Optimization of Multi-Intersection Traffic Signal Timing Model Based on Improved Q-Learning

Mingyue Yin1,*, Yunyun Wang2 and Ziqing Li3

1,2,3 School of Management, Wuhan University of Technology, Hubei, China

*Corresponding author e-mail: yinmingyue9840@126.com

Abstract. With the economic development and the increase of car ownership, major cities in China are facing plenty of problems, such as traffic congestion and environmental pollution caused by traffic congestion. The effective solution to solve this kind of problem is to adopt intelligent control to the traffic signal timing scheme. Based on traditional Q-learning, this paper adds the influence of phase difference and congestion factors, and uses the traffic condition data of adjacent intersections and this intersection to predict the next state, so as to adjust the timing scheme of signal lights at each intersection in the road network. The VISSIM simulation experiments show that this method can effectively reduce the average vehicle delay time and the average number of parking times, alleviate traffic congestion, and improve the traffic efficiency in the overall road network.

1. Introduction

The problem of urban traffic congestion is becoming increasingly serious in China. The average congestion delay index of Beijing, Chongqing, Harbin and other cities is greater than 2.0. At the same time, traffic jams in one direction at intersections and few vehicles in the other direction, which not only reduces the efficiency of vehicles at intersections but also wastes public resources. In order to solve these problems, we cannot all rely on expanding the road area and changing the existing infrastructure to solve the problem of traffic congestion. The most effective and resource-saving solution is to optimize the control and timing scheme of urban road traffic lights on the basis of existing facilities to ease traffic jams.

Urban road traffic signal control problems are characterized by variability, nonlinearity, and ambiguity. Traditional modeling and control methods are difficult to achieve ideal results [1]. Jianhua Tang used the improved reinforcement learning algorithm to control the multi-Agent system, and carried out a large number of simulation experiments on multiple intersections to prove that the method is robust and good scalability [2]. Pengtao Meng combined the fuzzy algorithm with the improved left-turn phase-based green waveband algorithm, and effectively reduced the delay time of vehicles at the intersection according to the actual modeling of the intersection [3].

Existing Q-learning still has the disadvantages of slower learning process and inability to respond to changing traffic flow in a timely manner in the control of regional multi-junction traffic, and traditional Q-learning often uses the means of "switching / maintaining" that is not suitable for controlling the large-scale urban traffic network.
2. Establishment of Traffic Signal Control Model for Multiple Intersections

2.1. Coordinated control

Through the coordinated control of the phase difference of the signal lights at adjacent intersections, multiple-signal linkage control is implemented to ensure that after passing the first traffic intersection, the vehicle can pass through multiple intersections without stopping during its exercise to obtain greater speed and smaller vehicle delays.

![Coordinated control system for multiple intersection signal lights.](image)

Based on the research of the existing single intersection adaptive signal timing scheme, the traffic phase information of the adjacent intersection and the intersection is used as the decision basis to adjust the phase difference. First calculate the phase difference between two adjacent intersections, then initialize the Q table based on the phase difference data, and then collect the data of the released phase traffic flow, long queues, average vehicle speed and other data of the current intersection and the adjacent intersection in the last two cycles, and the system calculation gives the individual timing scheme at this intersection, after analyzing and making decisions on the traffic conditions of adjacent intersections, select the timing scheme based on the phase difference from the Q table to adjust the original scheme. After the scheme is executed, the Q table is adjusted according to the execution effect data. Update the corresponding plan, and then repeat the above steps to time the signal lights.

2.2. Phase difference acquisition

This article selects the entry rules for the four-phase signal lights, as shown in Figure 2. Since all right-turning vehicles will not conflict with traffic in any other direction, there are no restrictions on right-turning lanes.

![Traffic rules for four-phase signal lights.](image)
The phase difference between two adjacent intersections is composed of the difference in travel time of the vehicles at the two intersections, the acceleration time of the first vehicle at the previous intersection, and the disappearance time of the stagnant vehicles at this intersection.

The distance between any two intersections is \( l \), and the expected speed of the section is \( v_0 \), \( a \) is the acceleration of the first car at the previous intersection, \( t_0 \) is the dissipation time of the queued vehicles at the intersection, and the phase difference formula of the adjacent intersection is

\[
T = \frac{l}{v_0} + \left(1 - \frac{a}{2} \right)v_0 - t_0
\]  

(1)

2.3. "Double Action" Q-learning Multi-road Signal Timing

For any traffic network with \( n \) junction, the definition \( i, j \in R \) represents two adjacent junction signal controllers, \( i \in \Gamma(j) \) and \( j \in \Gamma(i) \). \( \Gamma(i) \) is the set of intersections adjacent to \( i \). Using signal lights and intersection double action control, namely \((kla_i, a_j)\). Where \( kla_i \in \{\text{green, red}\} \) is the signal light \( k \) on the intersection \( t_i \) and \( a_j \) is the action of the intersection \( t_j \) adjacent to the passing lane of the intersection. Each intersection controller learns a function \( Q(s_k, kla_i, a_j) \) from the environment to represent the total delay time for the vehicle to reach destination \( des \) under the control of state \( s_k([kl, pos, des]) \) and double action \((kla_i, a_j)\). \( a \) represents the average delay between the vehicle in state \( b \) and the destination. represents the average delay between the vehicle in state \( s_k([kl, pos, des]) \) and the destination.

In the traffic network, the controller at any intersection continuously maintains and updates the \( Q \)-value and \( G \)-value lookup tables. The \( Q \)-value update formula is

\[
Q(s_k, kla_i, a_j) = \sum_{s'_{k_x}} P(s'_{k_x} | kla_i, a_j, s_k) [r(s_k, s'_{k_x}) + \gamma G(s'_{k_x})]
\]  

(2)

The \( G \)-value update formula is

\[
G(s_k) = \sum_{kla_i, a_j} P(kla_i, a_j | s_k) Q(s_k, kla_i, a_j)
\]  

(3)

The state transition probability \( P(s'_{k_x} | kla_i, a_j, s_k) \) indicates the probability that the vehicle will reach the next state \( s'_{k_x} \) in the state \( s_k \) and the double action \((kla_i, a_j)\). The state transition probability \( P(kla_i, a_j | s_k) \) indicates the probability that the vehicle will reach the next state \( s'_{k_x} \) in the state \( s_k \) and the double action \((kla_i, a_j)\). Based on the frequency of observations of actual vehicle conditions at intersections, the maximum likelihood probability is used for estimation and calculation.

\[
P(s'_{k_x} | kla_i, a_j, s_k) = \frac{C(s_k, kla_i, a_j, s'_{k_x})}{C(s_k, kla_i, a_j)}
\]  

(4)

\[
P(kla_i, a_j | s_k) = \frac{C(s_k, kla_i, a_j)}{C(s_k)}
\]  

(5)

Where \( C(s_k, kla_i, a_j, s'_{k_x}) \) represents the number of occurrences of the vehicle reaching the next state \( s'_{k_x} \) in the state \( s_k \) and the double action \((kla_i, a_j)\), \( C(s_k, kla_i, a_j) \) indicates the number of
times that the double action \((kla_i, a_j)\) is taken in the state \(s_k\) of the vehicle, \(C(s_k)\) means the number of times the vehicle status is \(s_k\).

\[ r(s_k, s'_k) \] represents the reward value of the vehicle moving from state \(s_k\) to state \(s'_k\) after the signal light changes,

\[
r(s_k, s'_k) = \begin{cases} 0, & s_k \neq s'_k \\ 1, & s_k = s'_k \end{cases}
\]

(6)

In the above formula, the reward value at \(s_k \neq s'_k\) is 0, which means that the vehicle can continue to move forward. When \(s_k = s'_k\) indicates that the vehicle is a red light, the vehicle needs to wait for a period, and the reward value becomes a penalty value, and the size is 1.

2.4. Congestion factor

Because the urban traffic network is complex and dense, the congestion of each lane has more or less influence on each other. In this paper, all lanes from the current location to the destination of the vehicle are numbered by \(1, 2, ..., m\), and the congestion factor corresponding to each lane is \(\omega_1, \omega_2, ..., \omega_m\). The congestion factor formula is

\[
\omega_i = \frac{V_{L_i}}{D_{L_i}}
\]

(7)

Among them, \(V_{L_i}\) is the length occupied by all vehicles in lane \(i\), and \(D_{L_i}\) is the actual length of lane \(i\), then the severity factor obtained by the vehicle is

\[
\omega = \frac{f_1\omega_1 + f_2\omega_2 + ... + f_m\omega_m}{f_1 + f_2 + ... + f_m}
\]

(8)

In the above formula, \(f_i\) is the correlation coefficient of each lane congestion factor, which reflects the degree of influence of different lanes on the vehicle's congestion factor. When the current lane congestion is more serious, the congestion factor is closer to 1. The congestion factor value is input to the main controller, which influences the controller's decision by changing the phase difference.

3. Results and Analysis

3.1. Multi-Intersection Traffic Signal Model Simulation Experiment

This article simplifies the simulation process. Only two types of buses and cars are produced, and traffic experiments are performed with the help of the VISSIM platform. The simulation was carried out according to the field survey of Luoshi Road, as shown in Figure 3.

![Simplified map of detected road sections.](Image)

Figure 3. Simplified map of detected road sections.
From the number of queuing theory and practice, survey the vehicle to reach the intersection of law, we are not reasonably assume that other external factors interfere with the vehicle, each time the vehicle reaches the intersection is random, discrete, Poisson distribution \[8\], that is,

\[ P[N(t) = x] = \frac{\lambda^x e^{-\lambda}}{x!} \quad (9) \]

Among them, \( P(x) \) is the probability of reaching \( x \) vehicles in the counting interval \( t \), and \( \lambda \) is the average arrival rate of the vehicles, that is, the number of vehicles averaged to a large intersection in a signal period, and \( \lambda < 0.4 \text{veh/s} \) is assumed according to the survey results.

If the vehicle arrives in accordance with the Poisson distribution, the headway is negatively exponentially distributed.

\[ P(h \geq t) = e^{-\alpha t} \quad (10) \]

Then

\[ t_{n+1} = t_n - \frac{\ln(\mu)}{\sigma} \quad (11) \]

In the formula, \( a \) represents the probability of reaching the headway time greater than or equal to seconds, \( b \) is a uniformly distributed random function generated by a random generator, and \( c \) is the average departure rate of the node at this road section, which represents the time when the car enters the road network.

According to the survey, the total traffic volume is 300~1000, the expected speed is 40~50 km/h, the maximum period of the signal light is 120 s, and the performance index adopts average delay time and average speed. During the simulation process, it is adjusted to be different from the actual situation. And calculate the statistics.

### 3.2. Comparative analysis of data

From the data, the original scheme (the phase difference reduction is 0s) is not the optimal scheme. Detection sections 1, 2, and 3 reduce 3s, 9s, and 5s based on the original phase difference, and have the least delay time, less parking time and number of stops, and higher average speed.

| Detection section | Phase difference reduction/s | Delay time / s | Improve effect /% | Number of stops/time | Improve effect /% |
|-------------------|-----------------------------|---------------|------------------|---------------------|------------------|
| 1                 | 0                           | 14.3          | 9.79             | 0.34                | 5.88             |
|                   | 3                           | 12.9          | 12.34            | 0.32                |                  |
| 2                 | 0                           | 23.5          | 11.53            | 0.19                | 21.05            |
| 3                 | 5                           | 16.1          | 11.53            | 0.19                | 21.05            |

The optimal phase difference reduction amount of each detection section is compared with the delay time and the number of parking times when the phase difference reduction amount is zero, and the results are shown in Table 1. It can be seen that compared with the original scheme, properly reducing the phase difference can reduce the delay time of detection sections 1, 2, 3 by 9.79%, 12.34%, and 11.53%, and reduce the number of parking times by 5.88%, 18.42%, 21.05%, thereby it can improve effectively the capacity of the transportation network.
4. Conclusion

This article establishes the improvement of Q learning based on the phase difference and the congestion factor, and regulates the timing of intersections under different traffic conditions, and then affects the decision of the timing of traffic lights in the entire road network. Experiments can effectively reduce the situation of intersections and improve vehicle traffic at intersections. And this solution is not limited to single or several traffic intersections, but starts from the city and performs overall regulation based on real-time traffic flow, which can effectively solve the problem of congestion in some sections of the city.

References

[1] C.J. Liu, Research on Coordinated Control of Multi-Intersection Traffic Signals Based on Q-Learning, Nanjing University of Posts and Telecommunications, 2018.
[2] J.H. Tang, Reinforcement learning and its application in urban traffic signal control, Xidian University, 2012.
[3] P.T. Meng, Research on Multi-Intersection Traffic Signal Cooperative Control Algorithm, Jiangnan University, 2018.
[4] A. Daechian, A. Haghani. Fuzzy Q-Learning-Based Multi-agent System for Intelligent Traffic Control by a Game Theory Approach, Arabian Journal for Science and Engineering, 43 (2018) 6-8.
[5] Y.F. Geng, C.G. Cassandras. Multi-intersection Traffic Light Control with blocking, Discrete Event Dynamic Systems, 25 (2015) 1-2.
[6] J.Y. Zhang, W.L. Xia, P.G. Research on Intelligent Dynamic Timing Model of Traffic Lights Based on Real-time Intersection, Technology and Innovation, 22 (2017) 26-28.