2015 IEEE 18th International Conference on Intelligent Transportation Systems

Adaptive Group-based Signal Control Using Reinforcement Learning with Eligibility Traces

Junchen Jin* and Xiaoliang Ma

Abstract— **Group-based signal controllers are widely deployed on urban networks in the European countries. However, group-based signal controls are usually implemented with rather simple timing logics, e.g. vehicle actuated timing. In addition, group-based signal control systems with pre-defined signal parameter settings show relatively poor performances in a dynamically changed traffic environment. This study, therefore, presents an adaptive group-based signal control system capable of changing control strategies with respect to non-stationary traffic demands. In this study, signal groups are formulated as individual agents. The signal group agent learns from traffic environments and makes intelligent timing decisions according to the perceived system states. Reinforcement learning with multiple-step backups is applied as the learning algorithm. Agents on-line update their knowledge based on a sequence of states during the learning process rather than purely on the basis of single previous state. The proposed signal control system is integrated into a software-in-the-loop simulation (SILS) framework for evaluation purpose. In the testbed experiments, the proposed adaptive group-based control system is compared to a benchmark signal control system, the well-established group-based fixed-time control system. The simulation results demonstrate that learning-based and adaptive group-based signal control system owns its advantage in dealing with dynamic traffic environments in terms of improving traffic mobility efficiency.**

I. INTRODUCTION

Traffic congestion problems can be alleviated by optimally performing traffic control and management strategies. Group-based signal controller is one of the most commonly used traffic control facilities in the European countries. Group-based control shows its advancements in allocating green times, especially in the case that traffic demands on different movements are unbalanced. Unlike stage-based control, group-based control assigns time settings to each single traffic movement rather than a group of compatible movements. Previous studies have pointed out that group-based signal control, in comparison to stage-based signal control, has the potential to improve traffic mobility and sustainability [1].

Junchen Jin is a Ph.D. student with the Traffic Simulation & Control Group, Division of Transport Planning, Economics and Engineering (TEE), KTH Royal Institute of Technology, Teknikringen 10, 10044, Stockholm, Sweden.

Xiaoliang Ma Ph.D. is with the Traffic Simulation & Control Group, Division of Transport Planning, Economics and Engineering (TEE), KTH Royal Institute of Technology, Teknikringen 10, 10044, Stockholm, Sweden.

*Corresponding author; e-mail: junchen@kth.se

To the best of the authors' knowledge, the existing group-based signal control systems apply rather simple timing logics. The signal parameters are pre-defined by traffic engineers using their expertise and experiences or by off-line optimization algorithms. Simple timing logics, coupled with pre-determined signal timing settings, usually are not able to handle sudden changes in traffic environment. Therefore, an adaptive group-based control system is paramount capable of self-adjusting control schemes regarding the changed traffic conditions. In this regards, some studies have reported that reinforcement learning (RL) framework is fundamentally well-suited for coping with signal control problems [2], [3], [4]. In the RL framework, signal group agents comply trials in the light of their own knowledges. The trails result in new observations from traffic environment. The agents learn new knowledge from such experiences using reinforcement learning algorithms. Signal group agents keep conducting learning process in the various operational conditions and become smarter and smarter.

Most RL-based signal control systems update knowledge merely based on the previous one state. However, in practice, decisions made by signal controllers have effects on several following states. For example, if the current phase is ordered to extend for three steps, this will have impacts on traffic performances for at least three succeeding states. Temporal difference algorithm with multiple-step backups is, thus, applied in this study. This type of reinforcement learning algorithm enables signal controller to look backwards all the way until the beginning of the defined learning horizon. Together with multiple-step backups approach, eligibility traces are used for recording the degree of eligibility for the undergoing learning process. The idea behind eligibility traces is that a state initiates a short-term memory (namely trace) each time when this state is visited.

Thorpe and Anderson firstly applied trace-based reinforcement learning method to signal control system [5]. However, they used a reinforcement learning method to optimize signal settings of a fixed-time control rather than designing a new signal control system. The authors in study [6] pointed out that eligibility traces do not significantly improve their RL-based signal control system. Their reward function is defined by the difference in travel delay between two successive decision points. Nevertheless, the benefit of

978-1-4673-6596-3/15 $31.00 © 2015 IEEE
DOI 10.1109/ITSC.2015.389

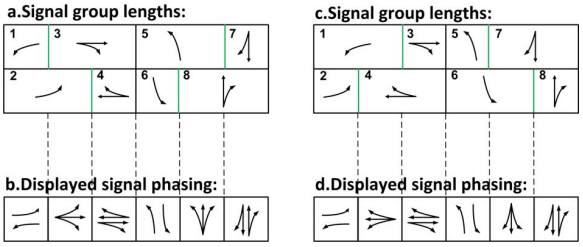

Fig. 1: Examples of group-based phasing techniques.

eligibility traces normally appears where rewards are delayed by many steps, while this cannot be revealed in their reward function. In the present study, signal group agents apply trace-based reinforcement learning framework. In Section II, signal groups are formulated as individual intelligent agents. Those agents learn from traffic environment and make intelligent signal timing decisions with response to real-time traffic conditions. Design elements of the proposed signal control system are described in detail in Section III. Section IV elaborates the experimental results of the proposed adaptive group-based signal control system.

II. ADAPTIVE GROUP-BASED SIGNAL CONTROL

A. Group-based signal control

In a group-based signal control system, signal group and phase are two basic components. Signal group is defined as a group of traffic movements that are simultaneously controlled by the identical indications. Phase is a combination of signal groups. Timings are directly assigned to signal groups, and phases are correspondingly generated. Fig. 1 gives a typical example of the ability of group-based phasing techniques. On the left diagram of Fig. 1, traffic flow associated with signal group 2 is higher compared with the traffic flow controlled by signal group 1. On the contrary, traffic flow for signal group 1 is relatively larger in Fig. 1(c). Signal group with larger traffic flow deserves longer time length. That means eastbound left-turn (signal group 2) deserves longer time in the first scenario while westbound left-turn (signal group 1) is supposed to be assigned more green time in the second scenario. Hence, group-based phasing differently generates the second phases regarding the above two scenarios (see Fig. 1(c) and Fig. 1(d)). Implementations of group-based signal control can be viewed in detail in the study [7].

Once a signal group is activated, signal control system automatically searches for another signal group as a substitute to switch to when the active signal group is terminated. Here, we name the substitute signal group as a 'candidate signal group'. A signal group cannot be nominated as the candidate if it was activated in the current cycle, or it has conflicts with the rest of signal groups in the current phase. If the candidate signal groups are not found, the ordered-to-terminate signal group has to wait until all of the other signal groups in the current phase are ordered to terminate. The signal group keeps passive green during the waiting period while no detections are reported to that signal group.

B. Reinforcement learning with multiple-step backups

In principle, signal groups can be formulated as individual agents. Specifically, a signal group agent perceives states and feedbacks from traffic environments, learns knowledge based on its learning algorithms and thereafter makes timing decisions. In a reinforcement learning framework, agent knowledge is represented by a long-run cumulative reward [8]. Dynamics of transportation system are difficult to model due to its complexity and changeability. Temporal Difference(TD), a popular class of RL algorithms, works on an on-line updating procedure by which signal group agents are able to learn directly from raw experiences without having any accessibilities to the dynamics of traffic environment. If the TD algorithm stores multiple-step backups and exploits them in the knowledge updating process, the reserved knowledge can also be taken into account. Agent knowledge is updated with the information that is received from the beginning of a user-defined episode. In this study, start point of the episode corresponds to the time point when signal group agent is activated. The episode lasts until the green period of signal group agent is terminated.

TD(λ) is the multiple-step backups version of TD algorithm. TD(λ) utilizes eligibility trace to achieve the average effects of multiple-step backups. The traces decay gradually over time, which matches the biological brain strategies for deciding how recently received stimuli should be used together with the current stimuli. SARSA(λ) learning is an on-policy TD(λ) learning method. Here, policy stands for the mapping from states to actions. Equation 1 and Equation 2 show the update process of cumulative reward for all of the state-action pairs at time step $t+1$. The cumulative rewards are updated by the previous cumulative rewards and temporal differences. Temporal difference is weighted by the degree of eligibility. Thus, the global temporal difference values trigger proportional to all recently visited states in the defined episode. For example, an agent takes action a_t at time t when it is in state s_t. Eventually, the cumulative reward in the end of the episode can be updated by Equation 4. This equation shows that all the following state-action pairs make contributions to update the cumulative reward for stage-action pair (s_t, a_t) until the end of the current episode.

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha\delta_t(s,a)e_t(s,a),\ s \in \mathcal{S},\ a \in \mathcal{A} \quad (1)$$

$$\delta_t(s,a) = r_{t+1} + \gamma Q_t(s_{t+1},a_{t+1}) - Q_t(s_t,a_t),\ s \in \mathcal{S},\ a \in \mathcal{A} \quad (2)$$

$$a_{t+1} = \arg\max_{a'} \pi(a'|s_t),\ a' \in \mathcal{A} \tag{3}$$

$$\begin{aligned} Q_T(s_t, a_t) \leftarrow & Q_t(s_t, a_t) + \alpha\delta_t(s_t, a_t)e_t(s_t, a_t) \\ & + \alpha\delta_{t+1}(s_{t+1}, a_{t+1})e_{t+1}(s_{t+1}, a_{t+1}) \\ & + \cdots + \alpha\delta_T(s_T, a_T)e_T(s_T, a_T) \end{aligned} \tag{4}$$

where $Q_t(s,a)$, $\delta_t(s,a)$ and $e_t(s,a)$ denotes the cumulative reward, temporal difference and the degree of eligibility when agent is in state s and takes action a at time step t, respectively; $\mathcal{S}$ and $\mathcal{A}$ denote state space and discrete action vector for all signal group agents; r_t represent the immediate reward at time t; $\pi(a|s)$ is the policy function that is the probability of taking action a when agent is in state s_t at time t; α and γ respectively represent learning rate and discount rate; T denotes the end of the current episode.

Two types of eligibility trace strategies, accumulating traces and replacing traces, are implemented in this paper. Values of eligibility trace are set to 0 for all of the state-action pairs in the beginning of the episode. Accumulating trace adds more credits to more recent events (recency) and to the events which have occurred more times (frequency). Trace decay builds up each time when the state-action pair is visited. Replacing trace only retains the recency property but discards the frequency property [9]. Definitions of the two trace strategies are listed in Equation 5 and Equation 6. It can be seen from the equations that both trace approaches implement exponentially-decaying memory. At each step, accumulating eligibility traces are increased by 1 for the current visiting state-action pair while decay by $\gamma\lambda$ for the other state-action pairs. Whereas replacing trace is set to 1 each time for the state-action pair which is currently visited regardless of the presence of a prior trace.

$$e_t(s,a) = \begin{cases} \gamma\lambda e_{t-1}(s,a) + 1, & \text{if } s = s_t \text{ and } a = a_t \\ \gamma\lambda e_{t-1}(s,a), & \text{otherwise.} \end{cases} \tag{5}$$

$$e_t(s,a) = \begin{cases} 1, & \text{if } s = s_t \text{ and } a = a_t \\ 0, & \text{if } s = s_t \text{ and } a \neq a_t \\ \gamma\lambda e_{t-1}(s,a), & \text{if } s \neq s_t. \end{cases} \tag{6}$$

where $s \in \mathcal{S}$ and $a \in \mathcal{A}$; $\lambda \in [0,1]$ denotes the trace decay parameter. The agent performs pure bootstrapping (SARSA learning) when $\lambda = 0$ while performs no bootstrapping (Monte Carlo algorithm) if $\lambda = 1$.

The generalization of a signal group agent can be represented by a tuple $sg = (\mathcal{S}[sg], \mathcal{A}[sg])$, where sg represents a signal group agent. From a practical point of view, signal group agents carry out learning process at discrete time. Pseudocode of SARSA(λ) learning for multi-agent signal control system is summarized in Fig. 2. Eligibility traces are firstly initialized as 0 for all the signal group agents with any state-action pairs. At each simulation step, all signal group agents are requested to obtain states from traffic environment. Reward is also computed according to the previous-step simulation.

```
Initialize meta-parameters α, γ, τ and λ;
for all sg ∈ signal_groups do
  for all s ∈ S[sg] do
    for all a ∈ A[sg] do
      Initialize Q[sg](s,a), s[sg] and a[sg];
      e[sg](s,a) = 0;
    end for
  end for
end for
for each simulation step do
  for all sg ∈ signal_groups do
    if sg.active() = True then
      Take action a[sg];
    end if
    Observe new states s' from environment;
    Compute reward r;
    if sg.active() = True then
      a'[sg] ← softmax(sg, Q, s, a);
      δ(sg, S, A) ← td(sg, r, Q, s, a, s', a');
      e(sg, S, A) ← trace(e, sg, S, A)
      Q(sg, S, A) ← update_Q(sg, Q, e, δ, S, A);
      s[sg] ← s'[sg]; a[sg] ← a'[sg];
    end if
    if sg.ordered_to_terminate() = True then
      for all s ∈ S[sg], a ∈ A[sg] do
        e[sg](s,a) = 0;
      end for
    end if
  end for
end for
```

Fig. 2: Pseudo-code of SARSA(λ) learning for the adaptive group-based signal control system.

Thereafter, the new action is chosen for active signal group agents based on softmax action selection policy. During the action selection process, a Boltzmann distribution is used to determine the probabilities to select actions when the agent is in state s_t at time t (see Equation 7). The highest selection probability is given to the greedy action which corresponds to the highest Q value. Equation 1 - Equation 6 demonstrate the update principle for cumulative reward $Q(s,a)$. If a signal group agent is ordered to terminate, the 'trace back' process will immediately stop and all eligibility traces are reset to 0.

$$\pi(a|s_t) = \frac{e^{Q(s_t,a)/\tau}}{\sum_{a'} e^{Q(s_t,a')/\tau}},\ a \in \mathcal{A},\ a' \in \mathcal{A} \tag{7}$$

where τ is a positive temperature parameter.

C. SILS framework for multi-agent system

In this study, software-in-the-loop simulation (SILS) framework is applied for evaluation purpose. The SILS framework was presented in detail in a previous study [10]. Two modules, distributed multi-agent signal control system and simulator-controller interface, are inte-

grated into the adaptive group-based signal controller software. SUMO version 0.19.0 [11] is employed as the microscopic traffic simulator in this computational framework. SUMO provides a socket connection interface, TraCI, allowing for on-line interactions. Through TraCI, detector information and vehicle characteristics are sent to the simulator-controller interface. Detection information is translated to states while vehicle characteristics are estimated as the feedbacks (rewards) to signal control system. States and feedbacks are the inputs to signal group agents for the purpose of making subsequent timing actions. Traffic light indications are interpreted by the timing actions and are sent to microscopic traffic simulator. Then signal controller in traffic simulator executes the received indications. Adaptive group-based signal control system is formulated as a multi-agent system. Signal group agents are able to receive information from other agents and incorporate the information into their decision-making process. Cooperations among agents are achieved by sharing partial information of the states with their neighbors. Two kinds of signal groups can be regarded as neighbors. One is the other signal group agents that operate in the current phase and the other one is the candidate signal group.

III. SIGNAL CONTROL SYSTEM DESIGN

A. State representation

State is defined to be fully observable under the current situations of road transport infrastructure. States, hereby, are either reported by detectors or provided by signal controllers. In this study, designations of signal controller and detector refer to Swedish standard configurations for signal control system. The Swedish detection system consists of long loop detectors locating close to stop line and short station detectors that are placed 50-80 meters upstream from the stop line. The functionality of short detectors is to measure the level of traffic flow by means of reporting time gaps between two passing vehicles. For instance, gap between vehicles will be reduced if traffic flow increases. Besides, occupancy status, determining whether vehicles are queuing before the stop line, is sent by long detectors. Signal controller also reports the other two states. They are elapsed green time and phase scenario. The elapsed green time is green time after minimum green time is passed. Range of elapsed green time is transformed from 0 to 50 seconds to a scale from 0 to 9. On the other hand, phase scenario represents whether the signal group agent has to wait for other signal group agents. Consequently, seven feature-based states, including the states sent by neighbors, are designed in the proposed adaptive group-based signal control system (see Equation 8).

$$\mathcal{S} = (g, o, gr, p, g_c, o_c, max\_gr) \tag{8}$$

where g, o, gr and p represent gap, occupancy, elapsed green time status and phase scenario states, respectively. g_c and o_c present gap and occupancy states for candidate signal group while max_gr represents the maximum value of green times among the other signal groups in the current phase.

B. Action definition

Generally, the actions of a signal group agent are either to be ordered to terminate or to extend the green time. Minimum recall mode is implemented so that a signal group agent is activated at least for minimum green time. In addition, maximum green time is defined to guarantee that signal extension, in principle, is not authorized all the time. Therefore, termination action is only valid when the minimum green time is passed while extension action is constrained by the maximum green time. It is assumed that the time duration for vehicles driving from the short detector to the tail of long detector is less than four seconds in a 50km/h - 60km/h speed environment. Action space is defined in Equation 9. If action is ordered to terminate, the subsequent signal group status is determined by whether signal group agent has to wait for other signal groups. The status changes to passive green if signal group has to wait for others. While signal group agent will terminate after accounting for the minimum green time, yellow time and clearance times if it is not required to wait for others.

$$\mathcal{A} = (0, 1, 2, 3, 4) \tag{9}$$

where $a = 0$ means that the agent is ordered to terminate; $a = 1$, $a = 2$, $a = 3$ and $a = 4$ respectively represent that green time of the signal agent is extended by one, two, three and four seconds.

C. Reward function

Signal group agents share the reward value at the same intersection. The proposed signal controller is designed to improve traffic mobility efficiency. Therefore, reward function is defined as the relative reductions in travel delay caused by the previous action (see Equation 10). Vehicles are counted to compute travel delay when they enter the position which is 200 meters upstream from an intersection. And vehicles are not counted any more when they pass the intersection. If the reward has a positive value, this implies that the immediate delay is reduced by executing the selected action. Similarly, a negative reward value indicates that the chosen action leads to an increase in total travel delay.

$$r_t = td_{ref} - ttd_t \tag{10}$$

where r_t is reward value at time t; td_{ref} is the user-defined reference of total travel delay and ttd_t is the total travel delay for the whole intersection at time t.

TABLE I: Traffic volume (vehicles/hour) for each turning movement on the study intersection.

Volume Level	period	Eastbound			Westbound			Northbound			Southbound		
		L	T	R	L	T	R	L	T	R	L	T	R
Medium	0-2h	50	800	75	50	800	75	30	500	40	30	500	40
High	2-4h	60	1200	80	60	1200	80	40	600	45	40	600	45
Medium	4-6h	40	700	70	40	700	70	25	500	35	25	500	35
Low	6-8h	30	400	30	30	400	30	20	300	20	20	300	20
Medium	8-10h	40	750	70	40	750	70	35	450	35	35	450	35
High	10-12h	70	1300	90	70	1300	90	45	600	50	45	600	50

L, T and R represent the turning rates of left-turn, through and right-turn movements, respectively.

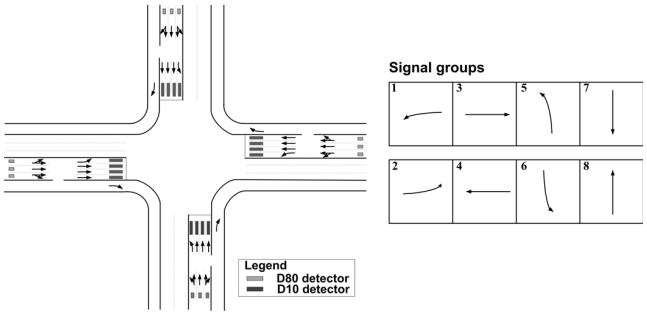

Fig. 3: Layout and signal groups of the study network.

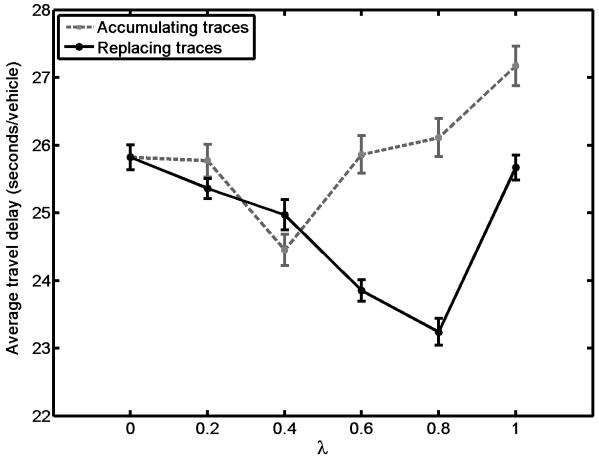

Fig. 4: Convergence results of adaptive group-based signal control with different λ settings.

IV. TEST-BED EXPERIMENTS

The proposed adaptive group-based signal control is tested on a four-armed isolated intersection. Eight signal groups are defined for this intersection. Fig. 3 shows the layout of study network and the signal groups. In the experiments, right-turn directions are not regulated by traffic lights. 1,000 one-hour simulation runs are performed to train the proposed adaptive group-based signal control system. For the purpose of avoiding the effects of initial vehicle loadings, signal group agents start learning after 900-second simulation.

Both accumulating trace and replacing trace are tested in this study. Sensitivity analysis of trace parameter (λ) is also carried out for both eligibility trace strategies. Fig. 4 shows the convergence results for SARSA(λ) learning using different settings of trace parameter. The mean value and standard deviation of average travel delay among the last 100 of 1,000 simulation runs are computed. Although accumulating traces based SARSA(λ) at most ($\lambda = 0.4$) improve only slightly over no traces ($\lambda = 0$) case (one-step backup), SARSA(λ) learning with replacing traces outperforms no-trace case at all except for no bootstrapping case ($\lambda = 1.0$). Besides, accumulating traces dramatically degrade convergence results when λ becomes larger. Replacing traces, on the other hand, significantly improve the efficiency of traffic mobility at high values of λ. Consequently, signal control agents will make better actions if they consider multiple steps backwards rather than updating knowledge based on single state from the previous step. In the following experiments, both accumulating trace and replacing trace are implemented with the λ that yields the lowest average travel delay.

Further, a 12-hour simulation experiment is carried out to compare the proposed adaptive group-based signal control system to a benchmark signal control system, a well-established group-based fixed-time control system. An off-line optimization framework is used to tune the signal control parameters for the benchmark signal control system [12]. Three different levels of traffic volumes are used in the experiments and the inflow volume changes every two hours (see Table I). Thirty randomly seeded simulation runs are conducted to make the evaluation results statistically significantly. To provide insights of the behaviors of signal control agents with respect to the changes in traffic environment, Fig. 5 demonstrates the total travel delay within three minutes for all vehicles. In comparison to the optimized group-based fixed time signal controller, adaptive group-based signal controllers perform much better in the case that level of traffic volume is not high. For example, level of traffic volume changes from medium to high to medium during simulation period 4-10 hours. Both two adaptive group-based control systems significantly outperform optimized group-

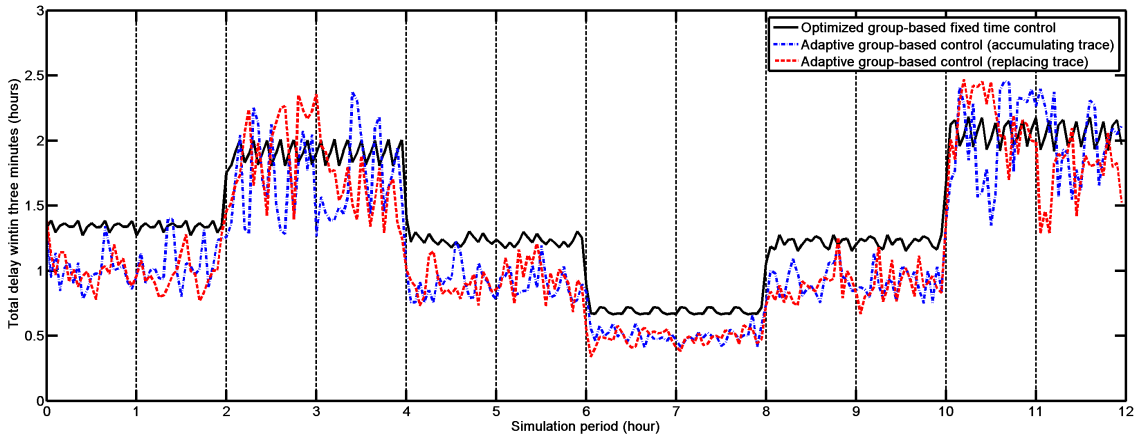

Fig. 5: Performance analysis of adaptive group-based signal control systems through 12-hour simulation.

based fixed time control. On the other hand, group-based fixed time control, despite being dynamic in generating phase pictures, is unable to adapt to the underlying changes of traffic demands. Learning-based group-based signal control system has the ability to gradually decrease travel delay after the level of traffic volume is changed. For instance, the level of volume changes from medium to high at time point '2 hours'. SARSA(λ) with replacing trace algorithm exhibits a reduction in travel delay after time point '3 hours'. However, SARSA(λ) with accumulating trace does not perform as well as that with replacing trace when traffic volume is relatively high. In conclusion, SARSA(λ) learning has the potential to improve efficiency of signal control system in the context of group-based phasing technique.

V. CONCLUSION

The present study is to extend group-based signal control based on the existing road transport infrastructures. The proposed adaptive group-based signal control system has the capability to deal with non-stationary traffic patterns. We treat signal groups as individual intelligent agents. Group-based signal control system is implemented using the distributed multi-agent architecture in which central level of manipulation is not required. Signal agents learn a behavior by observing traffic patterns. SARSA(λ) learning, an on-policy reinforcement learning algorithm with multiple-step backups, is implemented in this study. Such an algorithm is adaptive in nature and does not require model information of the traffic system dynamics. The proposed signal control system is integrated into a software-in-the-loop simulation framework for evaluation purpose. Experimental results successfully demonstrate that the proposed adaptive group-based control system has the potential to improve traffic mobility efficiency, compared to a well-established group-based fixed-time signal system.

References

[1] J. Jin, X. Ma, and I. Kosonen, "Stochastic optimization of group-based signal control and coordination using traffic simulation," in *Transportation Research Board Annual Meeting, 94th, 2015, Washington, DC, USA*, 2015, pp. 389–403.

[2] L. Prashanth and S. Bhatnagar, "Reinforcement learning with function approximation for traffic signal control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 412–421, 2011.

[3] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): Methodology and large-scale application on downtown toronto," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1140–1150, 2013.

[4] K. Prabuchandran, H. K. AN, and S. Bhatnagar, "Multi-agent reinforcement learning for traffic signal control," in *Proceedings of the 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2014, pp. 2529–2534.

[5] T. L. Thorpe and C. W. Anderson, "Traffic light control using SARSA with three state representations," Citeseer, Tech. Rep., 1996.

[6] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Design of reinforcement learning parameters for seamless application of adaptive traffic signal control," *Journal of Intelligent Transportation Systems*, vol. 18, no. 3, pp. 227–245, 2014.

[7] J. Jin and X. Ma, "Implementation and optimization of group-based signal control in traffic simulation," in *Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2014, pp. 2517–2522.

[8] A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 1998.

[9] S. P. Singh and R. S. Sutton, "Reinforcement learning with replacing eligibility traces," *Machine learning*, vol. 22, no. 1-3, pp. 123–158, 1996.

[10] J. Jin and X. Ma, "Adaptive group-based signal control by reinforcement learning," *Procedia-Social and Behavioral Sciences*, 2015.

[11] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO–simulation of urban mobility," *International Journal On Advances in Systems and Measurements*, vol. 5, no. 3 and 4, pp. 128–138, 2012.

[12] X. Ma, J. Jin, and W. Lei, "Multi-criteria analysis of optimal signal plans using microscopic traffic models," *Transportation Research Part D: Transport and Environment*, vol. 32, pp. 1–14, 2014.