UC Davis

Title

Dynamic traffic routing in a network with adaptive signal control

Permalink

https://escholarship.org/uc/item/7zf86446

Authors

Chai, Huajun
Zhang, HM
Ghosal, Dipak
et al.

Publication Date

2017-12-01

DOI

10.1016/j.trc.2017.08.017

Copyright Information

Peer reviewed

CrossMark

Dynamic traffic routing in a network with adaptive signal control

Huajun Chai [a], H.M. Zhang [a,*], Dipak Ghosal [b], Chen-Nee Chuah [c]

[a] Department of Civil and Environmental Engineering, University of California, Davis, Davis, CA 95616, USA
[b] Department of Computer Science, University of California, Davis, Davis, CA 95616, USA
[c] Department of Electrical and Computer Engineering, University of California, Davis, Davis, CA 95616, USA

ARTICLE INFO

ABSTRACT

In real traffic networks, travellers' route choice is affected by traffic control strategies. In this research, we capture the interaction between travellers' route choice and traffic signal control in a coherent framework. For travellers' route choice, a VANET (Vehicular Ad hoc NETwork) is considered, where travellers have access to the real-time traffic information through V2V/V2I (Vehicle to Vehicle/Vehicle to Infrastructure) infrastructures and make route choice decisions at each intersection using hyper-path trees. We test our algorithm and control strategy by simulation in OmNet++ (A network communication simulator) and SUMO (Simulation of Urban MObility) under several scenarios. The simulation results show that with the proposed dynamic routing, the overall travel cost significantly decreases. It is also shown that the proposed adaptive signal control reduces the average delay effectively, as well as reduces the fluctuation of the average speed within the whole network.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Regular vehicles like the ones most of us are driving today don't have real-time road traffic information. The routing strategies used are either based on past experience or based on limited local information. This leads to the fact that regular vehicles can't respond to road incidents in a timely manner to avoid long delay. However, with the emergence of connected vehicles technology it is possible to access both local and global real-time traffic information via the V2X infrastructure. It is an important and challenging research problem to study how to take advantage of the extraordinarily rich information we can get from the connect vehicle system. The main objective of this paper is to study the interaction between two major components: dynamic vehicle routing and adaptive traffic signal control in a connected vehicle environment. We consider different combinations of route choice strategies and various traffic signal control methods to obtain an effective joint vehicle routing and signal control scheme which will reduce the average travel time within the network.

Vehicle routing problems have been very well studied over the years. A great part of the existing papers studies how to route vehicles in a network efficiently to meet some constraints like location routing (Perl and Daskin, 1985; Min et al., 1998; Nagy and Salhi, 2007) or time constrained routing (Desrosiers et al., 1984, 1995; Solomon and Desrosiers, 1988; Nie and Wu, 2009). The routing problem in this paper is focused particularly on how to route vehicles in the network so that the total time

---

for the vehicles to get to their destinations can be minimized. The underlying problem for that is the well-known shortest path problem (SPP).

In the study of shortest path problem, the existing work can be categorized in two main groups: deterministic shortest path problem (DSPP) and stochastic shortest path problem (SSPP). Bellman (1956) proposed a dynamic programming method to solve the optimal route from one point to another with all link travel time to be deterministic and known ahead of time. SSPP is more interesting than DSPP as in real word link cost is usually not known deterministically but has many uncertainties. In stochastic scenarios, the shortest path problem can be further categorized into two groups according to Gao and Chabini (2006): path problem and optimal routing policy problem. Path problem aims to find a specific path (a deterministic link set) to destination that will attain a certain objective, such as least expected travel time (Miller-Hooks and Mahmassani, 2000) or maximum on-time arrival probability (Fan et al., 2005). Optimal routing policy problem, however, is more complicated than the path problem. According to Gao and Chabini (2006), a routing policy is defined as a decision rule that specifies which node to take next at each decision node based on realized link travel times and the current time. Compared with path problem, optimal routing policy problem in most cases can give a routing solution that is more efficient and reliable. The reason for that is because as the traveller travels in the network, he gains knowledge of the network (in this example, the travel time experienced after the traveller traverses a link). With this knowledge and the dependency of links being known, his anticipation of the future can be changed. The traditional path finding method does not take advantage of the newly learned information and the dependency between links. In order to take these advantages to improve the routing decision, the aforementioned optimal routing policy is needed. The optimal routing policy will not give one fixed path but a decision tree that will guide the traveller to the next node based on the current state(in most of the cases it is the arrival time at the current node) at the decision node. The optimal routing policy is particularly useful in stochastic and time-dependent networks. This routing policy is sometime known as hyper-path in some literature. Miller-Hooks and Mahmassani (2000) studied the least expected travel time problem using hyper-path algorithm in a stochastic and time-varying network. Fu (2001) also studied an adaptive routing approach with real-time information. Chen and Nie (2015) studied the stochastic optimal routing problem for vehicles with a limited travelling limit. The problem is formulated as a two-stage stochastic shortest path problem: both stages are a stochastic shortest path problem respectively. A label correcting based algorithm is used to solve the problem. However, their model does not consider time-variant link cost. Wu (2015) studied the travel reliability as an extension to the traditional shortest path problem in stochastic and time-dependent networks by adding the standard deviation to the mean travel time to represent the reliability of a certain route. However, the work does not explicitly consider the time-dependent problem in their formulation.

For signal control, most of the existing papers do not consider the interaction with traffic routing. The traditional control method, no matter adaptive or fixed-time, isolated or coordinated, only aims to reduce the delay or maximize the throughput of the intersection with known and perhaps time-dependent traffic demand (Rosdolsky, 1973; Hunt et al., 1982; Lo, 1999; Mirchandani and Head, 2001; Choy et al., 2003; Tatomir and Rothkrantz, 2004; Cheng et al., 2006; Haddad et al., 2013). This in reality might cause inefficient traffic routing, as traffic may oscillate between different routes due to the impact on travel delay caused by the signal control. Vehicles may make unnecessary reroute to avoid a red light at a certain intersection: for example when they see a red signal for the through movement, they may change their route and take a right turn in order not to wait at that intersection. This myopic behavior not necessarily guarantees to reduce the total travel time as the traveler may experience more red light stops (more delay) at the downstream intersections. It also causes fluctuations of road traffic as travelers are switching their routes too frequently. The fluctuation of road traffic has many negative effects on traffic control (Horowitz, 1984; Friesz, 1985; Zhang and Nagurney, 1996). So a good signal control strategy should take the interaction with vehicle routing behavior into consideration to achieve a better overall performance: not only reducing average travel time, but also maintaining a relatively stable on-road traffic. It requires the work to integrate traffic routing together with signal control. The problem is no longer a simple shortest path problem but a more complicated time-constrained shortest path problem (TCSPP).

In early literature, there have been many papers trying to solve the combined traffic assignment and traffic control problems, known as CTAC. Smith and Ghali (1990) studied the dynamics of traffic assignment and traffic control. When demand was constant, they were able to get a steady (equilibrium) state, but for dynamic demand they were unable to obtain such results. Later on, Yang and Yagar (1995) formulated the CTAC problem in a bi-level programming formulation. The upper level tried to minimize the system cost by varying the signal settings while the lower level modeled travellers' routing behavior which will give an equilibrium flow state given the signal settings. They also proposed an efficient method to solve the bi-level optimization problem. All the aforementioned papers were focusing on static networks. In reality traffic flows are varying from time-to-time and from day-to-day. A more sophisticated model should be used in order to solve the real world problem. Xiao and Lo (2014) formulated a joint dynamical traffic system that considered both travellers' route choices and traffic light control. The dynamical system permitted the signal controller to interact with and adapt to route choices of travellers, and vice versa in a day-to-day setting. Zaidi et al. (2015) applied the back-pressure algorithms from communication networks to the traffic network with adaptively re-routing traffic.

The aforementioned papers that jointly consider traffic assignment and signal control are all path based, and most of them consider static user equilibrium rather than adaptive control, which means they are unable to provide real-time routing guidance to individual travellers. In real world applications, it is more useful and has more significant impact to travellers if a system can offer good routing guidance at an individual level. This brings attention to combined traffic routing and signal control. Chen and Yang (2000) studied the shortest path problem in traffic-light networks. The constraints in their work were

called time-windows which were actually a mathematical representation of phases in traffic light cycles. Their shortest path algorithm could achieve a time complexity of $O(r \times n^3)$, where n denotes the number of nodes in the network and r is the number of different time-windows in a node. However, this work only considered deterministic and static link travel times. Yang and Miller-Hooks (2004) extended the work of Miller-Hooks and Mahmassani (2000) to incorporate the traffic signal operations. Their network flow was time-varying and stochastic. They studied two different sub-problems: one with actual signal timing known a priori, and the other with probabilistically known signal timings. They proposed algorithms for both problems to solve the adaptive routing in a signalized network. The underlying assumption in their paper was that the distribution of time-varying stochastic link travel times is known a priori. This assumption is too strong when applying to real world scenarios. In reality, travellers might know the distribution of link travel times in the near future (Zhang and Rice, 2003; Chien and Kuchipudi, 2003; Lin et al., 2004). But in a longer scale of time, it is impossible for travellers to know the future link travel time accurately.

The work in this paper addresses the issues of existing methods with the following contributions:

1. We have proposed a dynamic routing strategy that can constantly update travellers' knowledge of link travel time with consideration of adaptive signal control for efficient routing in real world transportation networks. The joint model of dynamic traffic routing and adaptive signal control developed in this paper is shown to reduce the average queue length and average travel time, as well as increase the average speed in the network. It shows significant advantage over the other traditional ways of routing and signal control.
2. We have proposed, tested and compared different signal control methods under different scenarios.
3. We have used the OmNet++ and SUMO simulation platform to study the benefit of VANET on the joint routing and signal control strategy proposed. The effect of using real-time information is studied and evaluated.

The rest of the paper is organized as follows: Section 2 exposes some difficulties and existing problems in routing and signal control that motivate us to study the topic. The challenges to address these issues are also stated later in the same section. The joint routing and control problem is stated and algorithms are proposed in Section 3. Some numerical experiments are designed and carried out in Section 4. Results and analysis are given in Section 5. Section 6 summarizes the main results of this work, and discusses future extensions to the current work.

2. Context and motivation

We have briefly talked about the problems associated with the traditional routing and signal control methods in Section 1. In this section, more details on the existing problems that motivate us to study the topic are discussed and explained. Fig. 1 shows a typical suburban road network in the future that are facilitated by Vehicular Ad hoc NETwork (VANET) technology. It consists of a freeway (in the middle), some main arterial roads and some minor local roads. The vehicles travelling in the network are all equipped with wireless devices that enable them to communicate with each other (V2V technology). Base
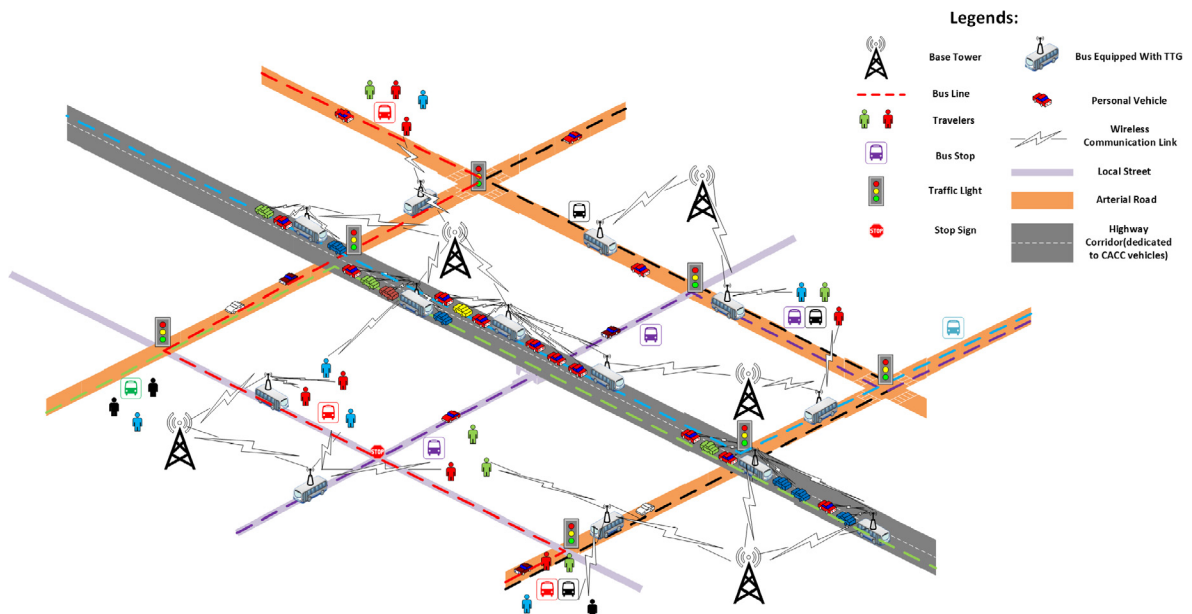

Fig. 1. Connected vehicles in VANET.

towers are served as communication hubs to collect and distribute information to individual vehicles (V2I technology). Technically, the VANET technology provides an opportunity to adopt more sophisticated routing and control methods. Here we list the three most critical aspects of the problems we are facing in joint routing and traffic signal control in real life. These are also the difficulties and challenges that drive us to study the problem.

- **Time-dependency:** The traffic on the road is evolving at all times. A road segment that is not congested during the last hour can become congested in the current hour. Moreover, for the same time-of-day, the same location can be congested today but not congested tomorrow. Such within day and between day fluctuations in traffic conditions and travel times make traffic routing and signal control a difficult task. A dynamic framework that can update the routing decision and signal control timing dynamically should be studied in order to fully consider the time-dependency.
- **Uncertainty:** There are many uncertainties on the road: accidents may happen unexpectedly; disasters like hurricane and flood may destroy and block road segments without early notice. These uncertainties can cause significant delays to travelers if there are no appropriate solutions to deal with them. Uncertainties are hard if not impossible to predict. Instead of predicting the unexpected incidents, we can come up with a good real-time updating framework that can respond to any unexpected incidents in a timely manner.
- **Stochasticity:** Even with the most up-to-date technologies we have today, it is impossible to capture the absolutely precise traffic state. For example, we can use floating cars to estimate the travel time for a certain road segment. But, the penetration rate of such floating cars limits the accuracy of the estimation. Uncertainties discussed previously can also impose difficulties to traffic state estimation. Thus, there is stochasticity residing in the joint routing and signal control problem: link travel time estimation and signal timing estimation. A comprehensive approach which should take careful consideration of such stochasticity is desired to make the routing and signal control strategy reliable, robust and effective.

We show two examples of the problems travelers might encounter if the aforementioned three aspects are not addressed appropriately. As a precursor to this study, we carried out simulations that use traditional routing and signal control methods to observe any possible problems. By traditional routing and signal control methods, we mean that the default traffic routing algorithm (deterministic shortest path algorithm) and the default signal control methods (fixed timing or simple adaptive control) implemented by SUMO are used in the simulation. Other inputs and parameters are the same as the simulation setups described in Section 4.

- **Longer travel time:** When we run simulations using the traditional routing algorithm, we observe that many travellers are not taking the "optimal" routes in terms of expected total travel time to destinations. The reason is that travellers have no access to the most up-to-date travel time information when they plan their routes, and the routes deemed optimal earlier may no longer be optimal later as traffic conditions change. The accuracy of the knowledge of the current and future traffic conditions becomes even more important when traffic signals are considered as one might encounter unexpected delays at intersections. A well designed joint routing and signal control model that can select routes with shorter travel times and avoid waiting at intersections is desired.
- **"Looping" en-route:** Another interesting observation is that a great number of vehicles are travelling back and forth on some links even when new link travel time is available to travellers. They loop in the network and spend much longer time before reaching their destinations. When travellers receive new information and re-plan their route, sometimes they will find that going back to the previous link will be a "best" choice at the moment. Thus, they would make a "U-turn" and travel backwards. This is due to uncertainties in the network and also the inaccuracy in estimating the future network states, especially the traffic signal timing plans and link travel times. The current link travel time alone is usually not a good resource to calculate the shortest path in a stochastic and time-dependent network. In order to better describe the link travel time characteristics, a well designed travel time updating model that considers both current information and historical information is needed.

With the help of rapidly developing VANET technology, it becomes possible to gain on-line information of on-road traffic like queues, average speed, delay at intersections and so on. In the future, most of the on-road vehicles will be equipped with wireless devices which will enable the vehicles to communicate with each other and with the control centers. These vehicles are known as Connected Vehicles. They use V2V (vehicle-to-vehicle) and V2I (vehicle-to-infrastructure) technologies to share information in the network. There will be two networks: one is the traditional physical transportation network, and the other is the so-called VANET that uses connected vehicles as mobile nodes. With real time traffic information being available, dynamic routing and adaptive signal control become possible. In our work, we also consider a stochastic and time-varying network. We propose a strategy that will constantly update the knowledge of network, then use this new information to update travellers' route choice. A hyper-path based solution is developed in this work. We also propose five different adaptive signal control strategies (including one method from other's work and one modified based on it) that are interacting with the dynamic routing to achieve a better system performance, which most of the existing papers do not consider.

3. Joint traffic routing and signal control

We have a general transportation network, which is denoted using graph representation $G = (V, A)$. V is the set of nodes in the network, $|V| = v$; A is the set of edges in the network, $|A| = m$. $V_s \subset V$, is a subset of nodes that are controlled by the traffic lights. $\tau_{ij}(t)$ is the link travel time for link (i, j) at time t. It is a time variant variable. In reality, link time cost is determined by many factors, including background traffic flow, link condition, driver's preference and so on. Table 1 shows all the symbols used in the following sections and their corresponding definitions.

There are two components interacting with each other in our problem: dynamic traffic routing and adaptive signal control. Each of these two components has specific inputs, and yields outputs which can be inputs to the other component. Fig. 2 illustrates the whole system framework for joint traffic routing and signal control. The simulations discussed in the later sections are all designed based on this system framework.

3.1. Dynamic traffic routing

There are many different adaptive routing strategies having been studied in the literature. To keep the problem as simple as possible first, this study uses the strategy stated as follows: when a vehicle arrives at a node, it will get updated link travel time, queuing length and traffic signal plans at the time it arrives. Based on that new information, the traveller will perform a new shortest path calculation (the algorithm used is given in the following parts of this section), and follow the new calculated shortest path thereafter until he reaches the next node. However, this procedure doesn't have to be performed every node for the purpose of calculation efficiency. Since the routing strategy here is changing dynamically over time, we refer it as *Dynamic Traffic Routing (DTR)* in the following sections in this paper. In contrast, *Adaptive Routing* without dynamic information updating is simply referred as *AR*.

In a general transportation network setting, traffic conditions are changing constantly. In the old days when real-time traffic information is not available, travellers plan their routes before their trips using the information that is available at that time. Once a route is selected, the choice is fixed. When some unexpected incidents happen on the links in the route, the travellers will experience longer than expected travel time. However, with the real-time traffic information being available, travellers are able to change their routes en-route. Their knowledge of traffic information will keep refreshing. Once the knowledge is updated, routes can also be updated.

Traditional routing usually uses deterministic shortest path algorithm to obtain new route. A well-known and widely used shortest path algorithm is Dijkstra's Algorithm (Dijkstra, 1959). A simple Dijkstra's Algorithm can be formulated as below:

$$\lambda_i = \min_{j \in \Gamma(i)} \{\lambda_j + \tau_{ij}\}$$

where λ_i is the minimum travel time from node i to destination; τ_{ij} is the link cost for link $(i, j$. This formulation assumes that link travel time is constant during the entire trip. However, this assumption is not realistic in most real world scenarios. Link cost varies from time-to-time and from day-to-day. It is intuitive to formulate link travel time τ_{ij} as a function of time t. So the link cost for link (i, j) at time t will be $\tau_{ij}(t)$. A modified Dijkstra's Algorithm is then

$$\lambda_i(t) = \min_{j \in \Gamma(i)} \{\lambda_j(t + \tau_{ij}(t)) + \tau_{ij}(t)\}$$

3.1.1. Stochastic link travel time

In the formulation above, we can see that calculation of the minimal travel time cost to destination depends on the travel cost to destination, i.e. $\lambda_j(t + \tau_{ij}(t))$, of the nodes that have not yet been traveled. This requires an estimation of the future link

Table 1
Symbols used in the paper.

| Symbol	Definition
A	Set of edges
V	Set of nodes
v	Number of nodes
m	Number of edges
V_s	Set of nodes that are controlled by traffic lights
$\tau_{ij}(t)$	Travel time on link (i, j) at time t
λ_i	Minimum cost from node i to destination node
λ_i^h	Minimum cost from node i to destination node with upstream node to be node h
π_i	Previous node of node i in the shortest path
ϕ	Intersection delay
$\rho_{ij}(t)$	Probability of link travel time $\tau_{ij}(t)$
ω_i^l	The lth phase at intersection i
Ω_i	The set of phases at intersection i
$\Gamma(i)$	The set of adjacent nodes of node i
T	The time horizon for stochastic and time-varying network, after which the network becomes static
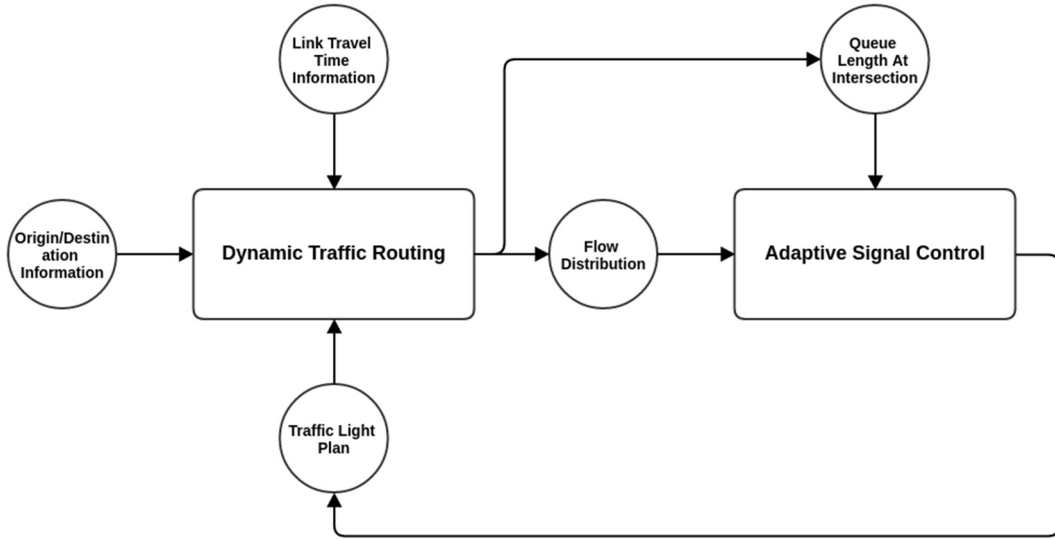
Fig. 2. Interactions between dynamic traffic routing and adaptive signal control.

cost. Due to the randomness and uncertainties mentioned in the previous sections, this estimated link cost subjects to some stochastic fluctuations. It is straightforward to formulate the problem in a stochastic manner.

A link (i, j) can have K possible costs:

$$\tau_{ij}^k(t), \quad k = 1, 2, \ldots, K.$$

Each of these link costs $\tau_{ij}^k(t)$ will have an associated probability:

$$\rho_{ij}^k(t), \quad k = 1, 2, \ldots, K.$$

3.1.2. Hyper-path based stochastic shortest path

Taking the stochastic link cost into consideration, the formulation for the hyper-path based shortest path formulation becomes:

$$\lambda_i(t) = \min_{j \in \Gamma(i)} \left\{ \sum_{k=1}^{K} \left\{ \left[\lambda_j(t + \tau_{ij}^k(t)) + \tau_{ij}(t) \right] \times \rho_{ij}(t) \right\} \right\} \tag{1}$$

The solution of the problem is a hyper-path (Miller-Hooks, 2001). Fig. 3 shows an example of a hyper-path tree. For every node, we maintain two different vectors: (1) the vector of next node to take at every time step; (2) the vector of the cost from current node to destination at every time step. By tracking down the tree by the labels, one can retrieve the shortest path to destination for any time step.

3.1.3. Link travel time updating

The fact that a dynamic shortest path algorithm outperforms a static shortest path algorithm is because the dynamic algorithm can use the most up-to-date traffic information to make the shortest path calculation more accurate. In order to achieve this, a dynamic link travel time updating scheme is needed. In the setting of this paper, every link maintains a set of possible link travel time realizations as well as their corresponding probabilities. After a certain period of time, every link will have a set of new link travel time realizations (they can be new values or values that are overlapped with the initial set). An updating scheme will merge the old set and the new set.

Fig. 4 shows the updating method used in this paper. Travellers will have different preferences towards old information and newly received information. It is characterized by the weights, i.e. b and c respectively, shown in the figure. This preference can be affected by travelers' experience, traffic conditions, time of day and so on. In the example in Fig. 4, weight for old information is $b = 0.5$, and weight for new information is $c = 0.5$. Updated travel time distribution can be calculated using the following equation:

$$\rho^{updated}(t = \tau) = \begin{cases} b \cdot \rho^{old}(t = \tau), & \text{if } \tau \in T^{old} \\ c \cdot \rho^{new}(t = \tau), & \text{if } \tau \in T^{new} \\ b \cdot \rho^{old}(t = \tau) + c \cdot \rho^{new}(t = \tau), & \text{if } \tau \in T^{old} \cap T^{new} \end{cases} \tag{2}$$
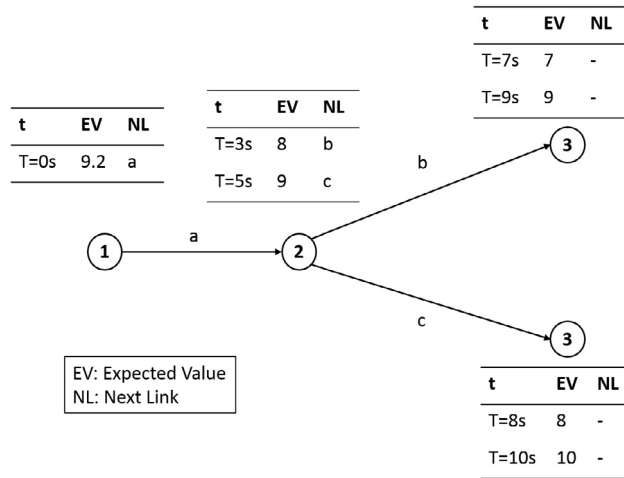
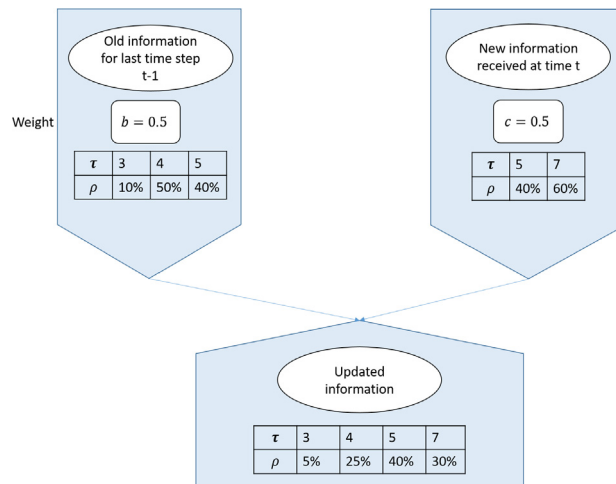**Fig. 3.** An example for hyper-path tree.

Fig. 4. An example for travel time updating.

$$T^{updated} = T^{old} \cup T^{new} \tag{3}$$

where T^{old}, T^{new} are the sets of old and new link travel times.

3.2. Adaptive signal control

Our research work considers a VANET environment which means the traffic information can be very detailed. The most important feature is that this information is real-time. In the old days, accurate real-time traffic information is not always available. Most traffic research papers use the so-called "forecast" or "predicted" data to design traffic control scheme. Due to the nature of the inaccuracy of the data source, the control algorithm is sometimes not working effectively especially under congestion scenarios in which traffic usually has larger unpredictability (Noland and Polak, 2002; Zheng and Van Zuylen, 2010).

With real-time traffic information being available, it now becomes possible for real-time adaptive signal control. Different adaptive control methods are proposed and tested in this study. The first one is a low-density control algorithm which is designed for low traffic volume situations. The second is a high-density algorithm which is for more congested situations. There is also a third control method called *Phase Selection Control* algorithm which chooses a phase candidate to switch to and determines its green duration to give the maximum throughput of the intersection. Varaiya (2013) proposed a traffic control algorithm called *Max Pressure* control. Their control algorithm selects the "optimal" phase based on the calculated

pressure for each stage. A stage that has the maximum pressure is then selected as the next stage for the intersection. In their paper, the pressure term is determined by the queues on the in-coming links, queues on the out-going links and the saturation flow rates for those links. Details on the model can be found in Appendix A and in Varaiya (2013). They claimed that their control method can achieve a stable network state in terms of average queue length in the network. However, there are several shortcomings in their work that need remedy. The way they determine green time split is too simple. By assigning most of the green time to the selected "optimal" phase will increase queue length on the approaches that are not permitted to pass in that phase, and hence increase travel time. This becomes quite significant especially in a homogeneously congested scenario in which each approach has almost the same demand. Another aspect is that their algorithm fails to consider upstream demand which is a very important component for designing an adaptive signal control algorithm, especially if coordination is considered. To address those problems, we further propose a control method called *Modified Max Pressure* control based on the aforementioned *Phase Selection Control* and *Max Pressure Control*, which considers pressures derived from both the queues at the current intersection and the upstream traffic demands. All the algorithms mentioned above are described in details in the following sections.

3.2.1. Low density control

When the traffic demand in the network is low, a low density control algorithm is used to control traffic. The algorithm is a fully vehicle-actuated control method. Traffic light will turn green for a certain approach if there are vehicles detected on the lanes of that approach (detectors are located at both ends of a certain link. See Fig. 8). Green time extension is also based on traffic actuation, which means if there are vehicles coming continuously the green time will extend correspondingly. Green time is subjected to minimum and maximum green times.

Our low density control algorithm works as Fig. 5:

3.2.2. High density control

When the traffic demand in the network is high, a high density control algorithm is used to control traffic. In this algorithm, phase split is set to be proportional to the incoming flow for each phase. (More rigorously, flow ratio, i.e. flow/saturation flow, should be used. However, in our scenario saturation flow rates are the same for each phase.) Before every updating period for the traffic signal timing, the incoming flows for every phase are collected by the detectors located on the lanes. Minimum and maximum green times are still respected in this case.

Our high density control algorithm works as Fig. 6:

3.2.3. Phase selection control

The aforementioned control algorithms do not change the phase sequence. In some cases, the choice of the next permitted phase can have a great effect on the traffic. Therefore, it is important to choose a good phase sequence for the time-being other than simply setting it to be fixed. The phase selection control proposed in this paper simultaneously determine the next permitted phase and the duration of that phase as following formulation states:

$$(\omega_i^l, G_{\omega_i^l}(t)) = \arg \max_{(\omega_i^l \in \Omega_i, t_g \in [G_{min}, G_{max}])} \left\{ \frac{N_{t_g}^{\omega_i^l}}{t_g} \right\} \qquad (4)$$

Note: $N_{t_g}^{\omega_i^l}$: is the estimated number of vehicles that will arrive at intersection i within time t_g that are also permitted in phase ω_i^l.
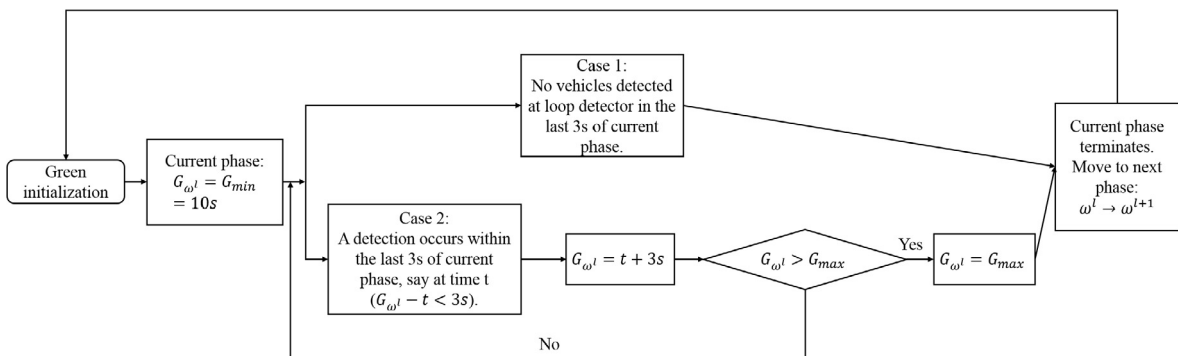

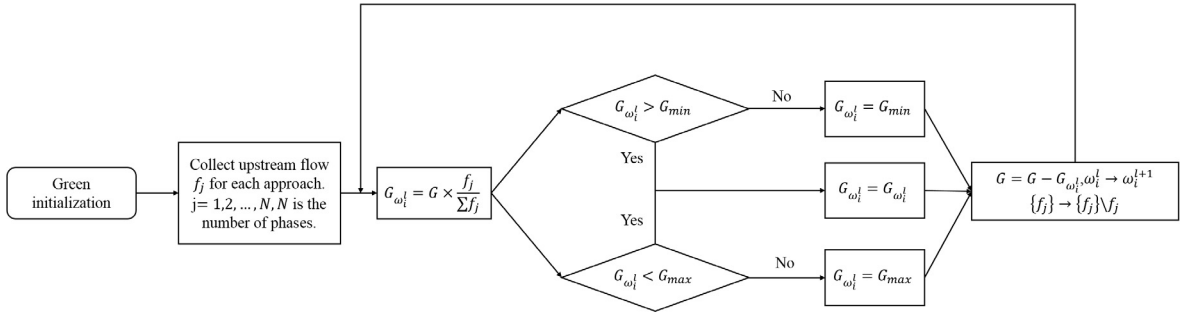**Fig. 5.** Low-density signal control algorithm.

Fig. 6. High-density signal control algorithm.

3.2.4. Modified max pressure control

As mention previously, a *Modified Max Pressure Control* is proposed as follows:

$$(\omega_i^l, G_{\omega_i^l}(t)) = \arg \max_{(\omega_i^l \in \Omega_i, t_g \in [G_{min}, G_{max}])} \left\{ \frac{\alpha \times N_{t_g}^{\omega_i^l} + \beta \times \gamma(\omega_i^l)}{t_g} \right\} \tag{5}$$

$$\gamma(\omega_i^l) = \sum_{n,m} c(n,m)w(n,m)S(n,m) \\ = \sum_{n,m;S(n,m)=1} c(n,m)w(n,m) \tag{6}$$

$$w(n,m) = x(n,m) + \sum_{p \in Out_m} \gamma(m,p)[d(m,p) - x(m,p)] \tag{7}$$

Note:

$N_{t_g}^{\omega_i^l}$: is the estimated number of vehicles will arrive at intersection i within time t_g that are also permitted in phase ω_i^l.
α, β: are weights on upstream flow and queue pressure, respectively.
m, p: link m and link p.
$c(n,m)$: saturation flow of movement (n,m), vehicles per period.
$\gamma(\omega_i^l)$: the portion of vehicles passing intersection i in phase ω_i^l. The $\sum$ in Eq. (6) sums up the number of passing vehicles in all the permitted movements (n,m) that belong to phase ω_i^l.
$d(m,p)$: the length of the corresponding lane on link m for the connection between link m and link p.
$x(m,p)$: the queue length on the corresponding lane on link m for the connection between link m and link p.
$w(n,m)$: the queue pressure for movement (n,m).

For comparison purpose, the original MP Control is included in the Appendix A. From the formulations for Phase Selection Control and Modified MP Control, one can see that the main difference between these two is the queue length term. Phase Selection Control only considers upstream demand to determine the phase sequence and duration, while Modified MP Control considers both upstream demand and queue length.

The *Modified Max Pressure Control* algorithm is presented as Algorithm 2 in Appendix A.

3.3. A combined adaptive signal control and DTR

Now we have a proper dynamic routing algorithm as well as several signal control schemes (low-density control, high-density control, phase selection control, max pressure control or modified max pressure control). While planning their routes, travellers now not only have to consider travel time on each link but also the delay caused by signal control at intersections. With that to be said, we adopted the non-adaptive hyper-path approach from Yang and Miller-Hooks, 2004 to incorporate the signal control methods and the Bayesian link travel time updating scheme proposed in the previous context to explicitly take into account signal delays and link travel time dynamics:

$$\mu_i^h(t) = \min_{j \in \Gamma(i)} \left\{ \sum_{k=1}^{K} \left\{ \left[\phi_{ij}^{\omega_i^l}(t) + \tau_{ij}^k \left(t + \phi_{ij}^{\omega_i^l}(t) \right) + \lambda_j^i \left(t + \phi_{ij}^{\omega_i^l}(t) + \tau_{ij}^k \left(t + \phi_{ij}^{\omega_i^l}(t) \right) \right) \right] \cdot \rho_{ij}^k \left(t + \phi_{ij}^{\omega_i^l}(t) \right) \right\} \right\} \tag{8}$$

$\mu_i^h(t)$: is the temporary label (travel cost to the destination) of node i at time t, the upstream node is h.

$\lambda_j^i(t)$: is the current label (travel cost to the destination) of node j at time t, the upstream node is i.

$\tau_{ij}(t)$: is the link travel time of link (i, j) at time t.

$\phi_{ij}^{\omega_i^l}$: is the delay caused by the signal light at intersection i at time t, ω_i^l is the lth phase of intersection i, the downstream node is j.

$\Gamma(i)$: is the set of all adjacent nodes of node i.

With the adaptive signal control methods and periodically updated link travel times, the traffic routing algorithm become adaptive and dynamic. The adaptive algorithm this study uses is described in Algorithm 1 in Appendix A:

3.4. Computation problems

Consider the DTR algorithm alone, the computational complexity is $O(V^2 \cdot T)$ per vehicle per re-routing if implemented using the priority-queue data structure. If we adopt vehicle-based DTR implementation, i.e. every vehicle will perform its own re-routing calculation, the complexity will be $O(V^2 \cdot T \cdot N)$, where N is the number of vehicles running in the network. When a network gets very large, the number of nodes (V) and number of vehicles (N) could also get very large which can make the computation so slow that the results it provides are not suitable for real-time applications.

One possible approach to conquer this computational barrier is to use O-D based DTR algorithm instead of vehicle-based. O-D based DTR means we compute a stochastic shortest path using the algorithm mentioned in Algorithm 1 for each possible O-D pair every updating interval t_{update}. This updating interval t_{update} can be wisely chosen that it can reduce computation time, while can still yield an acceptable result. Here we make some approximations. We assume the shortest path solution is the same during t_{update} as long as it's the same O-D pair. We store the stochastic shortest path for every O-D pair in a table, and set the table to be active. After t_{update}, old tables will be discarded by simply setting them to be in-active. Whenever a vehicle has a need for re-routing, it can look up in the current active table using its origin and destination information so that it can obtain a route (which is a hyper-path) to guide itself to the destination. In this way, the computational complexity becomes $O(V^2 \cdot T \cdot V)$. For networks with very large number of vehicles (N) and small number of nodes (V), it can reduce the computation time substantially.

4. Simulation tools and parameters

In this section, the DTR algorithm and adaptive signal control methods are tested on a synthetic signalized street network. We run simulations with different settings of parameters to explicitly compare the performance of our DTR algorithm and control methods with the traditional routing and control methods. The metrics we used include average travel time, average speed, average queue length and so on. The goal is to find an "optimal" solution for various traffic conditions. More details on the metrics are discussed later in Section 5.

4.1. Simulation tools

The VANET provides not only real-time traffic information that can be used for traffic routing and signal control purposes, but also the possibility to launch CACC programs so that freeway capacity and safety can be enhanced (Van Arem et al., 2006; Shladover et al., 2012). The work in this paper is based on the VANET setting. We use a simulation platform to test the algorithms and control methods proposed: a microscopic traffic simulation tool called SUMO is used to simulate the physical road network and the vehicles travelling in the network; a communication simulator called OmNet++ is used to simulate the communication network. These two simulators have interfaces that allow them to interact with each other. More information about the simulation platforms can be found here (Arellano and Mahgoub, 2013; Amoozadeh et al., 2015).

4.2. Simulation network

The network used is a 10×3 grid network which replicates a portion of a typical city road network. The middle East-West (EW) road and all the North–South (NS) cross streets are bi-directional. The other two EW roads are one-way roads. The EW roads have a speed limit of 40 mph or 17.78 m/s, while the cross roads have a speed limit of 25 mph or 11.11 m/s. The block size is 500 m long. There are a total of 30 nodes and 47 links within the network. All the intersections are controlled by traffic lights. The geometry and layout of the network are shown in Fig. 7. Each link in the network has double loop detectors located at both entry point and exit point of that link. Fig. 8 shows the location of loop detectors. The detectors can capture the flow and speed of the road traffic.

The default signal timings for the intersections are all fixed control but with different cycle lengths: the traffic lights on the middle main street have a default cycle length of 94 s; the traffic lights at the four corners of the network have a default cycle length of 34 s; the rest of the traffic lights have a default cycle length of 68 s. There is no signal coordination in the network. Through VANET, it assumes that each traveler has precise knowledge of current signal status and timing plans
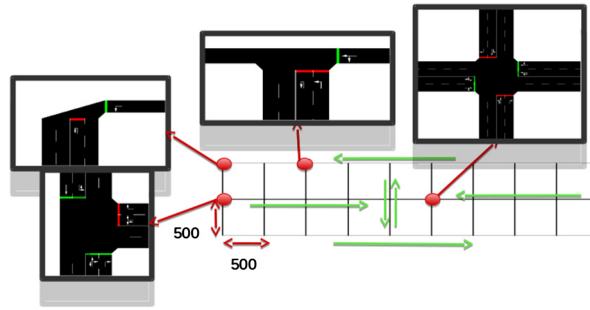
Fig. 7. A synthetic 10×3 grid network used for simulation.

throughout the network. The detailed implementations of the information acquisition process for each traveler, however, is not the focus of this paper and therefore is not described here.

4.3. Input parameters

Vehicles in the network are generated randomly with a uniform distribution in a certain period of time (different traffic demands will have different loading time, which is given in Table 3). The ODs are also randomly generated and assigned to vehicles. All nodes can serve as an origin as well as a destination of a vehicle. Vehicles are generated at nodes. In order to see the performance of the proposed algorithms under different levels of traffic loads, three different levels of traffic demand are used to replicate the free flow, the mildly congested and the heavily congested scenarios. Vehicles that reach their destinations will not re-enter into the network. The simulation terminates when the last vehicle exits the network.

Before the actual run of simulation, a set of small scale pre-run simulations are needed in order to generate the historical set of link travel times which serves as initial inputs for the DTR algorithm. For different traffic demands, the pre-run simulations should use different demand correspondingly. With trial-and-error, we find that usually a number of 10 or more pre-run simulations are needed in order to generate such a historical set. An example of link travel time set is given in Table 2. Between time $t = 0$ and time $t = T$, all links have stochastic link travel times. Each link may have a different number of possible link travel time realizations (in Table 2 all links have 3 different possible link travel times.).

In real world applications, this table is maintained by an information center. The center takes charge of maintaining and updating this table from time to time. In the setting of this paper, the updating frequency of the table is 1 s. Every time step, the center will collect the travel time realizations in the last second on all the links to obtain a new dataset, denoted as I_{new}. The dataset in the old table, denoted as I_{old}, is then updated to a new dataset ($I_{updated}$) using the method explained in Section 3.

$$b \cdot I_{old} + c \cdot I_{new} \rightarrow I_{updated} \tag{9}$$

Every time a vehicle needs to do a re-route, it sends request to the center to request the most up-to-date traffic information. Any new shortest path will be calculated based on that information. The simulation configurations are given in Table 3.

Vehicles are loaded onto the network randomly following a uniform distribution during time T_L. T_L is the vehicle loading time, which is given in Table 3. There are three levels of traffic demand respectively, which are 500 vehicles, 3000 vehicles and 6000 vehicles.

4.4. Simulation scenarios

We design different simulation scenarios to test our algorithms: different combinations of routing strategies and traffic signal control methods under different traffic demand levels.

For the purpose of comparison, the fixed timing control is also implemented. There are three different types of intersections (see network in Section 4). Each has a different timing plan.
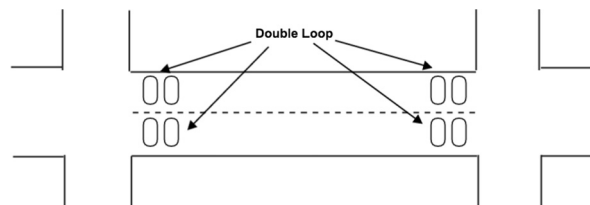

Fig. 8. Each link has double loop detectors at both ends.

Table 2
An example of historical set of link travel time.

		Link 1						⋯	Link N		
t = 0	τ (s)	10	6	8	2	3	4	⋯	4	5	6
	ρ	0.1	0.2	0.7	0.4	0.2	0.4	⋯	0.1	0.2	0.7
t = 1	τ (s)	5	6	7	2	3	5	⋯	5	6	7
	ρ	0.3	0.2	0.5	0.1	0.3	0.6	⋯	0.2	0.4	0.4
⋯	τ (s)	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
	ρ	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
t = T	τ (s)	10	15	20	5	7	9	⋯	8	9	10
	ρ	0.8	0.1	0.1	0.2	0.5	0.3	⋯	0.1	0.5	0.4

Table 3
Simulation parameters.

Link travel time table updating interval	1 s
Vehicle re-routing interval	Per intersection
Vehicle loading time (T_L)	3 min/5 min/10 min
Traffic signal updating interval	120 s

Table 4
Different combinations of simulation scenarios.

Number of Vehicles in The Simulation		Traffic Routing Method		Traffic Signal Control
500		Dijkstra's Algorithm		Fix Timing Control
				Low Density Control
3000	$\otimes$	Adaptive Routing	$\otimes$	High Density Control
				Phase Selection Control
6000		Dynamic Traffic Routing		Modified Max Pressure Control
				Max Pressure Control

- Cycle length = 94 s (see Table 5):
- Cycle length = 83 s (see Table 6):
- Cycle length = 68 s (see Table 7):

For each simulation run, there is a maximum simulation time limit which is preset to 1 h (3600 s). In most cases, all vehicles can reach their destinations before this time limit unless there is gridlock in which vehicles are stalled in the network. In rare cases where there are still vehicles in the network after the time limit, the simulation is forced to terminate. We will show in the following section an example of gridlock under a less effective signal control method.

5. Results and discussion

Results of different simulation scenarios (see Table 4) are analyzed here. To eliminate the effect of stochasticity in our problem due to stochastic link travel time, we run 10 times of simulations for each scenario and analyze the results based on the average of those ten simulation runs from different aspects to study the effect of that particular aspect. It consists of four sub-parts: effects of different traffic signal control methods, effects of different number of re-routing vehicles, effects of link travel time updating and effects on the macroscopic fundamental diagram. The main metrics are average travel time, average travel speed and average queue length in the network. These metrics are good indicators to evaluate the performances among different strategies. Strategies that have smaller travel time, higher travel speed and less queue are favored over the ones that have larger travel time, lower travel speed and more queue. The MFD analysis is to study how the joint traffic routing and signal control algorithm proposed can affect the network throughput in an aggregated way.

In order to see the performance of the proposed methods under some extreme conditions, an accident scenario is also tested and analyzed.

Table 5
Duration of phases for type 1 intersections.

	Phase # 1	Phase # 2	Phase # 3	Phase # 4	Phase # 5	Phase # 6	Phase # 7	Phase # 8
Phase duration (s)	31	5	6	5	31	5	6	5

Table 6
Duration of phases for type 2 intersections.

	Phase # 1	Phase # 2	Phase # 3	Phase # 4	Phase # 5	Phase # 6
Phase duration (s)	31	5	6	5	31	5

Table 7
Duration of phases for type 3 intersections.

	Phase # 1	Phase # 2	Phase # 3	Phase # 4
Phase duration (s)	31	3	31	3

## 5.1. Effects of different traffic signal control methods

First, we study the performance of the four different control methods with DTR: Low Density Control, High Density Control, Phase Selection Control, and Modified Max Pressure Control. Comparisons are made among the four together with Fixed Timing Control and the original Max Pressure Control under the three different traffic demand levels. Various metrics are studied in the following sections.

Figs. 9 and 10 show the results for different control methods under different percentage of travelers who seek re-routing. We can see that Phase Selection Control and Modified MP Control always have the smallest average queue length and largest average speed in all three scenarios with different traffic demands. The performances of Phase Selection Control and Modified MP Control are very similar to each other in all cases. This indicates that in this network, pressure from upstream demand has a dominant impact on designing the "optimal" signal timing, while the queue at an intersection is a less important factor. It is also interesting to note that the performance of Phase Selection and Modified Max Pressure controls are in most cases not sensitive to the number of travelers who engage in re-routing, while this is not true for other controls when traffic demand in the network is heavy. In the latter cases, more re-routing travelers does not necessarily lead to better network performance, an observation consistent with earlier literature on the adoption of Advanced Traveler Information Systems.

In the low demand scenario (500 vehicles), the rest four control methods perform very similarly, with the original MP Control working slightly better than the other three. This is because in such low density, there is very little probability to form a long queue at intersections. Incoming flows can be easily accommodated by even a fixed timing plan in every cycle. The benefit of allowing changing of phase sequence and duration, which is a key point of Phase Selection Control and Modified MP Control, becomes important in this case. That's why these two methods can outperform the rest.

In the mid-range demand scenario, the High Demand Control tends to work better than the Fix Timing Control, the Low Density Control and the original MP Control. In this range of demand, queues are formed from time to time. It is beneficial to give more preference to the movements that have greater demands to dissipate any potential queues by assigning longer green times to the corresponding phases. With the fact that High Density Control works better than the original MP Control, it confirms that upstream demand has a greater impact on designing traffic signal timing than the queue length in this demand level. Again, the Phase Selection and Modified MP control still outperform the other four control methods.

When the network is highly congested, like the case with 6000 vehicles, the performances of the rest four control methods (not including Phase Selection Control and the Modified MP Control) become complex. In this scenario, queues will be formed frequently at intersections, and normally can't be dissipated completely in one cycle when the cycle length is fixed, which is the case of all these four control methods. Performance will be random, and highly depend on the initial input which is stochastic in our simulation.
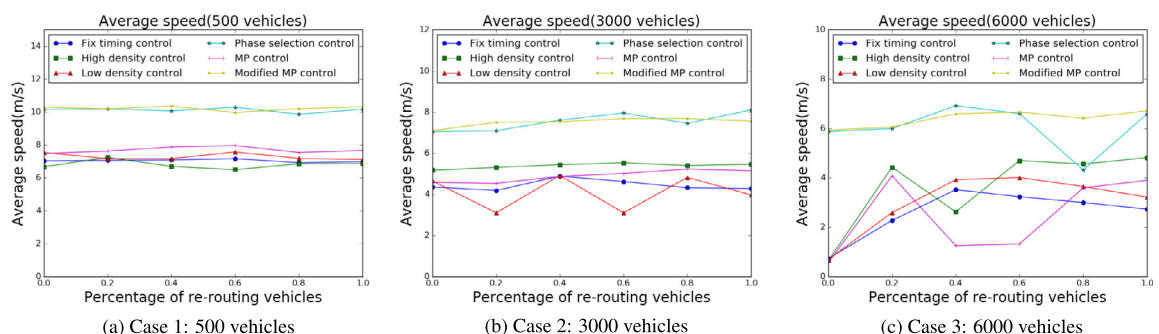
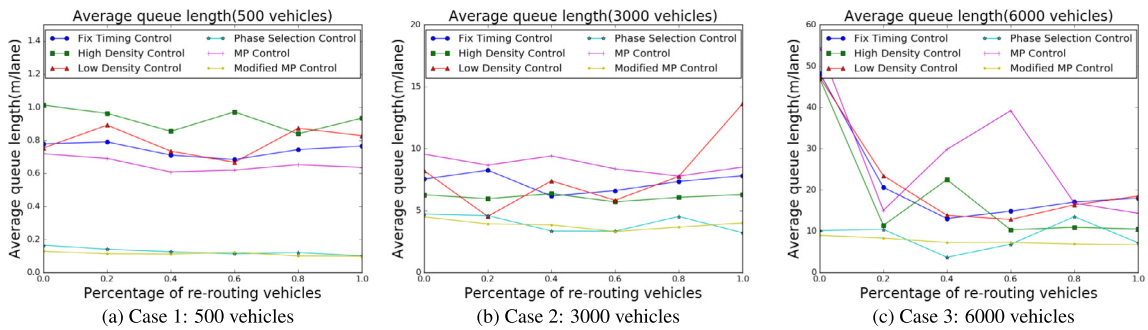**Fig. 9.** Average speed in the network with different number of vehicles.

Fig. 10. Average queue length in the network with different number of vehicles.

By examining the three different traffic demand scenarios, one important conclusion is that upstream demand plays a very important role in designing a good signal control method. Per signalized intersection point of view, the phase sequence and phase duration can be equally important when demand gets high. Failing to consider either one will result in degraded performance.

5.2. Effects of different number of re-routing vehicles

The x-axis in Figs. 9 and 10w is the percentage of re-routing vehicles in the network. Different number of re-routing vehicles can have quite a different effect on the overall performance even when the signal control remains the same. This is intuitive to understand when the percentages are 0% and 100%, representing all vehicles to be normal vehicles and all vehicles are re-routing vehicles, respectively.

In the low density case (The case of 500 vehicles), the average speed in the network is almost the same when the number of re-routing vehicles changes from 0% to 100%. However, the average queue length is slightly decreasing when the number of re-routing vehicles in the network becomes larger. When densities get higher (the cases of 3000 and 6000 vehicles), the relation is no longer monotonic. Average speed is highest, and average queue length is smallest when the percentage of re-routing vehicles is somewhere between 0% and 100%. This shows that in congested scenarios, it is not always true that the more re-routing vehicles the better. In those cases, the network is already packed with vehicles. It is hard for vehicles to move around. Any attempt to change route will incur further congestion burden to the network, which will possibly in turn exacerbate the current congestion level.

Furthermore, if all travellers have access to the same piece of information, after re-routing the result will be that the least congested links will be chosen as their preferred links. Hence traffic will flood to those links making them become congested. If another re-routing is performed, the newly allocated traffic will further shift to some other links that are not congested. So the traffic will periodically oscillate between different links, which will always leave some links congested and some links not congested. The consequence of this is underusing the capacity of the network. This is very similar to the concept of Herd Behavior in behavioral finance (Scharfstein and Stein, 1990; Banerjee, 1992). A sweet spot for those cases in terms of average speed and average queue length lies between 0% and 100% which depends on the demand, the signal control methods and many other latent factors yet to be identified. To overcome the negative effect of all-people re-routing, we can limit the accessibility to information so that only a portion of travellers can have the most up-to-date traffic information and be able to do re-routing. The proportion of re-routing travellers is an optimizable variable that can be closely related to traffic conditions and network geometry.

5.3. Effects of link travel time updates

As mentioned already in Section 1, the major difference between our work and the existing work is that the link travel time is constantly updated with the most current link travel time. Here, we study the performance difference between our method and the traditional one which does not have the functionality to update the link travel time. In Fig. 11, we have two different routing methods: Adaptive routing (AR) and Dynamic traffic routing (DTR). The former is a traditional stochastic routing policy mentioned in Section 1. It does not have the ability to update link travel time dynamically. The latter is the routing method we proposed. It allows dynamic updating of link travel time.

The dash lines are for AR, while the solid lines are for DTR. We compare these two under three different signal control methods: fixed timing control, low-density control and high-density control (No Phase Selection Control or Modified Max Pressure Control Method was tested in this case). Network traffic loads from 200 up to 2800 vehicles with an increment of 200 are tested. From the figure, we can see that all dash lines are above the corresponding solid lines (hence longer aver-
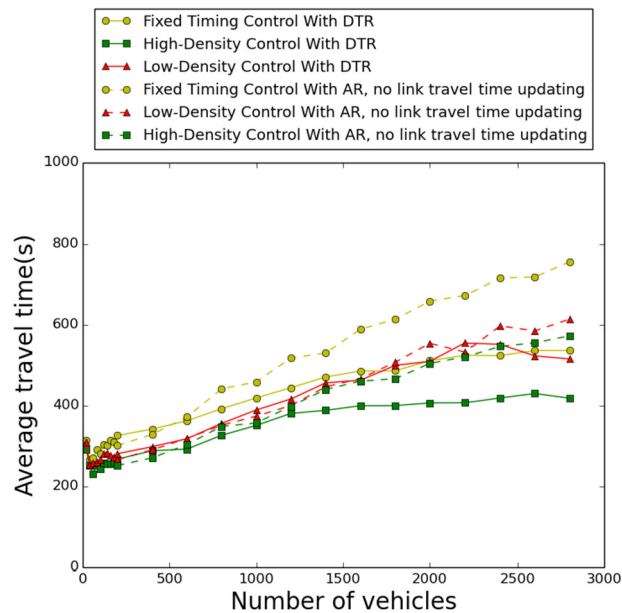
Fig. 11. Average travel time with and without link travel time updating.

age travel time) which means the DTR algorithm consistently works better than traditional AR algorithm. As traffic load becomes heavier, this benefit becomes larger.

5.4. Effects on the macroscopic fundamental diagram

Another metric to evaluate the overall performance of a network is the throughput. In this subsection, we measure the throughput (in veh/h) of the network under different signal control methods. The results are shown in Fig. 12. The x-axis is the number of running vehicles in the network, meanwhile the y-axis is the corresponding throughput (a.k.a. network exit flow) at that moment. we use different colors to indicate the temporal evolution of the simulation process: starting from the blue end and finishing in the red end. We combine simulation results from different demand levels into one figure: 500, 3000 and 6000 vehicles. In this way, we can construct an MFD as complete as possible to include both the free flow part and congested part. The 1-s data are quite noisy and moving average (with 50 steps) is applied to filter the noise. From the figure we can see that the relationship between throughput and number of running vehicles can be varying as the signal control method changes. In this experiment, 50% vehicles are re-routing vehicles and traffic signal control methods are: *Phase Selection Control, Max Pressure Control* and *Modified Max Pressure Control*. The small circles are for 500-vehicle case; the "+" signs are for the 3000-vehicle case and the small squares are for 6000-vehicle case. The colors are in a time sequence that corresponds to the simulation time. It starts from the bluish color and ends with the reddish color.

All three sub-figures have a portion of data points starting from the lower part of the figure (blue end), and are deviating from the "MFD curve". This portion belongs to the initial vehicle loading part while subjects to a very large uncertainty as the demand is generated randomly. This variation can be eliminated if we leave out the loading part (see Fig. 13). From the MFD plot, we can see that with *Phase Selection Control* and *Modified Max Pressure Control*, the MFD is closer to a trapezoidal shape than the *Max Pressure Control*. The *Max Pressure Control* case is more noisy. Under this control method, the network is gridlocked so the simulation terminates after the maximum simulation time is reached. The flow rate in the network is close to 0 while the density of the network is some positive value. This is represented as the rising part starting at a density of around 40 veh/km and a flow of about 0 veh/h (normal MFD should originate from (0,0)). The former two control methods also have a higher maximum flow which is the highest point in the MFD, which indicates that the network has a higher throughput under *Phase Selection Control* and *Modified Max Pressure Control*. Between these two control methods, there is no significant difference from the MFD point of view.

5.5. Traffic accident scenario

The previous simulations are all under normal traffic conditions, which showed that the routing algorithm and traffic signal light control algorithm we proposed work satisfactorily. We still want to see how the proposed algorithms will perform under unusual scenarios. Here, a traffic accident scenario is designed in order to test that. We manually pick three links on
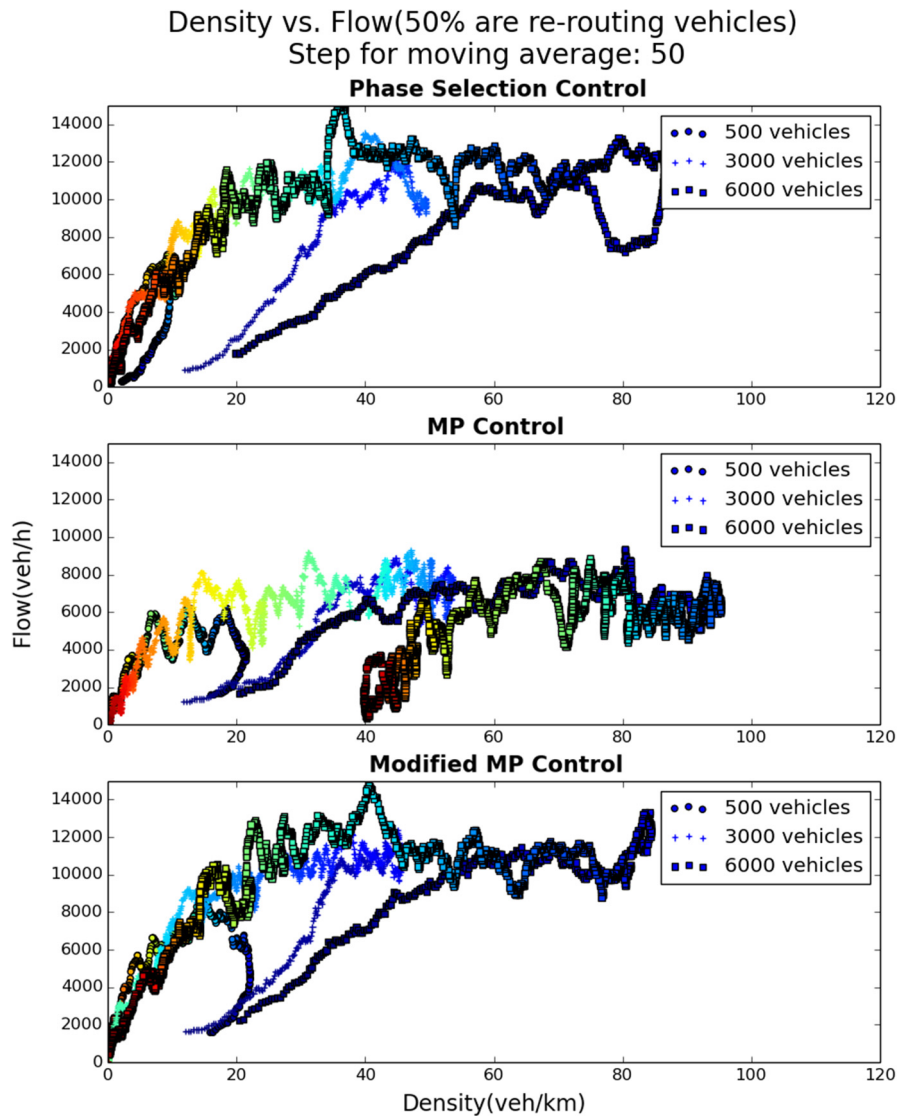
Fig. 12. Throughput vs. number of vehicles in the network.

the main arterial to be the locations where the accidents occur (see Fig. 14). This is to mimic a typical traffic accident scenario in down-town road networks. The links with accidents are assumed to be blocked during the time of accidents. Vehicles on those links are forced to stop behind the accident location. They are allowed to resume moving only after the accidents clear. The detailed information of duration and locations of accidents are given in Table 8.

We tested three different scenarios: 500 vehicles, 3000 vehicles and 6000 vehicles. The traffic light control method is *Phase Selection Control*. The results are shown in Fig. 15. From the figure, we can see that in the 500-vehicle and 3000-vehicle cases, the average travel time decreases as the percentage of re-routing vehicles increases; however, in the 6000-vehicle case the average travel time shows a slight "U" shape pattern where the minimum is reached when re-routing vehicle percentage is between 0.6 and 0.8. The observation of these simulations also confirms the explanation in the previous sections. When traffic demand is relatively low (500 and 3000 vehicles), most of the links are not congested yet. When some of the links are blocked or partially blocked by accidents, vehicles can reduce their travel time by taking some detours prior their entering into the blocked links. The network is not highly congested so these re-routing vehicles won't cause new congestions when they re-route and the network performs better when more vehicles re-route to avoid the accident locations. However, when the network is already heavily congested, if all vehicles are able to re-route, the result will not necessarily be the best since their routing decisions are not aimed at reducing network-wide travel time, and in some cases even becomes worse. The simulation suggests that in the 6000-vehicle case, a good choice will be allowing 60–80% vehicles to re-route. This percentage range, however, is likely to vary by network, traffic demand level, and driver population.
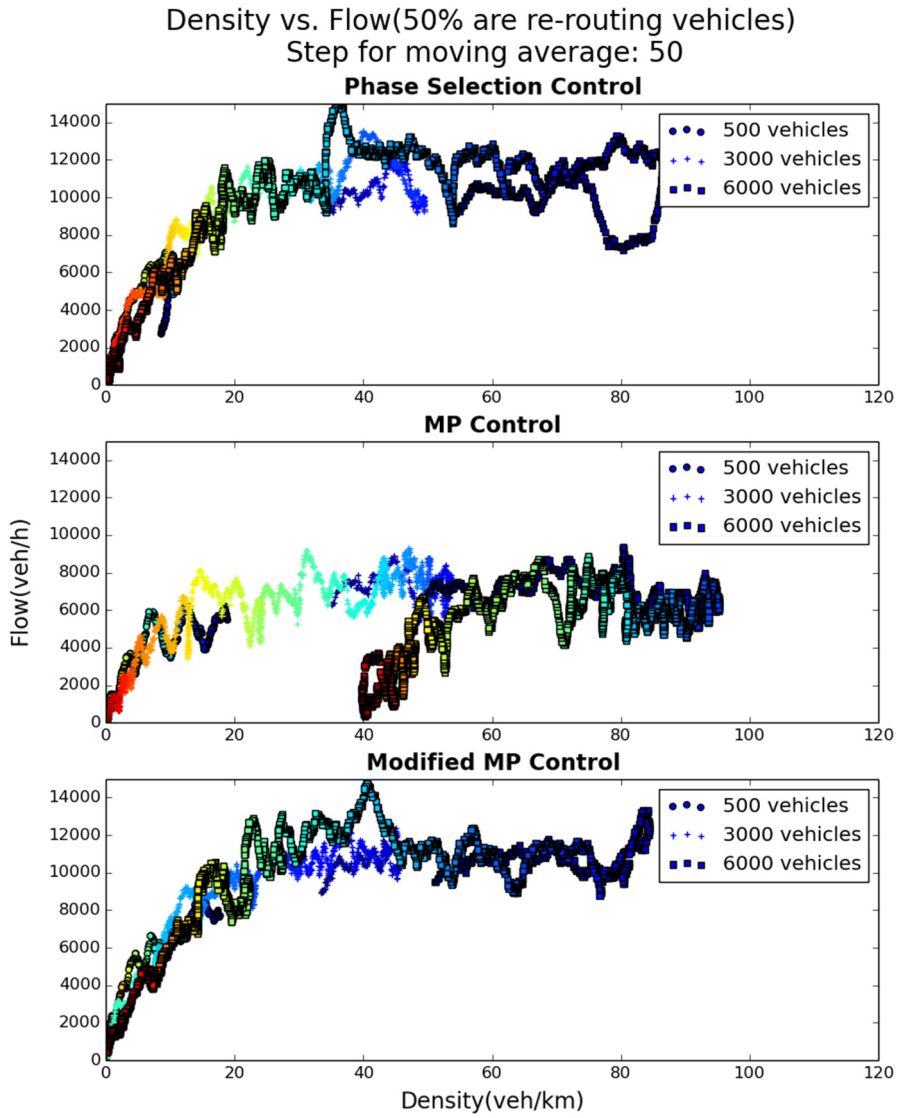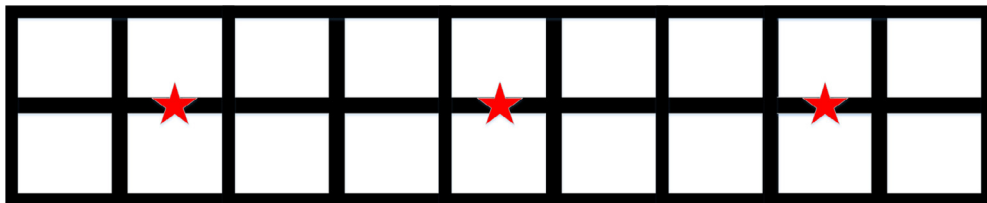
Fig. 13. Throughput vs. number of vehicles in the network (no initial loading).

★ : Accidents

Fig. 14. A traffic accident scenario: three links with accidents.

Table 8
Links with accidents.

Link	Begin time (s)	End time (s)	Position	Lane index
12–15	40	200	250	1
19–18	60	250	150	0
15–16	200	450	300	1

(a) Case 1: 500 vehicles

(b) Case 2: 3000 vehicles
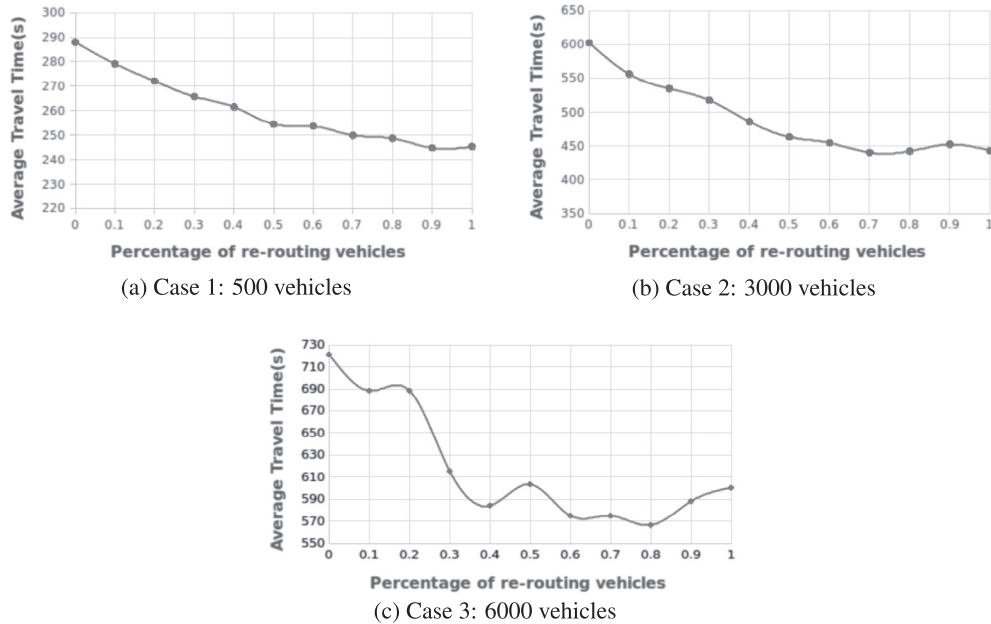
(c) Case 3: 6000 vehicles

Fig. 15. Average travel time under accident scenario.

6. Conclusions and future work

In this paper, we proposed a joint adaptive routing and traffic signal control algorithm to improve traffic operations in a VANET enabled traffic environment. Our dynamic routing algorithm (DTR) is an extension of the LET algorithm of Miller-Hooks and Mahmassani (2000) with periodic updates of link travel times. The proposed algorithm also takes into account the delay caused by real-time traffic signal operations. Besides several traditional traffic signal control strategies, namely fixed-timing, vehicle actuated control (known as low density control in our paper) and adaptive Webster's (known as high density control in our paper), we also proposed two new traffic signal control strategies, the Phase Selection Control and the Modified Max Pressure Control, to take into account the effects of both incoming demand and current queues on traffic signal operations.

The most difficult part of tackling the joint routing and signal control problem is modeling the interaction between these two components. The interaction is implicit, and thus hard to model analytically. Also, the routing problem is user-based and time-dependent. Hence, the computation cost for running simulation is not negligible as route needs to be re-calculated per traveller and per time interval which can be as small as a second depends on the resolution desired. This puts a significant burden on computation. In order to test the effects of different parameters on the performance of the proposed algorithms, a sizable number of simulations with various parameters need to be run. This is another computationally demanding task. So it needs a computationally efficient algorithm to speed up the simulations. Furthermore, some approximation techniques are needed to shorten the simulation time even more. These tasks are not trivial both from algorithmic point of view and implementation point of view.

Our simulation results shows that the DTR algorithm works well under higher demand scenarios together with the adaptive traffic signal control methods proposed in this paper. Enabling vehicle re-routing in the network can reduce the average travel time as well as reduce the average queue length at the intersection. The optimal re-routing ratio lies between 0 and 1.0, which our simulation tells us for 6000 vehicles this number is around 0.5. This number will likely vary over networks, traffic demand, and driver population. With the dynamic travel time updating model proposed in the paper, the re-routing algorithm can further reduce the average travel time in the network by taking advantage of the most current link travel time information.

The different signal control methods proposed and tested in this paper under different scenarios tells us that the Phase Selection Control and the Modified Max Pressure Control work better than the rest four control methods (including the original Max Pressure Control and fixed timing control). They tend to respond to traffic and accommodate traffic better than the rest. Average speed is higher and average queue length is shorter when these two control methods are applied. Among all the six control methods (including the original Max Pressure Control), the original Max Pressure Control performs the worst as its logic to find the optimal phase and its corresponding duration is not well designed. With the Phase Selection Control and the Modified Max Pressure Control, the MFD of the network is closer to a trapezoidal shape compared to that of the Original Max Pressure Control. For the latter, its MFD is more chaotic and fluctuating. The maximum flow rate is also lowered compared to the former two.

The joint dynamic traffic routing and adaptive signal control approach is tested against a traffic accident scenario. With links blocked or partially blocked by accidents, the dynamic traffic routing (DTA) can efficiently re-route traffic to other uncongested links to avoid long delay. According to the simulation results, the average travel time in the network can be reduced by 17—27% when the percentage of re-routing vehicles is chosen properly. The optimal percentage of re-routing vehicles in this case relies heavily on the traffic demand level. In uncongested or mildly congested scenarios, the more re-routing vehicles in the network, the less the average travel time. In a highly congested network, however, the optimal number of re-routing vehicles lies somewhere between 0% and 100%, which again depends on the traffic demand, driver population, network geometry and so on.

A possible extension of this work would be to consider global signal optimization. The signal control methods proposed in this work are all distributed ones. The phase selection method, for example, seeks local optimality but not global optimality. Coordination could be another extension to the current work. These new research directions will be even more challenging than the current joint routing and distributed signal control problem that we have dealt with in this paper and are worthy of serious investigation.

Acknowledgments

Appendix A. Algorithms

A.1. Algorithm 1: DTR algorithm with adaptive signal control

Algorithm 1. DTR algorithm with adaptive signal control

1:	**Procedure** Initialization
2:	**for** each node $i \in V$ **do**
3:	**for** each node $h \in \{\Gamma(i), i\}$ **do**
4:	**for** each $t \in S$ **do**
5:	**if** $i \neq D$ **then**
6:	$\lambda_i^h(t) = \infty$
7:	$\pi_i^h(t) = \infty$
8:	**else**
9:	$\lambda_i^h(t) = 0$
10:	$\pi_i^h(t) = 0$
11:	**end if**
12:	**end for**
13:	**end for**
14:	**end for**
15:	Create an empty vector H
16:	Insert D into H: $H \leftarrow D$
17:	**end procedure**
18:	
19:	**procedure** UPDATE LABELS
20:	**for** each $i \in \Gamma(j)$ **do**
21:	**for** each $h \in \{\Gamma(i), i\}$ **do**
22:	**for** each $t \in T$ **do**

23: $$\mu_i^h(t) = \min_{j\in\Gamma(i)}\left\{\sum_{k=1}^K\left\{\left[\phi_{ij}^{\omega_i^l}(t) + \tau_{ij}^k\left(t+\phi_{ij}^{\omega_i^l}(t)\right) + \lambda_j^i\left(t+\phi_{ij}^{\omega_i^l}(t) + \tau_{ij}^k\left(t+\phi_{ij}^{\omega_i^l}(t)\right)\right)\right]\cdot\rho_{ij}^k\left(t+\phi_{ij}^{\omega_i^l}(t)\right)\right\}\right\}$$

24:     **if** $\mu_i^h(t) < \lambda_i^h(t)$ **then**

25:         $\lambda_i^h(t) = \mu_i^h(t)$

26:         $\pi_i^h(t) = j$

27:         $H = H\bigcup i$

28:     **end if**

29:    **end for**

30:   **end for**

31: **end for**

32: Go to procedure INITIALIZATION

33: **end procedure**

34:

35: **procedure** CHECK VEHICLE STATE

36: **if** next node is reached and destination node is not reached **then**

37:   Go to Procedure Updating Queue Distribution

38: **else if** next node is reached and it is destination node **then**

39:   Go to Procedure Stop

40: **else**

41:   $t \leftarrow t + 1$

42:   Go to Procedure RE-ROUTING

43: **end if**

44: **end procedure**

45: **procedure** UPDATE LINK TRAVEL TIME

46:

$$\rho^{updated}(t=\tau) = \begin{cases} b\cdot\rho^{old}(t=\tau), & \text{if } \tau\in T^{old} \\ c\cdot\rho^{new}(t=\tau), & \text{if } \tau\in T^{new} \\ b\cdot\rho^{old}(t=\tau) + c\cdot\rho^{new}(t=\tau), & \text{if } \tau\in T^{old}\cap T^{new} \end{cases}$$
$$T^{updated} = T^{old}\cup T^{new}$$

47: **end procedure**

48:

49: **procedure** RE-ROUTING

50: Go to Procedure Initialization

51: **end procedure**

52: **procedure** STOP

53: STOP

54: **end procedure**

A.2. *Algorithm 2: Modified Max Pressure Algorithm*

Algorithm 2. Modified Max Pressure Algorithm

1: **procedure** CHOOSE PHASE AND DURATION

2:   **for** each $t \in [G_{min}, G_{max}]$ **do**

3:    **for** each movement m **do**

4:     $f_m(t) = \frac{\alpha\times N_t^m + \beta\times Q_t^m}{t}$

5:     (N_t^m: is the number of vehicles that will arrive within time t that will move in movement m)

6:     (Q_t^m: is the queue length on the link of movement m)

7:    **end for**

8:   **end for**

9:   $f_{high} = -1$

10:   **for** each $t \in [G_{min}, G_{max}]$ **do**

11:    **for** each phase p **do**

12:     sum the flow rates of each movement in the phase at that time

```
13:          if summed flow rate ⩾ f_high then
14:              f_high = summed flow rate
15:              next phase = p
16:              next phase duration = t
17:          end if
18:       end for
19:    end for
20: end procedure
```

A.3. Original MP control

The MP Control policy (Varaiya, 2013) $u^* : \chi \to \mathcal{S}$. For $X \in \chi$ assign the weight of each movement (n, m)

$$w(n, m)(X) = x(n, m) - \sum_{p \in Out_m} r(m, p) x(m, p)$$

and assign the pressure of each network signal control matrix $S \in \mathcal{S}$

$$\gamma(S)(X) = \sum_{n,m} c(n, m) w(n, m)(X) S(n, m) = \sum_{n,m; S(n,m)=1} c(n, m) w(n, m)(X)$$

The MP policy u^* is:

$$u^*(X) = argmax\{\gamma(S)(X) | S \in \mathcal{S}\}$$

References

Amoozadeh, M., Deng, H., Chuah, C.-N., Zhang, H.M., Ghosal, D., 2015. Platoon management with cooperative adaptive cruise control enabled by VANET. Vehic. Commun. 2 (2), 110–123.
Arellano, W., Mahgoub, I., 2013. Trafficmodeler extensions: a case for rapid VANET simulation using, OMNET++, sumo, and veins. In: 2013 High Capacity Optical Networks and Emerging/Enabling Technologies.
Banerjee, A.V., 1992. A simple model of herd behavior. Quart. J. Econ., 797–817
Bellman, R., 1956. On a Routing Problem, Tech. rep. DTIC Document.
Chen, P.W., Nie, Y.M., 2015. Stochastic optimal path problem with relays. Transport. Res. Part C: Emerg. Technol. 59, 48–65.
Chen, Y.-L., Yang, H.-H., 2000. Shortest paths in traffic-light networks. Transport. Res. Part B: Methodol. 34 (4), 241–253.
Cheng, S.-F., Epelman, M.A., Smith, R.L., 2006. Cosign: a parallel algorithm for coordinated traffic signal control. IEEE Trans. Intell. Transport. Syst. 7 (4), 551–564.
Chien, S.I.-J., Kuchipudi, C.M., 2003. Dynamic travel time prediction with real-time and historic data. J. Transport. Eng. 129 (6), 608–616.
Choy, M.C., Srinivasan, D., Cheu, R.L., 2003. Cooperative, hybrid agent architecture for real-time traffic signal control. IEEE Trans. Syst., Man Cybernet., Part A: Syst. Hum. 33 (5), 597–607.
Desrosiers, J., Dumas, Y., Solomon, M.M., Soumis, F., 1995. Time constrained routing and scheduling. Handbooks Operat. Res. Manage. Sci. 8, 35–139.
Desrosiers, J., Soumis, F., Desrochers, M., 1984. Routing with time windows by column generation. Networks 14 (4), 545–565.
Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1 (1), 269–271.
Fan, Y., Kalaba, R., Moore, J., 2005. Shortest paths in stochastic networks with correlated link costs. Comp. Math. Appl. 49 (9), 1549–1564.
Friesz, T.L., 1985. Transportation network equilibrium, design and aggregation: key developments and research opportunities. Transport. Res. Part A: Gen. 19 (5), 413–427.
Fu, L., 2001. An adaptive routing algorithm for in-vehicle route guidance systems with real-time information. Transport. Res. Part B: Methodol. 35 (8), 749–765.
Gao, S., Chabini, I., 2006. Optimal routing policy problems in stochastic time-dependent networks. Transport. Res. Part B: Methodol. 40 (2), 93–122.
Haddad, J., Ramezani, M., Geroliminis, N., 2013. Cooperative traffic control of a mixed network with two urban regions and a freeway. Transport. Res. Part B: Methodol. 54, 17–36.
Horowitz, J.L., 1984. The stability of stochastic equilibrium in a two-link transportation network. Transport. Res. Part B: Methodol. 18 (1), 13–28.
Hunt, P., Robertson, D., Bretherton, R., Royle, M.C., 1982. The scoot on-line traffic signal optimisation technique. Traffic Eng. Control 23 (4).
Lin, W.-H., Kulkarni, A., Mirchandani, P., 2004. Short-Term Arterial Travel Time Prediction for Advanced Traveler Information Systems. Taylor & Francis, pp. 143–154.
Lo, H.K., 1999. A novel traffic signal control formulation. Transport. Res. Part A: Policy Pract. 33 (6), 433–448.
Miller-Hooks, E., 2001. Adaptive least-expected time paths in stochastic, time-varying transportation and data networks. Networks 37 (1), 35–52.
Miller-Hooks, E.D., Mahmassani, H.S., 2000. Least expected time paths in stochastic, time-varying transportation networks. Transport. Sci. 34 (2), 198–215.
Min, H., Jayaraman, V., Srivastava, R., 1998. Combined location-routing problems: a synthesis and future research directions. Euro. J. Operat. Res. 108 (1), 1–15.
Mirchandani, P., Head, L., 2001. A real-time traffic signal control system: architecture, algorithms, and analysis. Transport. Res. Part C: Emerg. Technol. 9 (6), 415–432.
Nagy, G., Salhi, S., 2007. Location-routing: Issues, models and methods. Euro. J. Operat. Res. 177 (2), 649–672.
Nie, Y.M., Wu, X., 2009. Shortest path problem considering on-time arrival probability. Transport. Res. Part B: Methodol. 43 (6), 597–613.
Noland, R.B., Polak, J.W., 2002. Travel time variability: a review of theoretical and empirical issues. Transp. Rev. 22 (1), 39–54.
Perl, J., Daskin, M.S., 1985. A warehouse location-routing problem. Transport. Res. Part B: Methodol. 19 (5), 381–396.
Rosdolsky, H., 1973. A method for adaptive traffic control. Transport. Res. 7 (1), 1–16.
Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. Am. Econ. Rev., 465–479
Shladover, S.E., Su, D., Lu, X.-Y., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. Transport. Res. Rec.: J. Transport. Res. Board 2324 (1), 63–70.
Smith, M., Ghali, M., 1990. The dynamics of traffic assignment and traffic control: a theoretical study. Transport. Res. Part B: Methodol. 24 (6), 409–422.

Solomon, M.M., Desrosiers, J., 1988. Survey paper-time window constrained routing and scheduling problems. Transport. Sci. 22 (1), 1–13.

Tatomir, B., Rothkrantz, L., 2004. Dynamic traffic routing using ant based control. 2004 IEEE International Conference on Systems, Man and Cybernetics, vol. 4. IEEE, pp. 3970–3975.

Van Arem, B., Van Driel, C.J., Visser, R., 2006. The impact of cooperative adaptive cruise control on traffic-flow characteristics. IEEE Trans. Intell. Transport. Syst. 7 (4), 429–436.

Varaiya, P., 2013. Max pressure control of a network of signalized intersections. Transport. Res. Part C: Emerg. Technol. 36, 177–195.

Wu, X., 2015. Study on mean-standard deviation shortest path problem in stochastic and time-dependent networks: a stochastic dominance based approach. Transport. Res. Part B: Methodol. 80, 275–290.

Xiao, L., Lo, H.K., 2014. Combined route choice and adaptive traffic control in a day-to-day dynamical system. Netw. Spat. Econ., 1–21

Yang, B., Miller-Hooks, E., 2004. Adaptive routing considering delays due to signal operations. Transport. Res. Part B: Methodol. 38 (5), 385–413.

Yang, H., Yagar, S., 1995. Traffic assignment and signal control in saturated road networks. Transport. Res. Part A: Policy Practice 29 (2), 125–139.

Zaidi, A.A., Kulcsar, B., Wymeersch, H., 2015. Traffic-adaptive signal control and vehicle routing using a decentralized back-pressure method. In: 2015 European Control Conference (ECC). IEEE, pp. 3029–3034.

Zhang, D., Nagurney, A., 1996. On the local and global stability of a travel route choice adjustment process. Transport. Res. Part B: Methodol. 30 (4), 245–262.

Zhang, X., Rice, J.A., 2003. Short-term travel time prediction. Transport. Res. Part C: Emerg. Technol. 11 (3), 187–210.

Zheng, F., Van Zuylen, H., 2010. Uncertainty and predictability of urban link travel time. Transport. Res. Rec.: J. Transport. Res. Board 2192 (1), 136–146.