Journal of Southeast University (English Edition) Vol. 27, No. 3, pp. 311 – 315 Sept. 2011 ISSN 1003—7985

A distributed algorithm for signal coordination of multiple agents with embedded platoon dispersion model

Bie Yiming[1] Wang Dianhai[2] Ma Dongfang[1] Zhu Zibo[1]

([1] College of Transportation, Jilin University, Changchun 130025, China)

([2] College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China)

Abstract: In order to reduce average arterial vehicle delay, a novel distributed and coordinated traffic control algorithm is developed using the multiple agent system and the reinforce learning (RL). The RL is used to minimize average delay of arterial vehicles by training the interaction ability between agents and exterior environments. The Robertson platoon dispersion model is embedded in the RL algorithm to precisely predict platoon movements on arteries and then the reward function is developed based on the dispersion model and delay equations cited by HCM2000. The performance of the algorithm is evaluated in a Matlab environment and comparisons between the algorithm and the conventional coordination algorithm are conducted in three different traffic load scenarios. Results show that the proposed algorithm outperforms the conventional algorithm in all the scenarios. Moreover, with the increase in saturation degree, the performance is improved more significantly. The results verify the feasibility and efficiency of the established algorithm.

Key words: multiple agents; signal coordination; reinforce learning; platoon dispersion model

doi: 10.3969/j.issn.1003 – 7985.2011.03.017

As indicated in the newly released 2007 National Traffic Signal Report Card, an optimally operated traffic signal can reduce traffic delays by 15% to 40%, fuel consumption by up to 10%, and harmful emissions by up to 22%[1]. The signal coordination algorithm has become one essential component of traffic control systems. There has been a considerable amount of work on the subject of signal coordination, and many valuable results have been achieved. In general, these works can be broadly categorized into two groups: 1) Signal coordination from the perspective of conventional traffic approaches. Studies in this category employ traffic flow models to predict the platoon arrival pattern at each intersection and then develop a function among performance indices (PI) such as vehicle delay, stops, queue length and timing parameters[2-3]. The offsets among signals in the coordination subarea are optimized to obtain optimal PI. Some well-known adaptive control systems such as SCOOT and SCATS belong to this category[4]. 2) Signal coordination from the perspective of artificial intelligence. With the advent of artificial intelligence (AI), some researchers attempted to apply the AI technique to the traffic control field and also many results have been achieved. Studies in this category emphasize the similarities between AI and traffic control and solve the problem in the traffic domain using the AI technique, such as multiple agents, neural networks and fuzzy logic[5-9].

Due to the nonlinear and stochastic traffic processes in the network, it is difficult to precisely develop mathematical models for traffic control. However, some artificial intelligence methods take advantages in dealing with the nonlinear process, so recently many researchers have applied various AI techniques to develop better traffic signal control methods.

Over the past few years, multiple agent systems (MAS) have become a crucial technology for effectively exploiting the increasing availability of diverse and distributed information sources. Researchers adopt numerous techniques and use various tools to implement multi-agent systems for their problem domains. Also, some researchers apply the MAS to the traffic domain[10-12]. For example, Ma et al.[9] developed a signal coordination method based on the MAS technique, and game theory and social rules are applied to solve the coordination between two adjacent intersections. Oliveira and Camponogara[11] proposed a framework for a network of distributed agents to control nonlinear traffic systems. The framework decomposed the optimization problem into small subproblems to be solved by the agent network. Each agent sensed and controlled the variables of its subsystems, while communicating with agents in the vicinity to obtain neighborhood variables and coordinate their actions. In Ref. [12], multi-agent reinforce learning (RL) was applied to coordinate traffic signals at six intersections by constructing a vehicle-based model. Given different settings of traffic lights, the RL system estimates expected waiting time for cars.

Present studies mainly focus on developing architectures for the control system and establishing the relationship between MAS and traffic control. In fact, the communications between adjacent agents are determined by the traffic platoon that departs from the upstream stop line and moves towards a downstream signal. However, current studies ignore this essential feature of arterial traffic flow and the communications are not treated in detail. In the authors' opinion, basic traffic characteristics should be paid more attention to and only in such a way can traffic problems be solved. The main objective of this research is to address the weaknesses of current studies and develop a new signal coordination algorithm based on the MAS technique.

Received 2011-04-22.

Biographies: Bie Yiming (1986—), male, graduate; Wang Dianhai (corresponding author), male, professor, wangdianhai@sohu.com.

Foundation items: The National Key Technology R&D Program during the 11th Five-Year Plan Period of China (No. 2009BAG17B02), the National High Technology Research and Development Program of China (863 Program) (No. 2011AA110304), the National Natural Science Foundation of China (No. 50908100).

Citation: Bie Yiming, Wang Dianhai, Ma Dongfang, et al. A distributed algorithm for signal coordination of multiple agents with embedded platoon dispersion model[J]. Journal of Southeast University (English Edition), 2011, 27(3): 311 – 315. [doi: 10.3969/j.issn.1003 – 7985.2011.03.017]

1 System Model

1.1 Architecture of distributed signal coordination system

Take one subarea that includes three signalized intersections as an example. The system is composed of two levels of agents: local agent and central agent. Each level of agent contains different modules and bases. The architecture of the system is shown in Fig. 1.

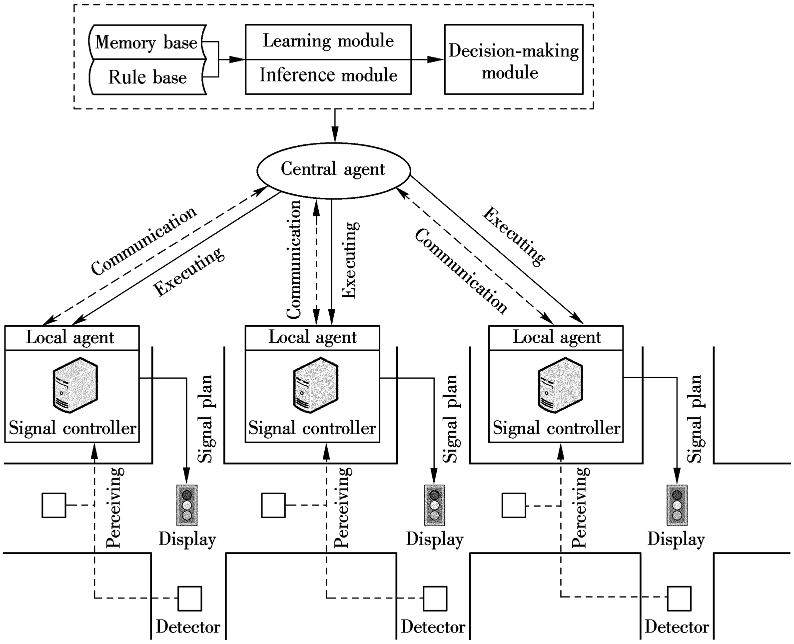

Fig. 1 Architecture of the distributed signal coordination system

1.1.1 Local agent

An isolated intersection in the subarea is designed as a local agent. It only carries out some simple functions such as detecting traffic flow information by local detectors and executing timing schemes that are assigned by the central agent. It includes the following two basic modules.

- **Perceiving module** The system perceives the dynamic changes in traffic conditions depending on loop detectors located on approach lanes of a local intersection. The detectors detect the traffic volume, vehicle speeds and occupation rates data and send these data to a signal controller.
- **Executing module** The central agent assigns optimal coordination schemes to the local agents. The local agents execute these schemes by dynamically changing the display of signal lights according to the schemes.

1.1.2 Central agent

The critical signal controller in the subarea or the signal control center is designated as the central agent. The central agent is the core of the distributed system and functions such as calculating optimal timing schemes for local agents, determining timing schemes (common cycle length, green split, offsets) for a subarea, and scheme assignments to local agents are realized by it. It includes the following modules:

- **Communication module** This module refers to the communication links of the system. It can upload traffic data obtained by a perceiving module to the central agent and download timing schemes from the central agent to local agents.
- **Rule base** The rule base stores general signal control information, such as signal control rules, link lengths, the number of approaches, saturation flow rates and the number of adjacent intersections.
- **Memory base** The memory base stores the historical traffic flow information and timing schemes and updates these data dynamically.
- **Learning module** This module learns from both traffic flow information and evaluation information and then obtains quantitative data, which can provide a theoretical foundation for the inference module.
- **Inference module** Based on the rule base of agents, this module estimates the dynamic changes in system states according to exterior information provided by a perceiving module. Then the inference module infers the optimal strategy for the distributed signal coordination system.
- **Decision-making module** Based on the information provided by the inference module, this module produces a set of optimal strategies. It also determines the optimal download occasion of timing schemes.

1.2 Definition of the RL elements

An RL problem refers to system states, actions and rewards.

1.2.1 System states

A mixed state representation that combines information at the controlled intersection is used. The general form of the equation that defines the state is as follows:

$$\boldsymbol{S}^k = \{X^k, V^k\} \tag{1}$$

$$\boldsymbol{X}^k = \{x_1^k, x_2^k, \ldots, x_n^k\} \tag{2}$$

$$\boldsymbol{V}^k = \{v_1^k, v_2^k, \ldots, v_n^k\} \tag{3}$$

where $\boldsymbol{S}^k$ is the system state in time step k; $\boldsymbol{X}^k$ is the saturation degree vector; x_i^k is the saturation degree of intersection i; n is the number of intersections in a subarea; $\boldsymbol{V}^k$ is the volume vector; v_i^k is the volume of the critical lane of the coordination phase in intersection i.

The reason for choosing these two factors is mainly because the saturation degree can reflect the traffic load of each intersection and is closely related to traffic volume, queue length, vehicle delay and speed. These performance indices refer to the whole state of the system. The volume of the critical lane of the coordination phase is related to the arriving platoon size and the assigned green time for the coordination phase, and both of them can affect coordination benefits.

1.2.2 Action set

The system takes an action each time step to reduce average vehicle delay in the subarea. Cycle length and offset are adjusted and actions include two aspects. One is to increase the common cycle length by 3 s, keep the current common cycle length, and decrease the common cycle length by 3 s. So the change set of the cycle length at each decision step is [+3, 0, −3]. The reason for choosing 3 s as the unit step length is that for large vehicles 2 s is less and 4 s is superfluous. The other one is to increase the offset by 3 s, keep the current offset, and decrease the offset by 3 s. So the change set of offset at each decision step is also [+3, 0, −3].

1.2.3 Reward function

The reward function is the most important component of the RL algorithm. Q learning, one of the most popular RL algorithms, is employed. Average vehicle delay in the subarea is used as a value function.

At isolated intersections, vehicles are usually regarded as arriving randomly. However, at arterial intersections, vehicles are affected by the signal lights and arrival rules of isolated intersections are not suitable for coordinated intersections. On an arterial link, vehicles are grouped into a platoon and one obvious characteristic of them is platoon dispersion. The Robertson dispersion model is usually used to depict the dispersion level and the model is shown as[13]

$$v_d(p) = \sum_{k=1}^{p-t} v_0(k) F(1-F)^{p-t-k} \tag{4}$$

where $F = 1/(1+0.35t)$ is the platoon dispersion factor; t is 0.8 times the travel time between two observation points in a unit time interval; $v_d(p)$ is the number of arriving vehicles at the downstream observation point during time interval p; $v_0(k)$ is the number of departure vehicles at the upstream observation point during time interval k.

The dispersion phenomenon can be depicted by Fig. 2. At the upstream point, when vehicles discharging during the green phase, the initial platoon is compact and the platoon length (time interval between the leading vehicle and tailing vehicle) is small. When traveling to the downstream intersection, the platoon disperses to some extent. The average arrival volume per time interval becomes smaller than before.

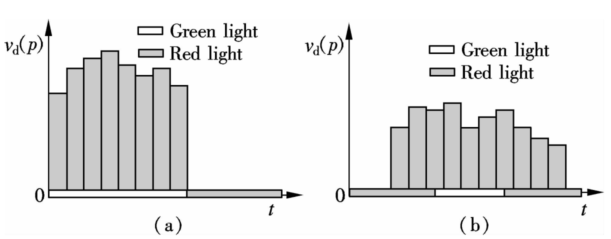

Fig. 2 Illustration of platoon dispersion on artery. (a) Upstream point; (b) Downstream point

The objective of this coordination system is to reduce arterial vehicle delay, so the first step is to develop a delay function. The delay of arterial vehicles is closely related to their arrival pattern at the intersection, and the Robertson dispersion model can be used for this because it can predict the vehicle arrival pattern at the downstream intersection.

HCM2000 provides a simple equation to calculate vehicle delay at arterial intersections, and the equation is shown as follows[14]:

$$D = D_1(P_F) + D_2 + D_3 \tag{5}$$

where D is the control delay per vehicle, s/veh; D_1 is the uniform control delay, s/veh; P_F is the uniform delay progression adjustment factor, which accounts for the effects of signal progression; D_2 is the random control delay, s/veh; D_3 is the initial queue delay, s/veh, and it is usually ignored when there is no initial queue vehicle at the beginning of each cycle.

$$D_1 = \frac{0.5C(1-g/C)^2}{1-\min(1, D_s)(g/C)} \tag{6}$$

$$P_F = \frac{(1-P)f_A}{1-g/C} \tag{7}$$

$$D_2 = 900T\left[(D_s-1) + \sqrt{(D_s-1)^2 + \frac{8eD_s}{C_{AP}T}}\right] \tag{8}$$

where C is the common cycle length, s; g is the effective green time for the study phase, s; D_s is the saturation degree; P is the proportion of vehicles arriving during the green phase; f_A is the supplemental adjustment factor for platoon arriving during the green phase, and it can be determined by looking up tables in HCM2000; T is the duration of the analysis period, h; C_{AP} is the capacity of the approach lane, veh/h; e is the adjustment factor and the recommended value is 0.5, and it can also be determined by looking up in the tables.

From Eqs. (4) to (8), we can find that factor P_F is the key point for calculating D and it is related to P. If P is determined, then D can be obtained directly. P is the proportion of vehicles arriving during the green phase and it is affected by the discharging time at the upstream intersection, the link travel time and the dispersion level. As illustrated in Fig. 2, the Robertson dispersion model can efficiently deal with these factors and P can be calculated by the model.

Based on the above analysis, the value function used in this paper is

$$P_I = \sum_{i=1}^{n} \frac{D_i}{n} = \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{\overline{D}_j}{n} \tag{9}$$

where P_I is the performance index, and it represents the average vehicle delay of arterial vehicles; n is the number of intersections in the subarea; D_i is the sum of average vehicle delay of all coordinated directions at intersection i; m is the number of coordinated directions; $\overline{D}_j$ is the average vehicle delay of coordinated direction j.

Based on the above function, the reward is defined as

$$r^{k+1} = \frac{P_I^k - P_I^{k+1}}{\max(P_I^k, P_I^{k+1})} \tag{10}$$

The reward ranges from -1 to 1, where positive reward values are received if the delay of decision time step $k+1$ is lower than that of time step k.

Given a perceived state, Q learning is utilized to quantify the preference and effectiveness of selecting an action. Following every selection of an action, the corresponding Q value is updated as

$$Q(s_k, a_k) \leftarrow Q(s_k, a_k) + \alpha[r^{k+1} + \gamma \max Q(s_{k+1}, a_t) - Q(s_{k+1}, a_k)] \tag{11}$$

where α is the learning rate; γ is the discount rate for the rewards; a_k is the action executed while in state s_k leading to the subsequent state s_{k+1}, and yielding a reward r^{k+1}.

2 Simulation Experiment

2.1 Simulation environment

The test road network consists of three signalized intersections. Sketches of the network and the phase diagrams are shown in Fig. 3. Each intersection contains three phases and phase 2 is selected as the coordinated phase. The saturation flow rate of the approach lane is 1 800 pcu/h; the lost green time of each phase is 3 s; the average vehicle speed is set as 50 km/h; the minimum and maximum cycle lengths are set as 70 and 120 s, respectively.

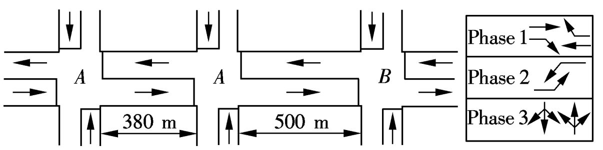

Fig. 3 Sketches of the test network and phase diagrams

To compare the benefits of the established algorithm, two algorithms are tested in this paper.

- **Base algorithm**

It is the conventional arterial progression algorithm without agents. Cycle length is determined by the Webster model. Offsets among intersections are determined by the numerical method.

- **Distributed algorithm**

It is the established algorithm in this paper. The learning rate α is set as 0.5; and the discount rate γ is set as 0.95. At the end of each cycle, the central agent makes a decision about the next action. Because the cycle length may be extended or shortened, the decision interval is not fixed.

In order to study the impact of network traffic load on control benefits, the following three sets of simulation scenarios are investigated in this paper: 1) Scenario 1, intersection saturation degree equals 0.7; 2) Scenario 2, intersection saturation degree equals 0.8; 3) Scenario 3, intersection saturation degree equals 0.9. Generally, when saturation degree is smaller than 0.7, the traffic load is so low that adjacent intersections cannot be coordinated. And when it is greater than 0.9, the traffic is on the edge of congestion, and signal coordination is not feasible.

2.2 Experimental results and analysis

In the Matlab environment, vehicles are generated using Poisson processes with predefined average arrival rates at the entrances of the network. Simulation duration lasts 4 h. The first 3 h are used to calibrate parameters of the distributed algorithm and in the last 1 h, evaluation data are collected every 5 min. Statistical data of the two algorithms in three different scenarios are shown in Fig. 4 to Fig. 6.

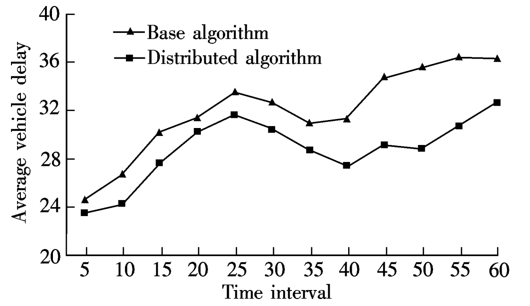

Fig. 4 Comparison of the two algorithms in scenario 1

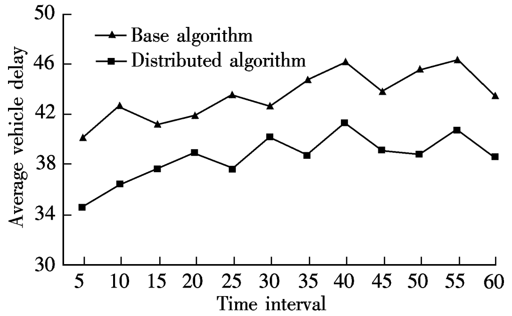

Fig. 5 Comparison of the two algorithms in scenario 2

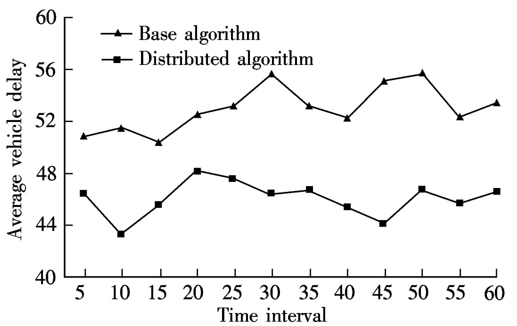

Fig. 6 Comparison of the two algorithms in scenario 3

The plots given in Fig. 4 show the network performance measured for scenario 1. During the simulation period, the average delay of the distributed algorithm is smaller than that of the base algorithm. Using the agents, the evaluation index is reduced by approximately 10.3%. For scenarios 2 and 3, the indices are reduced by 11.5% and 13.1%, respectively. The data also indicate that with the increase in saturation degree, the distributed algorithm performs more efficiently than the base algorithm. To a certain extent, this proves that the multiple agent architecture adapts itself according to the changing dynamics of the traffic network.

3 Conclusion

In this paper, a novel traffic coordination algorithm is developed using a cooperative multiple agent architecture. The uniqueness of this algorithm lies in the embedment of the Robertson platoon dispersion model, which can capture the dynamic movement characteristics of traffic flows on arteries. The performance of the algorithm is evaluated in a Matlab environment and the results show that it outperforms the conventional signal coordination algorithm in three traffic load scenarios. The promising results obtained in this paper provide a foundation for further applications of similar cooperative features in multiple agent systems.

References

[1] National Transportation Operations Coalition. National traffic signal report card [R]. Washington, DC: National Research Council, 2005.

[2] Little J D C, Kelson M D, Gartner N. MAXBAND: a program for setting signals on arterials and triangular networks [J]. *Journal of Transportation Research Record*, 1981, **75**: 40 − 46.

[3] Lin L T, Tung L W, Ku H C. Synchronized signal control model for maximizing progression along an arterial [J]. *Journal of Transportation Engineering*, 2010, **36**(8): 727 − 735.

[4] Xu Jianmin. *Traffic management and control* [M]. Beijing: China Communications Press, 2007: 230 − 240. (in Chinese)

[5] Choy M C, Srinivasan D, Cheu R L. Hybrid agent architecture for real-time traffic signal control [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 2003, **33**(5): 597 − 607.

[6] Liu A C, Urbanik T, Kohls A G.. Reinforcement learning-based multi-agent system for network traffic signal control [J]. *IET Intelligent Transportation Systems*, 2010, **4**(2): 128 − 135.

[7] Yang Q F, Yang C. On-line selection method of the traffic control and route guidance collaboration mode based on Q learning algorithm [J]. *Journal of Jilin University: Engineering and Technology Edition*, 2010, **40**(5): 1215 − 1219. (in Chinese)

[8] Huang Yanguo, Tang Jun, Xu Lunhui. Urban road traffic signal control method based on agent [J]. *Journal of Highway and Transportation Research and Development*, 2009, **26**(10): 126 − 129. (in Chinese)

[9] Ma Shoufeng, Li Ying, Liu Bao. Agent-based learning control method for urban traffic signal of single intersection [J]. *Journal of System Engineering*, 2003, **18**(3): 273 − 278. (in Chinese)

[10] Abdulhai B. Reinforcement learning for the true adaptive traffic signal control [J]. *Journal of Transportation Engineering*, 2003, **129**(3): 278 − 285.

[11] Oliveira L D B, Camponogara E. Multi-agent model predictive control of signaling split in urban traffic networks [J]. *Transportation Research Part C*, 2010, **18**(1): 120 − 139.

[12] Wiering M. Multi-agent reinforcement learning for traffic light control [C]// *Proceedings of the 17th International Conference on Machine Learning*. San Francisco, CA, USA, 2000: 1151 − 1158.

[13] Wang Dianhai. *Traffic flow theory* [M]. Beijing: China Communications Press, 2002: 146 − 147. (in Chinese)

[14] Transportation Research Board. Highway capacity manual [R]. Washington, DC: National Research Council, 2000.

基于车队离散模型的分布式多智能体信号协调控制算法

别一鸣[1]　王殿海[2]　马东方[1]　朱自博[1]

([1]吉林大学交通学院, 长春 130025)

([2]浙江大学建筑工程学院, 杭州 310058)

摘要:为减少干线协调交叉口的车辆延误,基于多智能体系统与增强学习算法(RL)建立了一种新的分布式交通信号协调算法.增强学习算法通过训练各个智能体与外界环境的交互能力达到减少车辆延误的目的.为更精确地描述干线上的车队运动规律,引入了罗伯逊车队离散模型,并基于该模型以及 HCM2000 中的干线交叉口车流延误计算公式建立了 RL 中的回报函数.在 Matlab 中仿真验证了所建算法的控制效果,并在 3 种不同交通负荷下与传统信号协调算法进行对比.结果表明,该算法较传统算法能有效降低干线车流延误;并且随着干线饱和度的增加降低幅度逐渐增大.该结果验证了所建算法的可行性与有效性.

关键词:多智能体;信号协调;增强学习;车队离散模型

中图分类号:U491