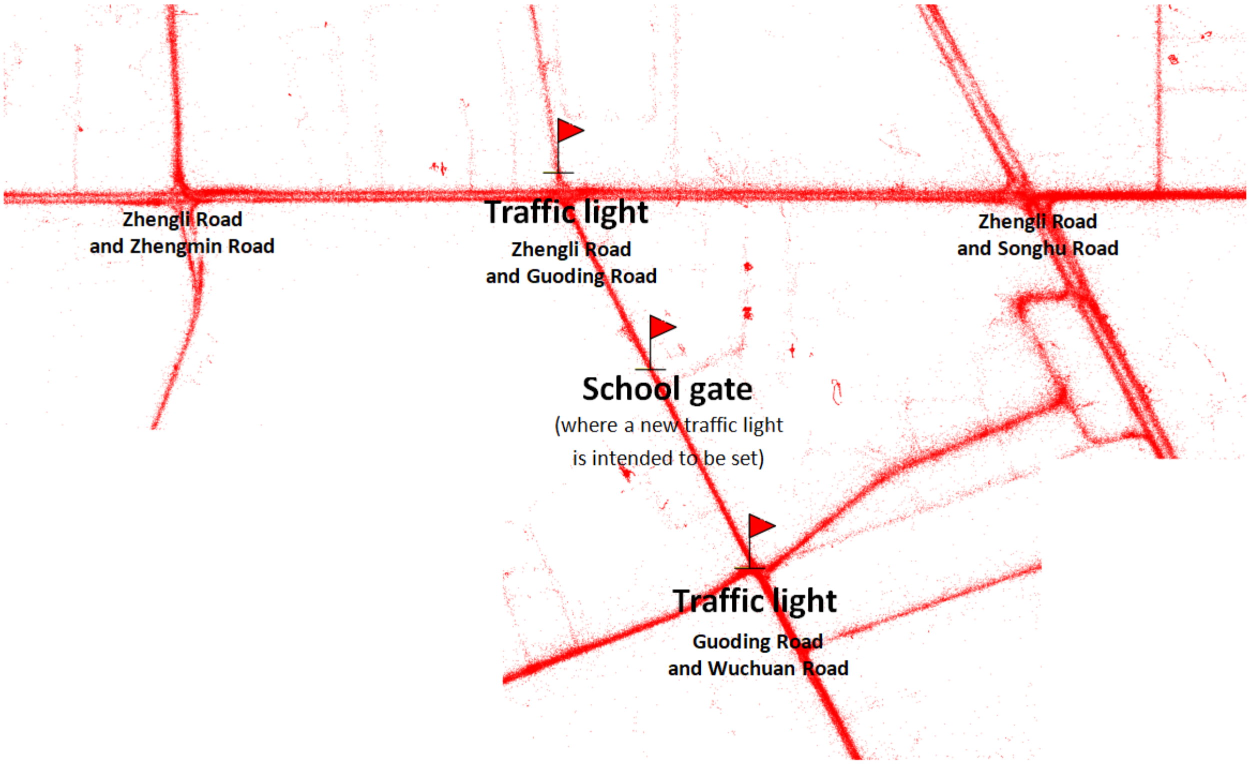


Figure 2: A map where all the red dots are a historical location of the taxis. This plot resembles the true map.

4 Model I. Multi-intersections Traffic Light Control System Based on Reinforcement Learning and Q-learning Algorithm

Define *signal status* as a set of traffic lights in all directions.For example, the north-south green light and the east-west red light are a state of the signal light.

Usually, the state of the signal light is a loop. There may be a loop of three to four or more states at a busy intersection. There are two sets of state loops at ordinary intersections. At the beginning of the project, we do not modify the status of the signal light. Based on existing signal state, we merely modify the duration of each state to observe optimization result.

Suppose that a certain intersection have s states. For example, cars coming from north to west and then cars coming from west to north is a single state. We call a sequential change of the s states a *loop*. The variable to be adjusted later in the model is the duration of each status signal. T_n, where $n \in [1, s], n \in N^+$ represent s states. We take s equals to 3 as in our paper.

Reinforcement learning is a type of machine learning algorithms which allows the agents or the machines to automatically determine the actions to take in a particular context, so as to enhance its performance to the best. The agent takes an action in the environment. Then the environment will change due to the action while also provide reward to the agent representing the feedback of its action. The agent employs trial and error to come up with the best action sequence and get the corresponding highest reward.

In this paper, the state matrix is defined as

$$s_t = (Y_{ij}^t)_{4*4} \tag{1}$$

where Y_{ij}^t represents the number of queuing cars in intersection i and direction j. The 4×4 state matrix represents the congestion level in an area of 4 intersections.

The reward given by the environment is

$$r_t = \frac{1}{\sum_{i \in I, j \in J} Y_{ij}^t} \tag{2}$$

That is, the fewer the total queuing cars are, the better the environment is. The action taken by the agent, or precisely speaking the traffic lights, is defined as

$$a_t = (a_{ij}^t)_{4*3} \tag{3}$$

where $a_{ij}^t = 1$ represents the color of the traffic light in intersection i and direction j.

Q-learning is a commonly used algorithms solving reinforcement learning model. It raises the concept of Q value to represent the total expected reward in the future conditional on the given $(state, action)$. Q-learning aims employs simulation and iteration to find out the optimum strategy to maximize the total expected return. Q value is defined as

$$Q_t(s, a)^{\pi} = E[\sum_{t' \geq t} \gamma^{t'} r_{t'} | s_0 = s, a_0 = a] \tag{4}$$

where γ is the discount factor and $\gamma \in [0, 1)$. A higher γ denotes the kind of agent putting emphasis on the future rewards, which a lower γ denotes the one cares more about the current rewards. It indicates how important the future long-term rewards are.

The targeted optimum strategy is

$$\pi^* = \arg\max_{\pi} Q_t^{\pi}(s, a) = \arg\max_{\pi} E[\sum_{t' \geq t} \gamma^{t'} r_{t'}] \quad (5)$$

Bellman iteration equation is proposed to approximate the global optimum solution of Q value, the direct calculation of which is impossible even in a rather small and finite set of $(state, action)$ since it is hard to completely know the future scenario. Hence, this paper employs the following iteration equations to estimate and approximate the optimum Q value:

$$\delta_{t+1}(Q) = R_{t+1} + \gamma \max_{a' \in \mathcal{A}} Q(Y_{t+1}, a') - Q(X_t, A_t) \quad (6)$$

$$Q_{t+1}(x, a) = Q_t(x, a) + \alpha_t \delta_{t+1}(Q_t) \mathbb{I}_{\{x = X_t, a = A_t\}} \quad (7)$$

, where $(x, a) \in \mathcal{X} \times \mathcal{A}$. Based on the model, the algorithm can be presented as below,

Algorithm 1 An example for format For & While Loop in Algorithm

```
for n in N iterations do
    for t in range(T) T is the max duration of time do
        Sample the arriving cars $X_t$
        Calculate the queuing cars $Y_t = Y_{t-1} + X_t$
        Compute $Q$ values after different actions
        Determine the best strategy based on $Q$ values $\pi$
        Carry out $\pi$ and update the state
    end for
    Iterate $Q$ value matrix
end for
```

5 Mode II. Single Intersection Dynamic Optimization System

In this section, we introduce a new model to compare with the reinforcement learning model. The model in this section is based on a discrete time frame. We divide the continuous time into discrete time period in which one unit is 5 seconds, expressed in t. We first explain the situation of the traffic light in this intersection, and then the vehicle flow probability model is established. Finally we determine the objective function.

5.1 Model Setup

Set the direction variable i, j, and $i, j \in S = \{East, South, West, North\} := \{1, 2, 3, 4\}$. Let $X(t)$ represent the number of vehicles entering the system from all directions in the t time period, which is a 4×4 diagonal matrix:

$$X^{(t)} = \begin{pmatrix} X_1^{(t)} & 0 & 0 & 0 \\ 0 & X_2^{(t)} & 0 & 0 \\ 0 & 0 & X_3^{(t)} & 0 \\ 0 & 0 & 0 & X_4^{(t)} \end{pmatrix} \quad (8)$$

The random variable $X_i^{(t)} \sim Possion\left(\lambda_i^{(t)}\right)$ indicates the number of vehicles coming from the i direction, and the parameter estimate $\hat{\lambda}_i^{(t)}$ will be estimated from actual data. We first estimate $\lambda_i^{(t)}$ for each hour separately, and then the subsequent time interval will be shortened to five or ten minutes. The high-frequency data are used to estimate this parameter more accurately.

After each car comes from the i direction, there is a certain probability of driving out in one of the other three directions. $Y_{ij}^{(t)}$ denotes the number of vehicles to go from i to j direction. Let $Y^{(t)} = (Y_{ij}^{(t)})_{4 \times 4}$. Then $Y(t)$ and $X(t)$ satisfy $Y^{(t)} = X^{(t)} P^{(t)}$, where $P^{(t)}$ is the probability matrix, of which the element represents the probability that the vehicle drives from i to j during the t time period. Its estimated value $P_{ij}^{(t)}$ is estimated by real data according to the following equation,

$$\hat{P}_{ij}^{(t)} = \frac{\text{\# of cars from } i \text{ to } j}{\text{\# of cars from } j} \quad (9)$$

Notice that $\sum_{j=1}^{4} P_{ij}^{(t)} = 1$. The number of vehicles leaving the system in all directions after turning from the intersection within t time period (5 seconds) is $W_{ij}^{(t)} = 5V_{ij}^{(t)}/L$, where L denotes the average length of the vehicle and $V_{ij}^{(t)}$ denotes the speed at which vehicles pass from i to j by the intersection at time t. Finally, the number of vehicles queued in a section is denoted as $Z^{(t)} = Y^{(t)} - W^{(t)}$. Up to now, the vehicle flow model has been established.

5.2 Objective function

For the objective function, we believe that the length of the traffic lights should be adjusted according to the number of vehicles lined up in real time, so as to minimize the total number of vehicles queuing in four directions in a loop and accordingly the traffic condition is improved. We first establish the relationship between traffic lights and queued vehicles: $T_n^{(t)} = a_n \sum_{i,j} Z_{ij}^{(t-1)}$. Here, i and j take the direction in which T_n is correspondingly releasing cars. This is a straight linear relation, which shows that at a certain time, the time length of a certain direction signal lamp depends on the number of vehicles waiting in line in the previous time period. The parameter a_n is the variable to be optimized. The objective function of the model is,

$$\min_{a_1,a_2,a_3} \quad E(T_1\sum_{i,j} Z_{ij}^{(t)} + T_2 \sum_{i,j} Z_{ij}^{(t)} + T_3\sum_{i,j} Z_{ij}^{(t)}) \tag{10}$$

In summary, we will adjust the parameter a_n to minimize the objective function (10) which matches the duration of the signal to the number of backlogs of the vehicle at the current intersection. In addition, when cars are released in one direction, the vehicles in the remaining directions will appear in the queue. $T_1\sum_{i,j} Z_{ij}^{(t)}$, $T_2\sum_{i,j} Z_{ij}^{(t)}$, $T_3\sum_{i,j} Z_{ij}^{(t)}$ denote the number of vehicles lined up on all sections of the road when the traffic light is in state n respectively. In the loop of three states of traffic lights at the intersection, the total number of vehicles in queue is the sum of the three, from which we get the above objective function.

6 Simulation and Results

6.1 Simulation and Results of Model 1

We define one unit of time to be 1 sec. The traffic lights in each intersection change their color once every 5 units of time. Each simulation takes 20000 units of time. The discount factor γ is set to be 0.95. The simulation is conducted in Python.

Figure 3 shows the simulation results in 4 intersections, comparing the original number of queuing cars with the optimization results. Clearly the original data have high peaks, representing the long queues and heavy congestion, while the reinforcement learning model significantly alleviate this situation since the average number of queuing cars is much lower than before. Hence, our model is significantly useful for shortening the queues of waiting vehicles and thus smoothing the traffic.

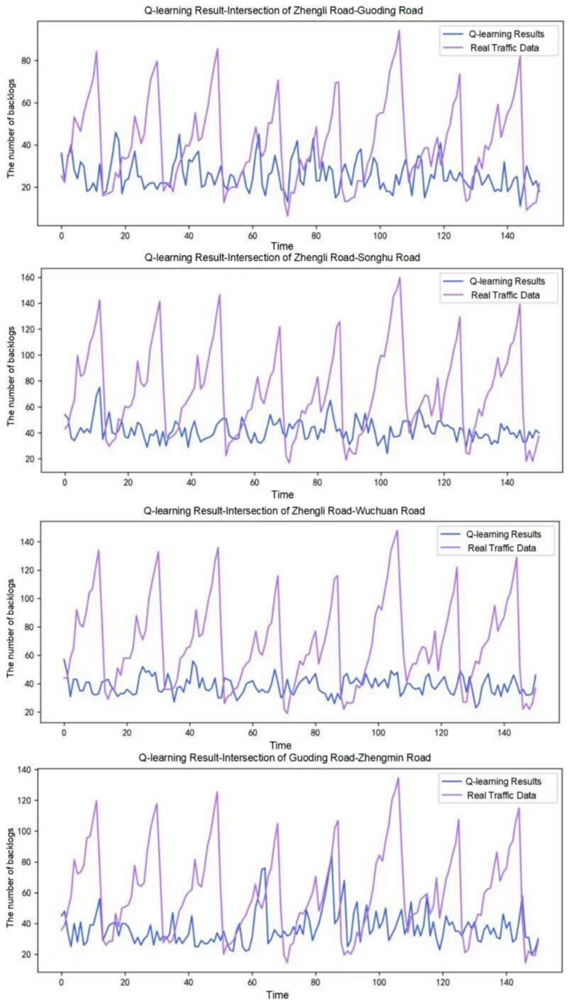


Figure 3: Simulation result of Q-learning algorithm

We compare the traffic performance without traffic light, the performance after adding a traffic light based on Model 1, and the one based on Model 2. The road is about twenty meters wide and seven hundred meters long. We allow thirty seconds for pedestrian to cross the road and sixty seconds for vehicles to pass. In rush hour it will take cars about 140 seconds to go from the light in the north to the south, while in other periods it will take from fifty to ninety seconds.

6.2 Simulation and Results of Model 2

To do the simulation, we should first find out the distribution of the speed of the cars, as shown in Fig. 1b. The speed at which the vehicle heading j passes through the intersection is determined by the actual data. In one direction, the visualization of $V^(t)_{ij}$ data is shown in figure 8, and the rest is similar. It can be seen that, $V^(t)_{ij}$ obeys the normal distribution, and the visualization results of its mean and variance are shown in figure 6 and 7. According to the visual results, the driving speed of Zhengli Road is generally faster. However, the turning speed of Zhengli Road to Guoding Road is slower.The speed variance of the straight direction of Zhengli Road is large, which indicates the instability of the traffic flow. In addition, because of the complexity of the signal light state at the intersection, it is difficult to express the signal lamp length in a matrix form similar to the above norm; in other words, a variety of different vehicles situation exist in one signal state. Therefore, we need to determine the status of the signal light when the vehicle in the system passes through the intersection, so as to determine the subscript n, and then allocate the time T.

Since we have data of Qiangsheng taxis, in order to realistically simulate all vehicle conditions, we use the multiplier 50 to process the number of taxis to estimate the total number of vehicles. Then we use Python for computer simulation to implement the model. In order to get the initial condition of the road, we first let the system perform ten loops, and take the average of the 11th to 20th waiting numbers as the objective function of the required optimization.

At the same time, in order to solve the peak congestion problem, we take the time interval from 11:00 to 12:00. Finally, using the heuristic algorithm, we get the optimal values of a1, a2, a3 are 0.08, 0.10, 0.09 respectively. Under this condition, we can optimize the queued waiting vehicle value from 210 in the previous cycle to 107.Congestion has eased by 50%.

6.3 Comparison of Two Models

The results above can be compared to demonstrate the performance of the two models. In the proposed model based on Q-learning, the congestion can be reduced by 50% over the four intersections. And this model is more applicable since any number of intersections can be added to it and there is no extra assumptions on the transportation condition. However, in the comparative model based on mathematical method, it can only deal with singular intersection once and reduce the congestion by 43%. Considering the applicability and performance of the two models, it is reasonable to use the model based on Q-learning in the future application.

Figure 4 compares the two model outcomes discussed above, the horizontal axis of which represents the 24hrs in a day and the vertical axis shows how long it will take to go through the road with 700 meters long.

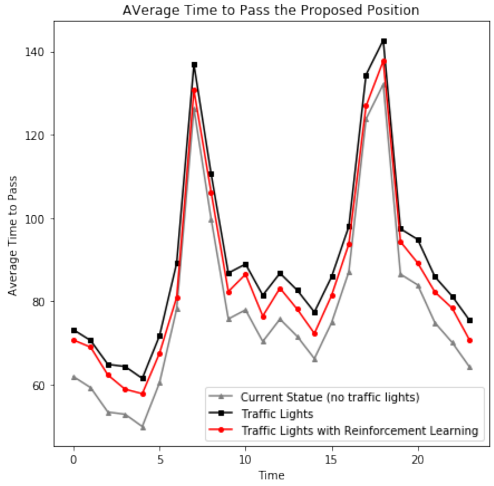


Figure 4: The average time required to pass the proposed position. The grey line is the current state without any traffic lights. Results show that our reinforcement method could do better than simply adding a new traffic light to the proposed position.

Trivially, it takes the shortest time for cars the pass through the road if there is not a new traffic light in the middle. Model 1 shows that an average of additional eleven seconds are necessary to establish a new light. However, if reinforcement learning are applied in the north and south intersections, 3.8 seconds can be deduced from the 11 seconds. That is, adding a traffic light in front of the school gate will increase the vehicles' passing time by 7.2 seconds.

7.2 seconds is quite acceptable, if we take into consideration that in reality it will take some time for vehicles to shift down or even stop to wait for pedestrians' passing without light, or more importantly, if we compare the cost of traffic accidents. Therefore, setting a traffic light in front of the university gate while also applying reinforcement learning model to the traffic light control system will guarantee the passengers' safety as well as ensure a low level of congestion.

7 Conclusion

In this paper, we proposed a multi-agent system for controlling traffic lights in multiple intersections based on reinforcement learning by taking into account the number of waiting cars in all of the target intersections. Additionally, the number of involved intersections can be adjusted in order to adapt the total traffic condition at the regional level instead of focusing on only one intersection.
In order to compare the proposed method with the system on singular intersection, we implement our system in the real-life data in Shanghai and our model is therefore easily adaptive to the real application. Based on the comparison of the two models on real data, the multi-agent system has proved to perform better than the singular system without the assumptions on the traffic conditions, rendering the proposed model to be flexible and efficient. More importantly, this model can simultaneously solve the problem of whether we should set traffic light in front of the gate of SHUFE, symbolizing the fact that our model can be directly used to tackle the real-life problem.
Now our system has been set as an open source for the government applying to construction of traffic system.

8 Discussion and future work

In this paper, we have a full data set from all of the taxis in Shanghai Qiangsheng Taxi Company. However, under real circumstances, there are many other participants on the road. We should take all types of vehicles into consideration, including trucks and private cars. Though collecting data from all types of vehicles is unrealistic, our method is to approximate the ratio of all cars over taxis. We also consider that the speed of the cars could be too simplified. We have not taken accelerating and decelerating process into consideration, making our model less robust when it comes to real scenario. In the future adaptation of this model into real application, our model can be used to offer a example of how to construct more compelling and practical models based on the raw data, but there are also many improvements to make with more reliable data.

References

[1] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, “Design of reinforcement learning parameters for seamless application of adaptive traffic signal control,” *J. Intell. Transp. Syst. Technol. Planning, Oper.*, 2014.

[2] C. J.C. H. Watkins and P. Dayan, “Q-learning,” *Machine Learning*, vol. 8, no. 3, pp. 279–292, 1992, ISSN: 1573-0565. DOI: `10.1007/BF00992698`. [Online]. Available: `https://doi.org/10.1007/BF00992698`.

[3] B. Abdulhai, R. Pringle, and G. J. Karakoulas, “Reinforcement learning for true adaptive traffic signal control,” *Journal of Transportation Engineering*, 2018.

[4] M. A. Wiering, “Multi-agent reinforcement learning for traffic light control.,” *Proceedings of the Seventeenth International Conference on Machine Learning*, 2000.

[5] M. Steingrover, R. Schouten, S. Peelen, E. Nijhuis, and B. Bakker, “Reinforcement learning of traffic light controllers adapting to traffic congestion,” 2005.

[6] X. Liang and X. Du, “Deep reinforcement learning for traffic light control in vehicular networks,” *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, 2018.

[7] T. Chu, J. Wang, L. Codeca, and Z. L. Member, “Multi-agent deep reinforcement learning for large-scale traffic signal control,” *Arxive*, 2019.

[8] M. Guo, P. Wang, C.-Y. Chan, and S. Askary, “A reinforcement learning approach for intelligent traffic signal control at urban intersections,” *Arxiv, e-print*, Arxiv:1905.07698, 2019.

[9] X. Liang, X. Du, G. Wang, and Z. Han, “Deep reinforcement learning for traffic light control in vehicular networks,” *IEEE Trans. Veh. Technol., vol. XX, no. XX,*, 2018.

Appendix

A Data Preprocessing

This part will demonstrate how we preprocessed the raw data and proffer insights about how to dispose with the raw high-frequency data into the data used in our model for the government to construct traffic lights based on their collected data.

To solve the problem, our research area is the intersection of Zhengli Road and Wuchuan Road, Zhengli Road and Guoding Road, Zhengli Road and Songhu Road and Guoding Road and Zhengmin Road, and 13 connected roads. 3 of these 13 roads connected with each other upon the four intersections, and 10 are connected to the external roads.

Our raw data is the whole month data from Shanghai Qiangsheng Taxi Company. All taxi automatically connect to the headquarters every 10 seconds to transmit data features including vehicle status, GPS coordinates, speed, and car direction. There are 3.425 billion pieces of data in the whole month, and each piece of data has 11 variables.

The final output of the data preprocessing is the traffic flow and the average speed of the 13 part of roads in the research area at different time periods, and the probabilities of vehicles turning to the four directions at each intersection.

The specific steps of data pre-processing are as follows,

1. Firstly, we filter the raw data, removing the vehicle data that does not appear in the research area. After this step we have 10.71 million pieces of data left. The heatmap drawn by the filtered data are shown in Figure 2, demonstrating the distribution of the cars at the four intersections.
2. We adjust the latitude and longitude coordinates of the original data. According to the longitude 1° is about 95 km, the latitude 1° is about 111 km when the latitude is 31°, we do unit conversion and set the Guoding Road and Zhengli Road intersection (31.31 °N, 121.4985 °E) as the origin point. The coordinates unit has been changed to meter now.
3. Next, we separate the research area into different roads and intersections. The research area is divided into 18 sections, including 4 intersections, 13 parts of roads and an “other”. We first calculate the coordinates of the four intersections, and the data within a certain range from the intersection will be classified into the intersection parts. We have also set up 13 road center-lines, and according to the width of the road, data in the range of 20 or 30 meters from the center-lines will be classified into this road part. The rest of the data will be categorized into "other" section, typically because of the positioning problem or other reasons that the coordinates of the vehicle cannot be properly classified.
4. Then we sort all the data by vehicle ID and time. Since the raw data is based on each timestamp and each vehicle, and we are going to study the movement process of the vehicle in the research area, the data based on

timestamp of the vehicle are synthesized to data based on process. We define a group of time point data that satisfies the following two conditions as a process:

- Having same vehicle ID
- After sorting by time sequence, the data between two adjacent time points data in the research area does not exceed 5 minutes

5. Next, we deal with the vehicle's moving path. We save the moving path of the vehicle in text format, so that we can get the data we need by string processing. According to the settings, if only 1-2 timestamp data in one process are classified as "other" section, while other timestamp data can be classified normally, we will only use these normal data. Then we use regular expressions to ensure that the vehicle process data is "entry—road-road crossing-road-exit" situation. We calculate and output 221,000 eligible processes.
6. Finally, according to the requirements from following models, we get three types of data,
   (a) **Traffic flow**. Traffic flow data is the number of vehicles arriving from one of the four directions of four intersections in unit of time, which is a 4×4 matrix. It can be calculated from splitting preprocessed data. The data will change over time and there will be significant differences between weekdays and weekends. It should be noted that since the data we have are taxi data from one company, which only accounts for a very small part of all vehicles, in the model we will set a multiplier to calculate the number of all vehicles in the research area.
   (b) **Vehicle turn probability**. Vehicle turn probability is the ratio of vehicles arriving from one direction of an intersection leaving from four directions. This data is four 4×4 matrices, and the sum of every row of matrix is 1, and the data does not change with time.
   (c) **Average vehicle speed**. Average vehicle speed is calculated by all vehicles arriving from one direction and leaving from another. It is also four 4×4 matrix. The average speed will not include the data when the vehicle stops.