

**MULTI-INTERSECTION TRAFFIC CONGESTION CONTROL METHOD BASED ON REINFORCEMENT LEARNING****LIN LI, YONGHENG WANG*, SHAOFENG GENG**

College of Information Science and Engineering, Hunan University, Changsha 410082, China

E-mail:291977418@qq.com, wyh@hnu.edu.cn, 13630375@qq.com**KEYWORDS:**Reinforcement learning; Traffic control; Bilayer Control Model**ABSTRACT**

Multi-intersection traffic congestion control is always affected by two problems. The first one is multi-intersection will lead to multi-system status. The second one is the influences between adjacent intersections. Under the influence of the two problems, it is hard to control traffic congestion, therefore, in this paper, we propose the bilayer control model, the upper layer can control the road network area and the lower layer can control the area intersection in parallel in this model. Through combining upper and lower system control, the model can solve the problem of multi-system state effectively. At the same time, the model uses a message-passing mechanism to solve the problem of influences between intersections when the intersection is related to the direct connected intersection. Experiments show the bilayer control model can solve the problem of multi-intersections traffic congestion control effectively.

INTRODUCTION

With the development of society, the number of vehicles increased sharply and the traffic congestion problem attracted more and more people's attention. Expanding road will waste lots of material and manpower, besides, it can't solve the problem of traffic congestion effectively for a long time. Therefore, a effective intelligent method is needed for relieving traffic congestion problem.

At present, there is a method for controlling traffic congestion, one is the machine learning based method for single or a few intersection. However, when the number of intersection is increasing, the number of system state is also increasing, so the problem will become complicated. The influence between intersection will intensify the complexity and result in unsatisfactory result. So the problem we need to solve mainly focus on system state and influence between intersection.

Some studies have proposed methods to deal with the problem. Classification and Clustering can resolve the number of system state, but it can only control district macroscopical and ignore the control of intersection microcosmic. The influence between intersection can be resolved by inter agent control around agent, which has a risk invalidation of control agent. Transfer information between intersection can also resolve the influence between intersection, which will result in slow convergence and can not get real-time control results.

For above condition, we propose a model of layered combination control. In the upper layer, the zone node can get Vehicle distribution situation between zone by superior control. In the lower layer, intersection node realize the control of single intersection by combining upper layer shunt addition, namely. Through the combination control of two layer to implement the traffic congestion control and solve the number of system state number. At the same time, The upper and lower layer just assume that node has relation with direct connected node, the node only send

This work is supported by the National Natural Science Foundation of China (No.61371116).



messages to the directly connected nodes, Through message passing to solve the intersection problem, and because of the hierarchical reduce state quantity, avoid messaging slow convergence of the nodes.

RELATED WORKS

Fuzzy control, artificial neural networks, genetic algorithms, reinforced learning and other machine learning algorithms are often used to solve traffic congestion control problem[1]. Fuzzy control based on the traffic flow input to control green light time through a fuzzy controller, that is to say, more green time corresponding to the more vehicle. Most of the applications of Fuzzy control are used in single intersection and are rarely used in large complex transportation network[2]. Artificial neural networks are widely used in traffic control study because of its ability for modeling and learning[3]. Spall proposed an adaptive traffic control method based on artificial neural network and it can get the best time to plan traffic lights according to traffic information[4]. Xiupin proposed self-learning control system based on neural network. The system consists of two-layer neural network and is better able to adapt to the multi-phase intersection[5]. Genetic algorithm is a heuristic algorithm for stochastic optimization. It get the approximate optimal solution by simulating the natural law of survival of the fittest. Fitsum Teklu used genetic algorithm to control traffic system through optimizing routing[6]. D. Sun optimized the traffic light control to reduce the delay time by using multi-objective genetic algorithm[7]. However, when an increase in the number of intersection, neural networks and genetic algorithms are difficult to obtain the desired results. Reinforcement learning firstly obtain environmental information and system take a certain action to environment. Then environmental status change and send a reward to system. Then the system repeat the process. Finally, the system selects the maximum total reward action set. Reinforcement learning can be divided into reinforcement learning based on model and non - based model. Markov decision process is a typical reinforcement learning based model and Q-learning is a typical non model reinforcement learning[8][9][10]. They can be applied in transportation systems[11][12][13]. Compared to the model-based reinforcement learning, reinforcement learning non - based model requires fewer parameters and is more flexible. Therefore, system choose reinforcement learning non - based model as basic learning algorithm.

Reducing the number of states is achieved by clustering or hierarchical structure according to certain rules[14][15]. The intersection is merged to reduce the total number of intersection with the method mentioned above so as to achieve large-scale network control[16][17]. But after clustering, a lot of low-level information will be discarded and the accuracy of the final control result will be affected. Interaction between the junctions can use individual intersection control. But this way take no into account between the intersection, that is to say, individually optimized junction may cause adjacent intersection congestion. Center node control method is that the center node control all its adjacent intersections[18]. However, this method has a risk that once the central node failure, adjacent intersections will lose control. Road also network can be converted into a traffic map. In the map, the node represent intersection and the edge represents relationship between intersections and the Edge weight represent message value between intersections[19]. Optimization of the entire system will be converted to optimize each node in map. When the message is passed between nodes tends convergence, node is complete optimization and system get the approximate optimal solution. But when the intersection number is large, convergence slows down even real-time results may not be obtained.



BILAYER CONTROL MODEL

The method of hierarchical control mainly include two layers, each node of the lower layer is interconnected, the single node stands for single intersection. The upper layer also like this and connect with lower layer, the node of upper layer can be got from lower layer and the lower node can get control information from upper layer. The model as shown in Figure 3.1

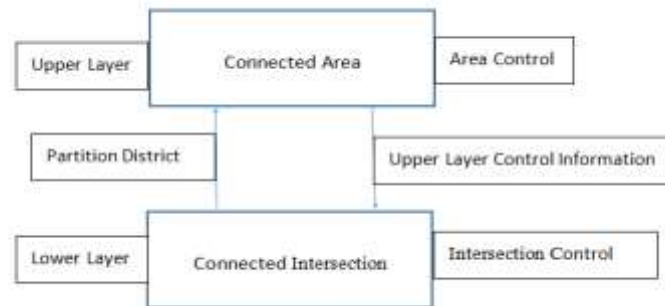


Figure 3.1: hierarchical control model

The whole model includes three steps, we will describe them in detail as below:

1) Partition District

Before dividing district we should divide the important intersection firstly, the number of intersection can be defined n , the information of all intersections can be defined $N=(num_1, num_2 \dots num_n)$, num_n is the average traffic flow of intersection n , num_n can be defined important intersection when num_n is larger than the threshold num_thred . In the lower layer, the road net is regarded as an undigraph $G_low=(V_low, E_low)$, V_low is a node and stands for intersection, E_low is ligature between nodes and stands for distance. We can use two-dimensional matrix M to describe the road net. M has n lines and n columns, $M_{ij}(0 \leq i \leq n-1, 0 \leq j \leq n-1)$ is the distance between intersection i and j . Then we will set a threshold named $r_threshold$ and center on the important intersection, set $r_threshold$ as radius, the intersection in this section will be merged into one section. Besides, we can set a threshold named $rr_threshold$ if the important intersection is existed in this section, when the important intersection is existed in the section which centered on important intersection and regarded $rr_threshold$ as radius, the two intersection can be combined. We can abandon the intersection when the intersection is not in this section that is centered on important road and regarded $r_threshold$ as radius. If intersection located in the section that centered on intersection and regarded $r_threshold$ as radius, we can confirm the district according to the serial number of important intersection, once the intersection is divided into one district, it will not divided into other new district.

2) Upper Layer Control

After getting district structure, we can get the shunt into information between districts. We define upper layer district as undigraph $G_high=(V_high, E_high)$, the node V_high is district, E_high is the relation between districts. $u_{ij}(a_j)$ is the value of E_high , i and j are the node of undigraph, a is the action between node, a is belong to A , $A=(V_{src}, V_{des}, Rate)$ stands for the vehicle of Rate number that V_{src} sent to V_{des} . $u_{ij}(a_j)$ is local optimization information that i sent to j , and the influence between node will be considered by u_{ij} . Due to we consider node just



has relation with direct connected node, the total reward of system can turn into the sum of reward of every node, the whole system will reach optimal when every node reach optimal. Then the optimization of the whole system will be converted to the optimization of each node. As a result, the news u_{ij} between the node will be very important, when the relay messages between the nodes tend to be convergent, nodes will reach the optimal state, the whole system get the approximate optimal state. At this point, the corresponding action, that is, the optimal control action, namely the optimal regional shunt direction. Due to fewer area number, the system can guarantee convergence within certain steps. If area number is large, you can set the same limited steps, this system can also be obtained after the end of the approximate solutions. So the whole process of upper control is the messages between nodes, when the message tend to convergent, we get the optimal control action, namely the optimal area bypass action.

3) Lower Layer Control

After getting the upper layer control area shunting action, the key point is to deal with the single intersection within district according to the upper control action. In order to considering the mutual influence between intersections, we assume that the mutual influence between intersection only has relation with direct connected intersection, the specific influence relationship is introduced in upper control. Combined with the reinforcement learning content that the second section introduced, the single control of intersection we use the reinforcement learning model. Defining intersection information as $R = (S, A_light, A_High, M, R)$, In this formula, S represent the vehicle number of intersection, $A_light = (L_phase, L_time)$, L_phase is phase, L_time represents the green time, A_High represents the shunt direction of upper control, according to A_High , it will affect green time of every phase, $M = (m1, m2, mn)$ represents information that n adjacent intersections sent to the current intersection, R is the reward of intersections. In the end, when the transfer message between intersection tends to convergence, R will get approximate maximum, now A_light is the optimal time of traffic lights control. In addition, because of the optimal shunting action has been got in different regions, so optimization in each area can be executed in parallel and reduce the total optimization time of system.

At last, the whole process is shown in algorithm 1 as below. In this algorithm i_num is average traffic flow, $num_threshold$ is the threshold. The process of this algorithm is the three steps which is introduced.

Algorithm 1. Hierarchical Control Algorithm

Input: I is the intersection in the lower layer

Output: A_area is the area control action, A_isec is the intersection control action

Begin

```

for  $i \in I$ 
  if  $i\_num \geq num\_threshold$ 
     $I\_important \leftarrow i$ 
  end for
for  $i \in I\_important$ 
  do  $I\_area \leftarrow$  partition district
end for
for  $i \in I\_area$ 

```



```

while message is not convergent
    do reinforcement learning
    A_area←strategy
end for
for i∈ I_area
    for j∈ i
        while message is not convergent
            do reinforcement learning
            A_isec←strategy
        End for
    End for
End

```

4)summary

Through the study of the upper layer, the combined area get optimal distribution control action according to the global situation, achieve global optimal; Through the study of the under layer, on the one hand, a single intersection realizes own local optimum, on the other hand, due to the combination of shunt processing action for the first time reinforcement learning, and also meets the conditions of the global optimum, and take full advantage of all the collected environmental information. Finally, when crossing the rise in the numbers, general machine learning methods can't solve, through layered combination and intersection traffic control can be resolved effectively optimize, and considering the problem of the mutual influence between nodes to achieve global optimization.

EXPERIMENTAL

4.1 Simulation tools

For the study of traffic congestion control, directly in the actual environment traffic light control is unrealistic, so the use of simulation tools for simulation and control traffic. Green Light District is a microscopic traffic simulator , it is mainly divided into two functions, one is network editing, the other is traffic control. When network editing, network mainly consist of roads and motor vehicles, both sides of the road is composed of two kinds of nodes, one kind is a crossroad, another kind is edge node, the function of edge node is to produce different types of vehicles according to the different frequencies, the motion of the vehicle speed and driving route is setting before the experimental running the simulation. In the traffic control function, mainly is the traffic light control, such as fixed time control, fuzzy control, etc., for traffic simulation control mainly through the control of the traffic lights. The road network as shown in Figure 4.1, which can set different frequency of edge nodes to realize the different intersection size and to achieve the comparison result of different traffic control algorithms through the output evaluation parameters .

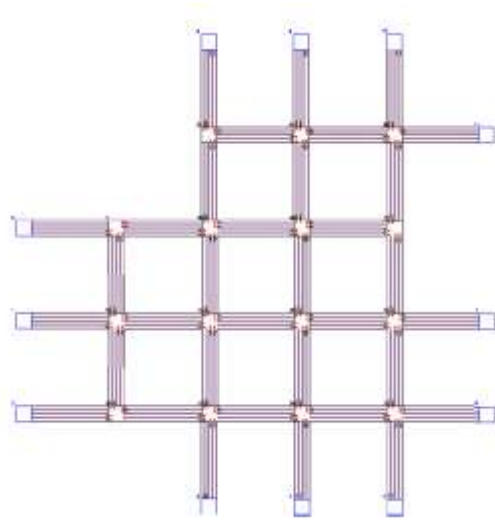


Figure 4.1: the traffic road network of 15 intersections

4.2 Road markers and merge

First of all, intersections can be divided into big crossroads and small crossroads according to the intersection traffic, then make the small intersections within 10km near the big intersections into the center of the intersection to get a large area. Crossroads take place layered combination control after the merger of region.

4.3 The upper layer learning

Status: On a single road, road section has its saturation of vehicles, for the combined area also has its saturation of vehicles. Suppose there are a combined area to withstand the number of vehicles more than 5000 would be congestion, the upper limit is 10,000 vehicles, so when there are 6000 vehicles, congestion level is 20%, when there are 7000 vehicles, congestion level is 40%, when there are 10000 vehicles, congestion level is 100%. Under the assumption that the non-congested state is 0, each more than 10% is set to a state, percentage of the molecules are in less than 10 by rounding operation, then all the individual state $s_n = (s_0, s_1, \dots, s_n)$, where $n = 0, 1, \dots, 10$, s_n represents the total number of vehicles over the standard $10 * n * \%$. Assuming the number of combined road is k , then the total state $S_k = (S_1, S_2, \dots, S_k)$, where S_k represents the k -th state of intersections after the merger, and is one of the s_n .

Action: The upper layer learning mainly for the vehicle shunt between the intersections, the amount of vehicles of current combined area are assigned to the connection area. Assuming the total number of vehicles in the k -th combined area is Total Number, shunt operation is setting at 0.5% of the total number of vehicles each time, it is divided into 11 gears, that is $A_n = (A_0, A_1, \dots, A_n)$, the A_n represents to its connection crossroads shunt Total Number * 0.5 * $n * \%$ cars.

Remuneration: Total compensation is defined as the sum of the vehicle saturation of each combined region, because the vehicle saturation is directly related to the degree of congestion, when the total vehicle saturation is minimal, the degree of congestion is minimized. Therefore, when the vehicle saturation calculation reaches convergence, the system achieves the global optimal.



Strategy: The corresponding shunt operation when the total vehicle saturation of system is minimum.

4.4 The under layer learning

Status: Each section of vehicle saturation, assuming a 500 meters long driveway more than 50 vehicles would be congestion, up to 100 vehicles. So when there are 60 vehicles, congestion level is 20%, when there are 100 vehicles, congestion level is 100%. Under the assumption that the normal state is 0, each more than 10% is set to a state, percentage of the molecules are in less than 10 by rounding operation, then all state corresponds to $s_n = (s_0, s_1, \dots, s_n)$, where $n = 0, 1, \dots, 10$, s_n represents the total number of vehicles over the standard $10 * n \%$, since each intersection are optimized separately, so the status of each intersection is indicated separately.

Action: Intersection traffic lights control time, mainly green light control time L , $L = L_{base} + L_{now} + L_{one}$, L_{base} guarantees intersections with a minimum through time, that is when a driveway rarely such as when a vehicle can also be in a certain time period through, avoid excessive waiting for long time. L_{now} represents the current green time of intersection, L_{now} is an integer multiple of 2s, divided into 10 levels, namely a minimum of 2s, up to 20s, L_{one} is decided by the action of reinforcement learning for the first time, for example, split by 0.5 percentage points to extend 5s, then split $n * 0.5$ percentage points, the green time extension $5 * n$ seconds.

Remuneration: The current traffic congestion of intersections is directly related to the traffic situation, the remuneration is set to the degree of congestion, it requests system directly optimize the congestion.

Strategy: The total remuneration is optimal that is the corresponding green light control time when the degree of congestion is minimum, because on the one hand, ensure the optimization of intersection here, and in the movement process to take into account the global optimization, the current traffic light control is the actual control action.

4.5 Experimental results

As shown in Figure 4.2, the horizontal axis represents time iteration number, the vertical axis represents the intersection average waiting time, the waiting time is increasing when system started. With the passage of time, waiting time reduced and tended to balance, it is consistent with results during regional convergence algorithm mentioned. The experimental results show that bilayer congestion control method for controlling a large intersection has a certain effect.

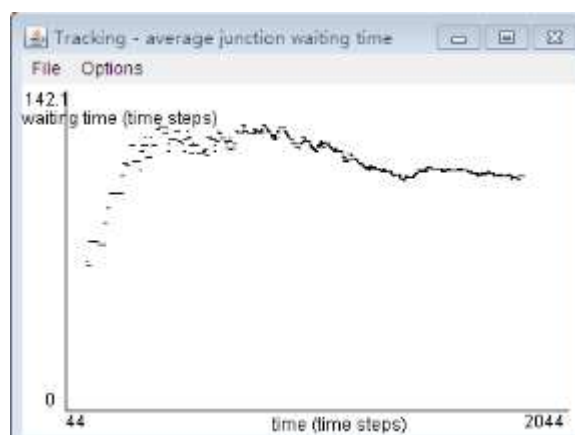




Figure 4.2: experimental result

CONCLUSION

In order to solve the problem of congestion control in multiple crossroads, this paper presents a bilayer control model to emphatically solving the intractable problem and coordination problems between intersections due to the number of intersections. In bilayer control model, the number of total regions in the upper layer will reduce through the combination of lower layer information. At the same time, each region in the lower layer can realize independent control, which avoids the problem of quantity and make full use of the underlying collection information. Coordination problems between intersections were affected by crossing number, coordination problems between intersections will reduce complexity when the number reduced. Only by assuming its intersection intersection directly connected and related messaging mechanism tore solve the problem of coordination between intersections. Through the experiment we found that the double combination control of intersection traffic have certain effect.

Because of the control of double combination related to each individual control as well as the communication of upper and lower layers, in this paper, the lower used the results of the upper and ignored the communication problems in a timely manner between the upper and the lower. Therefore, in the future work, we should consider the timely communication problems between layer and layer more and the monolayer details problem of control.

ACKNOWLEDGE

First, this work is supported by the National Natural Science Foundation of China (No.61371116).

Second, I want to appreciate my supervisor, Yongheng Wang, who gives me precious suggestions.

I also appreciate my classmate who help me a lot.

Finally, I want to express my gratitude to my family, thanks their supporting.

REFERENCES

- [1]Araghi S, Khosravi A, Creighton D. A review on computational intelligence methods for controlling traffic signal timing[J]. Expert Systems with Applications, 2015, 42(3): 1538-1550.
- [2]Gokulan B P, Srinivasan D. Distributed geometric fuzzy multiagent urban traffic signal control[J]. Intelligent Transportation Systems, IEEE Transactions on, 2010, 11(3): 714-727.
- [3]Heinen M R, Bazzan A L C, Engel P M. Dealing with continuous-state reinforcement learning for intelligent control of traffic signals[C]//Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on. IEEE, 2011: 890-895.
- [4]Chin D C, Spall J C, Smith R H. Evaluation of system-wide traffic signal control using stochastic optimization and neural networks[C]//American Control Conference, 1999. Proceedings of the 1999. IEEE, 1999, 3: 2188-2194.
- [5]XU Q, YANG R, HE N. A Traffic Signal Control Method for Isolated Intersection Applying Fuzzy Neural Network [J][J]. Journal of Huaihai Institute of Technology, 2004, 1: 006.
- [6]Teklu F, Sumalee A, Watling D. A genetic algorithm approach for optimizing traffic control signals considering



- routing[J]. Computer - Aided Civil and Infrastructure Engineering, 2007, 22(1): 31-43.
- [7]Sun D, Benekohal R F, Waller S T. Multiobjective traffic signal timing optimization using non-dominated sorting genetic algorithm[C]//Intelligent Vehicles Symposium, 2003. Proceedings. IEEE. IEEE, 2003: 198-203.
- [8]Haijema R, van der Wal J. An MDP decomposition approach for traffic control at isolated signalized intersections[J]. Probability in the Engineering and Informational Sciences, 2008, 22(04): 587-602.
- [9]Prashanth L A, Bhatnagar S. Reinforcement learning with function approximation for traffic signal control[J]. Intelligent Transportation Systems, IEEE Transactions on, 2011, 12(2): 412-421.
- [10]Pham T T, Brys T, Taylor M E, et al. Learning coordinated traffic light control[C]//Proceedings of the Adaptive and Learning Agents workshop (at AAMAS-13). 2013, 10: 1196-1201.
- [11]Abdoos M, Mozayani N, Bazzan A L C. Traffic light control in non-stationary environments based on multi agent q-learning[C]//Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on. IEEE, 2011: 1580-1585.
- [12]Prashanth L A, Bhatnagar S. Reinforcement learning with average cost for adaptive control of traffic lights at intersections[C]//Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on. IEEE, 2011: 1640-1645.
- [13]Yaping W, Zheng Z. A method of reinforcement learning based automatic traffic signal control[C]//Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on. IEEE, 2011, 1: 119-122.
- [14]Bazzan A L C, de Oliveira D, da Silva B C. Learning in groups of traffic signals[J]. Engineering Applications of Artificial Intelligence, 2010, 23(4): 560-568.
- [15]Abdoos M, Mozayani N, Bazzan A L C. Holonic multi-agent system for traffic signals control[J]. Engineering Applications of Artificial Intelligence, 2013, 26(5): 1575-1587.
- [16]Prashanth L A, Bhatnagar S. Threshold tuning using stochastic optimization for graded signal control[J]. Vehicular Technology, IEEE Transactions on, 2012, 61(9): 3865-3880.
- [17]Abdoos M, Mozayani N, Bazzan A L C. Hierarchical control of traffic signals using Q-learning with tile coding[J]. Applied intelligence, 2014, 40(2): 201-213.
- [18]Arel I, Liu C, Urbanik T, et al. Reinforcement learning-based multi-agent system for network traffic signal control[J]. Intelligent Transport Systems, IET, 2010, 4(2): 128-135.
- [19]Kuyer L, Bakker B, Whiteson S. Multiagent reinforcement learning and coordination for urban traffic control using coordination graphs and max-plus[D]. MS Thesis, 2007.