Back-PressureおよびQ-Learningに基づく適応的交通制御アルゴリズム

Arnan Maipradit[1,a)] Gao Juntao[1,b)] Tomoya Kawakami[1,c)] Minuro Ito[1,d)]

Abstract: 近年，交通渋滞は特に問題となり，移動時間の長大化や大気汚染などを引き起こしている．Back-Pressureに基づく交通制御アルゴリズムが交通渋滞を緩和するために提案されているが，それらは交通情報が不正確な場合や局所的な場合，効果的なスケジューリングが行えない．そこで本研究では，交通渋滞を効果的に緩和するため，Back-PressureとQ-Learningに基づく適応的な交通制御アルゴリズムを提案する．提案手法はリアルタイムの正確な交通情報とQ-Learningで学習した広域的な情報に基づいて，交通制御を行う．提案手法はシミュレーションにより評価し，交通渋滞が緩和されることで，車両の平均移動時間が比較手法よりも最大で25%削減されることを確認した．

Keywords: Back-pressure, Q-Learning, vehicle routing

1. Introduction

Traffic congestion in urban areas become a serious problem, which increases vehicle travel time [1]. Many road traffic network use fixed cycles to control traffic lights, which easily causes traffic congestion because they do not consider real-time traffic information. Moreover, drivers can not select optimal routes to avoid traffic congestion because the drivers can not get real-time global traffic information. Survey showed that traffic congestion can be mitigated by efficient traffic signal control and vehicle routing [2].

Some metropolises implement adaptive traffic signal control, such as SCOOT (Split Cycle Offset Optimization Technique) [3], SCATS (Sydney Coordinated Adaptive Traffic System) [4], RHODES (Real-time Hierarchical Optimizing Distributed Effective System) [5], etc. These systems adjust parameters of signal control, such as phase and cycle length, according to real-time traffic situation [6]. Although the results of these systems showed some improvement, most of the systems cannot guarantee performance [14]. Control theory have been implemented for improving traffic signal control algorithms [7-12]. However, none of these systems can be implemented in large urban traffic networks because large urban traffic networks require decentralized algorithm while algorithms [7-12] are centralized.

Recently, decentralized traffic control algorithms based on back-pressure have been proposed [13-17]. Back-pressure was applied in traffic signal control in [14], and was showed superior to fixed cycle signal control. These back-presssure based traffic control algorithms do not consider adaptive control of vehicle routes, e.g., some use fixed shortest path for route management, which easily results in traffic congestion especially during rush hours. Some research considered joint adaptive signal control and dynamic vehicles routing [19] [21]. However, they focused on providing individual vehicles adaptive route guidance only. Coordination between different vehicles may further reduce traffic congestion.

Originally back-pressure algorithm was developed for routing and scheduling in communication networks [22], and has been used to jointly control traffic signals and vehicle route in road networks [18]. However there is a big difference between road network and communication network: communication time between two nodes is almost zero, while vehicle travel time between two junctions is significant. Directly applying back-pressure algorithm to traffic control as [18] is not appropriate. Based on this observation, [24] proposed an adaptive traffic control algorithm which considers the difference between road network and communication network. Specifically, they control traffic signal and vehicle routes based on real-time traffic information, like vehicle speed and vehicle position. As a result, their algorithm significantly reduces traffic congestion.

However, their work controls traffic lights and vehicle routes based on only local traffic information, i.e., every control agent considers information of vehicles only at one junctions. Therefore, their algorithm is short-sigthed. For more efficient traffic control, global traffic information and coordination between different junction agents are needed. In this paper, we extend the work [24] and propose an adaptive traffic control algorithm that controls traffic based on accurate real-time traffic information and global traffic information, where neighboring junction agents exchange traffic information with each other and learning global

1 Nara Institute of Science and Technology Nara, Japan
a) maipradit.arnan.lu0@is.naist.jp
b) jtgao@is.naist.jp
c) kawakami@is.naist.jp
d) ito@is.naist.jp

traffic information by Q-learning.

2. SYSTEM MODEL

2.1 Road Network

Every road network atleast consist of Road (R), Junction (J), Lane (L), Origin (o), Destination (d) and Flow (f). From previous parameters we can create our road network model as : $\mathbb{J} = \{J_1, J_2, J_3, ...J_{max}\}$ is the set of roads, each R_i has three L_{ij} because we assume that each road has three lanes as show in Fig. 1. Normally vehicles must have origin and destination, so we group the vehicles with the same origin and destination into a same f. We define $\mathbb{F}$ as the set of all flows, $\mathbb{O} = \{o(f), f \in \mathbb{F}\}$ as the set of all origins, $\mathbb{D} = \{d(f), f \in \mathbb{F}\}$ as the set of all destinations and $\lambda_f(t)$ as the number of vehicle s exogenously at time slot t.

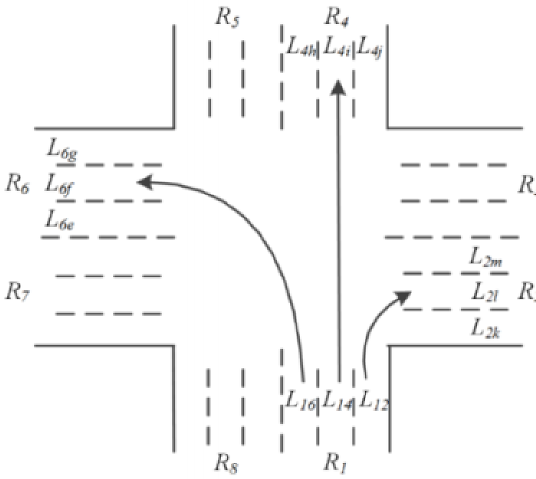

Fig. 1 Possible traffic movement from road R_1 at a junction J_1.

2.2 Traffic Rules

Vehicles travelling along a road network from origin to destination. It may pass a junction from road R_i to road R_j, We define it as traffic movement (R_i, R_j). Every junction that has traffic light, it should consist of traffic phase p_i^a. p_i^a is the set of all possible traffic movement that can happen in the same time period. We define $\mathbb{M}_a$ as the set of all possible movement and $\mathbb{P}_a = \{p_1^a, p_2^a, ...p_{max}^a\}$ as the set of all possible traffic phase. In traffic signal control process, agent at each junction have to activate p_i^a by selecting from $\mathbb{P}_a$. An example of all possible phases at a junction is given in Fig. 2.

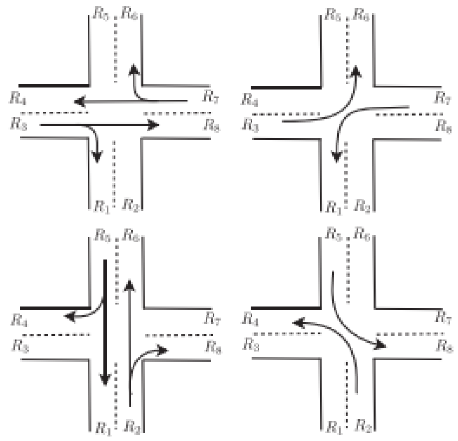

Fig. 2 All possible phase at a junction.

3. BACK-PRESSURE BASED ADAPTIVE TRAFFIC SIGNAL CONTROL AND VEHICLE ROUTING

In this section, we describe our back-pressure and Q-Learning based traffic control algorithm, which reduces traffic congestion using global traffic information in road network.

Each junction has a control agent that collects information of vehicle speed and vehicle position every time slot for traffic control. At each time slot, the control agent at each junction performs three tasks sequentially. Task 1 (Learning Global Congestion Information): It exchanges congestion level information with neighboring agents. Based on exchanged congestion information, the agent updates its own congestion estimate based on Q-learning. Through this kind of congestion information exchange and update, all agents will finally obtain global congestion information which can be used in the following two tasks. Task 2 (Traffic Phase Selection): It selects a traffic phase to control traffic signals based on back-pressure algorithm. Task 3 (Vehicle Routing): After a vehicle passes through the junction under the traffic phase in task 2, the agent determines which lane the vehicle should join. Since each lane determines vehicle turning direction, i.e., going straight, turning left or turning right, the process of determining lanes for a vehicle to join forms the routing process for that vehicle. To help agents to do the three tasks, we need the following shadow network.

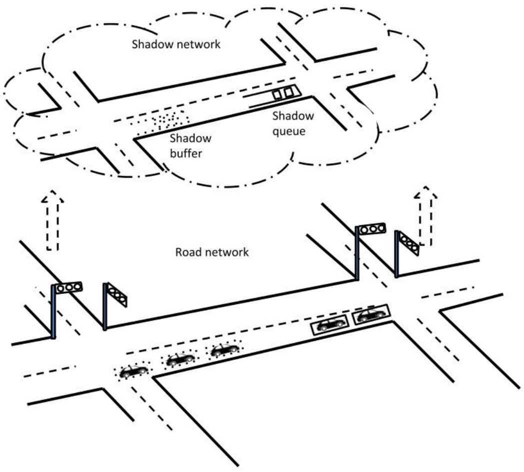

Fig. 3 Illustration of shadow network.

3.1 Shadow Network

An example of shadow network is given in Fig. 3, where a virtual shadow vehicle corresponds to one real vehicle in road network, a shadow buffer corresponds to the beginning part of one real road (a vehicle just passing through a junction will enter this part of road) and a shadow queue corresponds to the end part of one real road (a vehicle running close to next junction will enter this part of the road).

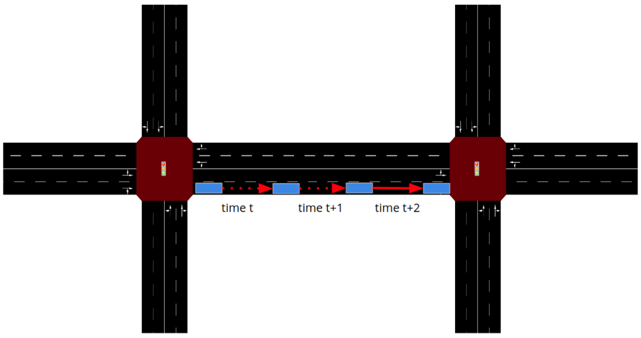

Fig. 4 A vehicle needs time to travel across a road.

In the shadow network, whenever a real vehicle enters the road network, a shadow vehicle is generated and enters the shadow network. Furthermore, one more shadow vehicle is generated with probability $\epsilon, 0 < \epsilon < 1$ and also enters the shadow network. This operation makes sure that algorithm is stable, i.e., queue size will not go to infinite [18,20].

When a real vehicle enters road network from origin road R_i at time slot t and wants to go to destination $d \in \mathbb{D}$, corresponding shadow vehicles will also enter shadow buffer $\bar{B}_i^d(t)$ associated with destination d and road R_i. After that vehicle runs close to the end part of road R_i, the corresponding shadow vehicle first leaves shadow buffer $\bar{B}_i^d(t)$ and then enters shadow queue $\bar{Q}_i^d(t)$ associated with destination d and road R_i. We say a vehicle runs close to end part of one road if its speed is less than 5 Km/h or the distance between vehicle and next junction less than 100 meters.

Similarly, after a real vehicle destined for destination $d \in \mathbb{D}$ leaves road R_i and enters adjacent road R_j at slot t, a corresponding shadow vehicle will also leaves shadow queue $\bar{Q}_i^d(t)$ of road R_i and enters shadow buffer $\bar{B}_j^d(t)$ of road R_j.

Movement of virtual shadow vehicles in the shadow network can be seen as control information exchange, based on which a agent performs its three tasks (details are given in the following section). Movement process of shadow vehicles in our model is quite different from work [18], where a vehicle enters road R_j and its corresponding shadow vehicle immediately enters shadow queue $\bar{Q}_i^d(t)$. This misleads agents in judging congestion level at each road.

3.2 Adaptive Traffic Control Algorithm Based on Back-Pressure and Q-Learning

Our algorithm is decentralized and the agent at each junction runs the following algorithm independently. At every time slot t, each agent perform the three tasks sequentially.

3.2.1 Task 1 Learning Global Congestion Information

The agent at a junction is responsible for estimating route congestion level $R_{ij}^d(t)$ for route to destination d and by the way of road i and the neighbor road j. Each agent maintains a table R_i to store the value of $R_{ij}^d(t)$. At the beginning of each time slot, the agent exchanges information of its own queue length $\bar{Q}_j^d(t)$ and the table R_i with neighboring agents. After exchanging those information, the agent updates its route congestion estimate $R_{ij}^d(t)$ as floows:

$$R_{ij}^d(t) \to (1-\alpha)R_{ij}^d(t-1) + \alpha[\bar{Q}_j^d(t) + \gamma \min_k R_{jk}^d(t)] \tag{1}$$

where α and γ are Q-learning parameters, $0 < \alpha, \gamma <= 1$. If $R_{ij}^d > C_{max}$, set $R_{ij}^d = C_{max}$, C_{max} is a positive constant. Each agent calculates a bias quantity $C_i^d(t)$ as follows:

$$C_i^d(t) = \min_j R_{ij}^d(t) \tag{2}$$

Finally, the bias quantity $C_i^d(t)$ will be used in Traffic Phase Selection. The following task 2 Traffic Phase Selection and task 3 Vehicle Routing are the same with that in work [24] (please refer to [24] for details), except that traffic pressure in our work is defined as follows:

$$w_{ij}^d(t) = max\{(\bar{Q}_i^d(t) + C_i^d(t)) - (\bar{Q}_j^d(t) + C_j^d(t)), 0\} \tag{3}$$

4. SIMULATIONS

In this section, we do simulations to evaluate the performance of our algorithm and compare it to other algorithms as follows.

- Fixed-cycle (FC) signal controller
- Back-pressure based signal controller with shortest path routing (SP-BP) [14]
- Back-pressure based adaptive traffic signal control and vehicle routing without real-time control information update (AR-BP) [18].
- Back-pressure based adaptive traffic signal control and vehicle routing with real-time control information update (ARD-BP)[24].
- Adaptive Traffic Control Algorithm Based on Back-Pressure and Q-Learning (ARD-BP-Q).

4.1 Simulation Setup

We implemented our algorithm in an open-source simulator SUMO (Simulation of Urban MObility)[23]. This is the set of vehicle arrival rate the we use to simulate our algorithm : {100, 200, 300, 400, 500, 800} vehicle/hour.

Structure of road network for simulation is shown in Fig. 5, where each road has different length ranging from 250 meters to 950 meters and different speed limit ranging from 60 km/h to 140 km/h. Each road consists of three lanes: lanes for turning left, right , and going straight as show in Fig. 1. There are 8 origin and destination pairs $\{(o_1, d_1), (o_2, d_2), (o_3, d_3), ..., (o_8, d_8)\}$ as shown in Fig. 4. Vehicle arrival rates are set to be the same for all pairs ranging from 100 vehicle/hour to 1100 vehicle/hour. System time slot is set to be 15 seconds. Shadow vehicle generating parameter ϵ is set to be 0.02 and vehicle routing parameter β is set to be the same as ϵ.

During simulation we collect the following data: vehicle speed, number of vehicles in network, number of arriving vehicles and vehicle traveling time. Vehicle traveling time is the time it takes a vehicle to travel from its origin to its destination.

We separate simulations into two types: algorithm with and without warm-up.

4.1.1 Simulation without warm-up

FC and SP-BP, We run simulation for these algorithms 7200 + 5000 = 12200 second (3.4 hours). We do not collect simulation data of vehicles that enter road network after 7200 seconds, because 5000 seconds left in simulation will make wrong estimate average vehicle travelling time.

4.1.2 Algorithm with warm-up

For algorithms AR-BP , ARD-BP and ARD-BP-Q, we run simulation for 6000 + 7200 + 5000 = 18200 seconds (5 hours). We do not collect simulation data of vehicles that enter road network during the first 6000 seconds, because these algorithms need warm-up state to learn vehicle routing probabilities and reach a stable routing policy. We do not collect simulation data of vehicles that enter road network after 6000 + 7200 = 13200 seconds.

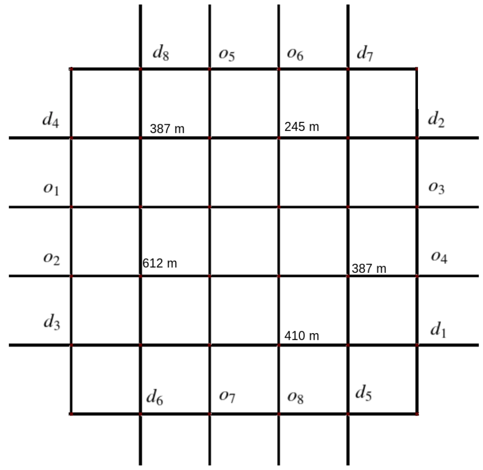

Fig. 5 Road network with 8 pairs of origin and destination, where each road has different length and speed limit.

4.2 Simulation Result and Analysis

All simulation results will summarized in figure. From Fig. 6, we can see that our algorithm ARD-BP-Q has almost lowest average travelling time compare to other algorithm in almost vehicle arrival rates.Compare to ARD-BP our ARD-BP-Q reduce average vehicle travelling time by percentage ranging from 15% to 25% respectively. However, average travelling time under ARD-BP-Q is only slightly lager than SP-BP at low vehicle arrival rates. Because under low vehicle arrival rates do not make a lot of congestion. The better choice should be SP-BP.

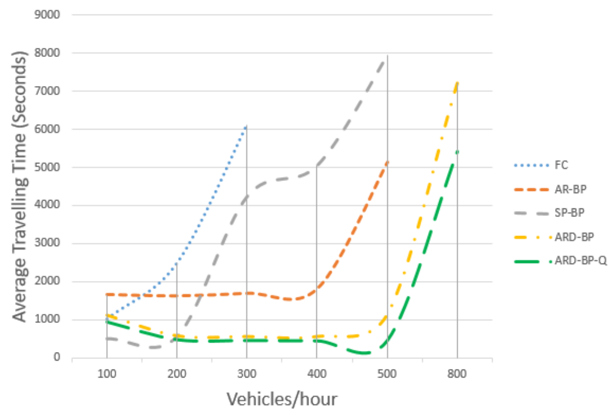

Fig. 6 Average vehicle travelling time with increasing vehicle arrival rate under different algorithms. For AR-BP, ARD-BP and ARD-BP-Q, α = 2.5.

Fig. 7 shows simulation results of average vehicle speed. From the figure, we can see that average vehicle speed under ARD-BP-Q is higher than other algorithms when vehicle arrival rate is more than 200 vehicles/hour, indicating that vehicles under our algorithm ARD-BP-Q are less congested.

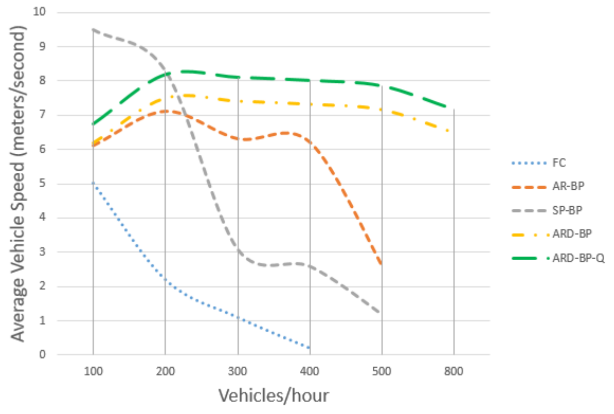

Fig. 7 Average vehicle speed with increasing vehicle arrival rate. For AR-BP, ARD-BP and ARD-BP-Q, $\alpha = 2.5$.

Fig. 8 shows simulation results of average number of vehicles in network. This figure shows that the number of vehicles in road network under ARD-BP-Q algorithm is smaller than other algorithms, meaning less traffic congestion. Fig. 9 shows that more vehicles can arrive at destinations under our algorithm ARD-BP-Q, meaning that more vehicles under other algorithms get stuck in road network.

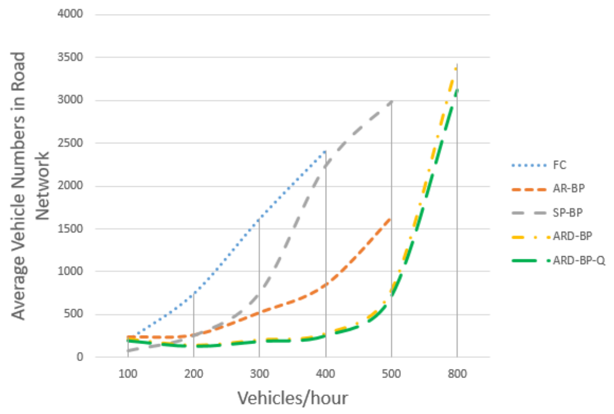

Fig. 8 Average number of vehicles in road network with increasing vehicle arrival rate. For AR-BP, ARD-BP and ARD-BP-Q, $\alpha = 2.5$.

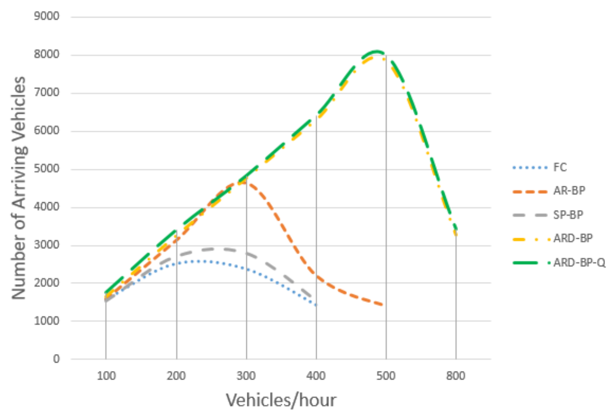

Fig. 9 Number of vehicles arriving at destinations. For AR-BP, ARD-BP and ARD-BP-Q, $\alpha = 2.5$.

Finally, we run simulations to check the effect of α on ARD-BP-Q performance. The result is summarized in Fig. 10. We can see that there exists the optimal value of α such that ARD-BP-Q achieves the lowest average traveling time.

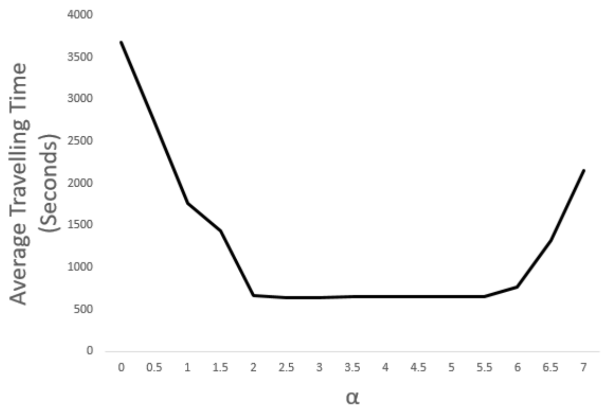

Fig. 10 Effect of parameter α on average travelling time under ARD-BP-Q. Vehicle arrival rate is set to be 450 vehicles/hour.

In summary, our Adaptive Traffic Control Algorithm Based on Back-Pressure and Q-Learning greatly reduces traffic congestion as verified by simulations.

5. CONCLUSIONS

In this paper, we proposed an adaptive traffic control algorithm based on back-pressure and Q-learning. Our algorithm controls traffic based on accurate real-time traffic information (achieved by using shadow network) and global traffic information (achieved by using Q-learning). Our algorithm greatly reduces traffic congestion and vehicle traveling time as verified by simulations and is superior to other four algorithms.

Acknowledgments This work was supported by JSPS KAKENHI Grant Number 16K01288 and Hoso Bunka Foundation, Japan.

References

[1] L. Li, D. Wen and D. Yao, " A survey of traffic control with vehicular communications, " IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 1, pp. 425–432, Feb. 2014.
[2] A. Ejaz and H. Gharavi, "Cooperative Vehicular Networking: A Survey," IEEE Transactions on Intelligent Transportation Systems, vol.19, no.3, pp: 996–1014, 2018.
[3] D.I. Robertson, I. Dennis and R.D. Bretherton, "Optimizing networks of traffic signals in real time-the SCOOT method," IEEE Transactions on vehicular technology, vol.40, no.1, pp. 11–15, 1991.
[4] P.R. Lowrie, Scats, sydney co-ordinated adaptive traffic system: A traffic responsive method of controlling urban traffic, Darlinghurst, NSW, Australia : Roads and Traffic Authority NSW, Traffic Control Section, 1990.
[5] P. Mirchandani and L. Head, " A real-time traffic signal control system: architecture, algorithms, and analysis, " Transportation Research Part C: Emerging Technologies, vol. 9, no.6, pp. 415–432, 2001.
[6] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos and Y. Wang, " Review of road traffic control strategies, " Proceedings of the IEEE, vol. 91, no.12, pp. 2043–2067, 2003.
[7] K. Aboudolas, M. Papageorgiou and E. Kosmatopoulos, " Store-and forward based methods for the signal control problem in large-scale congested urban road networks, " Transportation Research Part C: Emerging Technologies, vol. 17, pp. 163–174, 2009.
[8] C. Diakaki, M. Papageorgiou and K. Aboudolas, " A multivariable regu- lator approach to traffic-responsive network-wide signal control, " Control Eng. Pract, vol. 10, no.2, pp. 183–195, 2002.
[9] L.B. Oliveira and E. Camponogara, " Multi-agent model predictive control of signaling split in urban traffic networks, " Transp. Res. C, Emerging Technol, vol. 18, no.1, pp. 120–139, 2010.

[10] H. Ezawa and N. Mukai, " Adaptive traffic signal control based on vehicle route sharing by wireless communication, " International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, vol. 6279, pp. 280–289, 2010.
[11] T. Tettamanti, T. Luspay, B. Kulcsr, T. Peni and I. Varga, " Robust-control for urban road traffic networks, " IEEE Trans. Intell. Transp. Syst, vol. 15, no. 1, pp. 385–398, 2014.
[12] I. Alvarez, A. Poznyak and A. Malo, " Urban traffic control problem a game theory approach, " in Proc. IEEE Conf. Decision Control, pp. 2168–2172, 2008.
[13] T. Le, P. Kovcs, et al., " Decentralized signal control for urban road networks, " Transp. Res. C, Emerging Technol, vol. 58, pp. 431–450, 2015.
[14] T. Wongpiromsarn, T. Uthaicharoenpong, Y. Wang, E. Frazzoli and D. Wang, " Distributed traffic signal control for maximum network throughput, " in Proc. IEEE ITSC, pp. 588–595, 2012.
[15] J. Gregoire, E. Frazzoli, A. de La Fortelle and T. Wongpiromsarn, " Back-pressure traffic signal control with unknown routing rates, " IFAC Proceedings Volumes, vol. 47, no. 3, pp. 11332–11337, 2014.
[16] J. Gregoire, X. Qian, E. Frazzoli, A. deLaFortelle and T. Wongpiromsarn, " Capacity-aware back-pressure traffic signal control, " IEEE Trans. Control Netw. Syst, vol. 2, pp. 164–173, 2015.
[17] Kulcsr, K. Ampountolas and A. Dabiri, " Single region robust perimeter traffic control, " in Proc. IEEE Eur. Control Conf, pp. 2628–2633, 2015.
[18] A. Zaidi, B. Kulcsr and H. Wymeersch, " Back-pressure traffic signal control with fixed and adaptive routing for urban vehicular networks, " IEEE Transactions on Intelligent Transportation Systems, vol.17, no.8, pp. 2134–2143, 2016.
[19] H. Chai, H. M. Zhang, D. Ghosal and C. N. Chuah, " Dynamic traffic routing in a network with adaptive signal control, " Transportation Research Part C: Emerging Technologies, vol. 85, pp. 64–85, 2017.
[20] E. Athanasopoulou and L. Bui, et. al., "Back-pressure-based packet-by-packet adaptive routing in communication networks," IEEE/ACM transactions on networking, vol. 2, no. 1, pp. 244–257, 2013.
[21] S. Kim, M.E. Lewis and C. White, " Optimal vehicle routing with real- time traffic information, " IEEE Trans. Intell. Transp. Syst, vol. 6, no.2, pp. 178–188, 2005.
[22] L. Tassiulas and A. Ephremides, " Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks, " IEEE Transactions on Automatic Control, vol.37, no.12, pp. 1936–1948, 1992.
[23] D. Krajzewicz, G. Hertkorn, C. Rssel and C, P. Wagner, " SUMO (Simulation of Urban MObility)-an open-source traffic simulation, " in Proceedings of the 4th middle East Symposium on Simulation and Modelling, 2002, pp. 183–187.
[24] Y. Liu, J. Gao and M. Ito, "Back-Pressure Based Adaptive Traffic Signal Control and Vehicle Routing with Real-Time Control Information Update," 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Madrid, 2018, pp. 1–6.