

A fast method to prevent traffic blockage by signal control based on reinforcement learning

Meng-Jia Shen

Lanzhou University, Lanzhou, Gansu, China
Email: shenmj13@lzu.edu.cn

In this paper, we present an efficient and fast way to prevent traffic blockage by controlling traffic signal. A new model is adopted to our program which is fused and developed by probabilistic model and cellular automatic model (CA model). Based on this model, we used wavelet neural network (WNN) for predicting traffic flow and use this data to improve the green or red light time setting. Q-learning algorithm, as one of the reinforcement learning methods, also is applied to project the phase of traffic light from different intersections. Considering the requirement of model response rate, we also develop this algorithm for a fast respond to the complex and fickle traffic condition. Finally, a simulation study is carried out to evaluate our method and the result shows that this method can avoid traffic blockage efficiently.

Keywords: Q-Learning Algorithm; Traffic Modeling; Wavelet Neural Network; Traffic Signal Control.

1. Introduction

Nowadays, the growth of vertical and complexity of city traffic network lead to an increasing of traffic congestion. How to prevent traffic blockage has become a stern problem in contemporary society. In this paper, we want to find out an efficient way to work out this problem. Considering the complex and fickle traffic condition, we need to find a method which has a fast respond to the current situation.

Congestion problem has been studied for a long time and traffic signal control for intersection is one of the most popular issues in this area. Christofa monitored the vehicle queue status and the movement of the end-of-queue using connected vehicle information and IntelliDrive data[1]. Agbolosu Amison proposed a dynamic signal optimization method by detecting the gap between clusters of vehicles[2]. He used online data to detect vehicles platoons and then optimized the signal using mixed integer linear program[3]. Smith discussed this problem and compared various strategies[4,5].

In this paper, we proposed a method to prevent traffic blockage by planning the setting of traffic signal. A improved model is proposed and applied in our research. Q-learning algorithm is an important part in our method, and we adapt this reinforcement learning method to setting signal between different intersections. WNN model is used to predict traffic flow data and makes the learning algorithm have a better performance.

2. Method

The proposed method can be described by follow chart Figure 1.

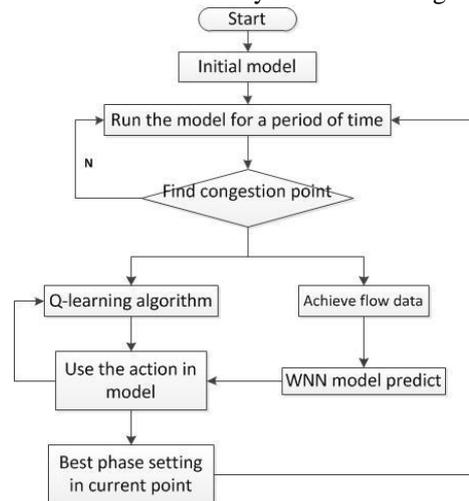


Fig. 1 flow chart of proposed method

2.1. Transportation model

There are many classical but quite complex traffic models and all of them want to find out a most accuracy way to describe intersections in traffic network. While, intersection is complicated and can be easily affect by numerous unstable and random factors. We believe the best solution for this kind of problem needs to based on systematology viewpoints. It is no need to calculate the traffic flow in intersections exactly. In this model, we seen the intersections as a system, a black box. Each car drives into the intersection has a probability to put out toward some roads connected to the intersection. It is easy and convenient to convert the tough traffic model problem into a distribution problem[6].

In addition, we applied CA model to simulate the traffic condition of the road in city which is widely used in this field[7]. All of the multi-lanes in city can be separated into several ones-way lanes in different direction. Every lane has a

input signal and a output signal. Considering that there may some department, school or some crowded place by the road side, the input number of cars may not equal to the output in a lane. So, in this model, every lane has a parameter to describe the probability of some vehicles which are not counted by the statistic equipment of the traffic flow in intersection enter into this section. In this model, there are four possible phase in each road which are shown as follow

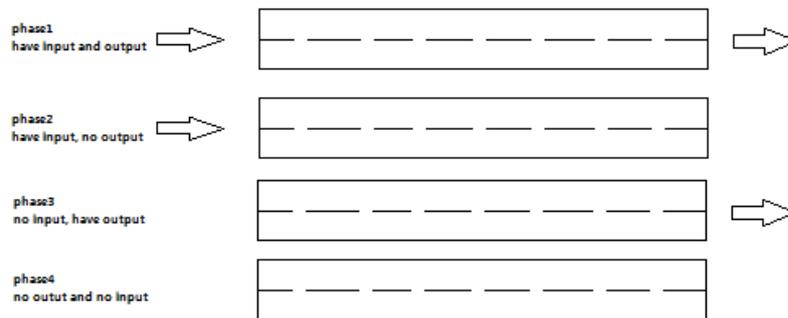


Fig. 2 example of phases in a lane

Phase 1 and phase 3 are able to decrease the dwell time of cars in current section, and phase 2 and phase 4 can not. Apparently, Phase 3 and phase 1 are better choice to ease the traffic pressure and phase 3 can do better than phase 1 as well. So, in projecting of traffic signal in a congestion, we need to facilitate the forming of phase 1 and phase 3.

2.2. Predict model by WNN

Predicting traffic flow data is an important part in signal control and signal light time sitting. If we are able to predict that the flow data would have an increase in the near future, some readjustment could be make up to prevent blockage. After getting flow data from the transportation model, we applied this data into WNN predict model. WNN is a combination of classic neural network and the wavelet analysis [8]. It has two parts: model selection and variable significance testing [8]. WNN has three layers in structure and wavelet activation functions are in hidden layer instead of traditional activation functions. WNN structure is showed in Figure 4, with output $F(x)$, inputs $p\{x_1, x_2, \dots, x_p\}$ and m number of mother wavelet [9].

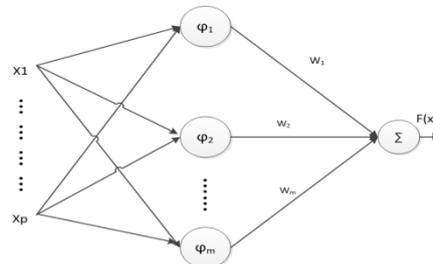


Fig. 3 WNN model

The output of a standard form WNN is given by: [9]

$$F(x) = \sum_{i=1}^M w_i \Psi_i(x) = \sum_{i=1}^M w_i |a_i|^{-1/2} \Psi((x-b_i)/a_i) \tag{1}$$

In *i*th unit of the hidden layer, Ψ_i is the wavelet activation function and w_i is the weight connecting the *i*th unit of the hidden layer to the output layer unit. Note that for the *n*-dimensional input space, the multivariate wavelet basis function can be calculated by the tensor product of *n* single wavelet basis functions as follows. Back-propagation used in training the WNN. [9]

$$\Psi(x) = \prod_{i=1}^n \Psi(x_i) \tag{2}$$

2.3. Reinforcement learning

Q-learning algorithm is a model-free learning method in reinforcement learning. It can find out an optimal action-selection policy for given Markov decision process (MDP) [10,11,12]. The Q-learning algorithm works as follow:

Set parameters and initialize reward and Q matrix

For each episode:

Select a random initial state and do while the goal state has not been reached.

2.1 Select and use one action among all possible actions, consider going to the next state.

2.2 Get maximum Q value for this next state based on all possible actions.

2.3 Compute: $Q(\text{state}, \text{action}) = R(\text{state}, \text{action}) + \text{Gamma} * \text{Max}[Q(\text{next state}, \text{all actions})]$

2.4 Set the next state as the current state

End Do

End For

It is hard and slow to calculate the reward function in the whole traffic model. We improved the reward function and focus on the area which already has or would have a blockage. In the above model, if a road from the congestion area is in a right traffic signal phase (phase 1 and phase 3), it will have a good traffic appease ability and be beneficial to the local traffic condition. In another

word, in a specific road, the time of each phase can express the improvement degree to the congestion in this area. So, the better setting of traffic phase, the better effect to prevent traffic jam. While, different roads have different influence to the transport. A parameter I is needed to present this kind of imparity between various roads. The reward function R is defined as follow. Function P is the score of traffic signal phase setting. V is the average speed of current road.

$$R=[P(\text{phase,time})+V]*I \tag{3}$$

3. Experiment

3.1. WNN model test

In experiment, the test data is from our traffic model which is initialized by random and run a period of time. Every 4 hours in simulation time, we record 30 data value (during 30 minute in simulation time) as our test data. The WNN model has 2 input mode, 2 output mode and 12 hidden layers. The relationship between test data and prediction shows as follow. Horizontal axis shows the time in each simulation and the vertical axis shows the number of traffic flow data of a road in model. Model shows a good performance in predicting and the result shows that the maximum error rate is 6.8%.

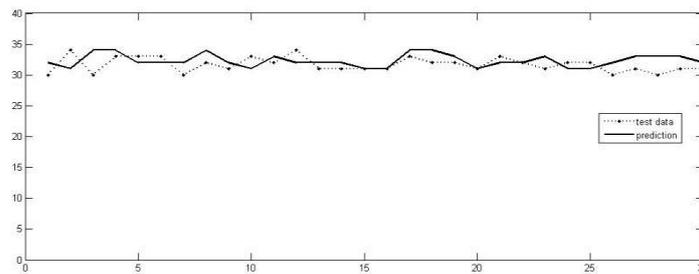


Fig.4 Test data and prediction Figure-1

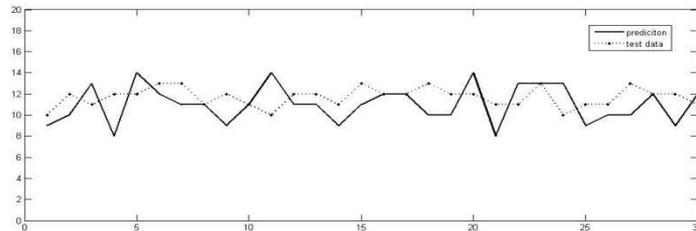


Fig.5 Test data and prediction Figure-2

3.2. Method test

The method experiment result is shown as follow. Horizontal axis shows the time in each simulation and the vertical axis shows the average delay time. It can be found that, due to the improvement of reward function of Q-learning, our method has a faster respond than the Q-learning function with a reward function of average delay time. Meanwhile, the blockage condition has been efficiently prevented.

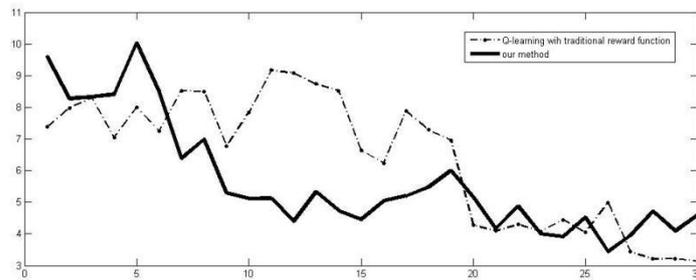


Fig.6 Method Experiment Result Figure

4. Conclusion

In this paper, we proposed a method to prevent traffic blockage based on reinforcement learning. A improved model was used in our research and WNN model was used to make a prediction of the flow in our traffic model. This method show an efficient way to work out traffic congestion in given model, but we ignored many factors in city traffic system. There are many aspects need to have a future study.

Acknowledgment

This work is supported by research funds for National Training Programs of innovation and Entrepreneurship for Undergraduates.(Lanzhou University, project number: 201610730084).

References

1. E.Christofa, J.Argote, and A.Skabardonis, Arterial queue spillback detection and signal control based on connected vehicle technology *Transp. Res. Rec.*, *J.Transp. Res. Board*, vol. 2356,pp. 61-70, Nov.2013.
2. S.J.Agolosu-Amison, B.Park, and I.Yun, Comparative evaluation of heuristic optimization methods in urban arterial network optimization. in *Proc. 12th Int. IEEE Conf. Intell. Syst. (ITSC)*, Oct.2009, pp.1-6.
3. Q.He, K.L.Head, and J.Ding, PAMSCOD:platoon-based arterial multi-model signal control with online data *Transp. Res. C, EmergTechnol.*, Vol.20,pp.163-184,Feb.2012.
4. B.L.Smith, R.Venkatanarayana, H.Park, N.Goodall, J.Datesh, and C.Skerrit, IntelliDrive traffic signal control algorithms *Center Transp. Studies, Univ. Virginia, Charlottesville, VA, USA, Tech. Rep. Final Report*,2010.
5. W. Sun, X. Wu, Y. Wang, and G. Yu, A continuous-flow-intersectionlite design and traffic control for oversaturated bottleneck intersections, *Transp. Res. C, Emerg. Technol.*, vol. 56, pp.18-33, Jul. 2015.
6. N. V. Findler and J. Stapp, Distributed approach to optimized control of street traffic signals, *J. Transp. Eng.*, vol. 118, no. 1, pp. 99-110, Jan./Feb. 1992.
7. M. Maher, A comparison of the use of the cell transmission and platoon dispersion models in TRANSYT 13, *Transp. Planning Technol.*, vol. 34,no. 1, pp. 71-85, Feb. 2011.
8. Antonios K.Alexandridis, Achilleas D.Zapranis. Wavelet neural networks: A practical guide. Preprint Published in *Neural Networks*, 42,pp.1-27,2013.
9. Mohammed Awad. Chaotic Time series Prediction using Wavelet Neural Networks *Journal of Artificial Intelligence: Theory and Application(VOL.1-2010/Iss.3)* Awad/Chaotic Time series Prediction Wavelet Neural Network/pp.73-80.
10. <https://en.wikipedia.org/wiki/Q-learning>.
11. M.Abdoos, N.Mozyano, and A.L.C. Bazzan. Traffic light control in non-stationary environments based on multi-agent q-learning. In *intelligent*

- Transportation System(ITSC), 2011 14th international IEEE conference on, pages 1580-1585. IEEE, 2011.
12. J.C.Medina and R.F.Benekohal. Traffic signal control using reinforcement learning and the max-plus algorithm as a coordinating strategy. In Intelligent Transportation Systems(ITSC), 2012 15th International IEEE Conference on ,pages 596-601.IEEE,2012.