Article:

Ngoduy, D orcid.org/0000-0002-0446-5636, Hoang, NH, Vu, HL et al. (1 more author) (2016) Optimal queue placement in dynamic system optimum solutions for single origin-destination traffic networks. Transportation Research Part B: Methodological, 92 (Part B). pp. 148-169. ISSN 0191-2615

https://doi.org/10.1016/j.trb.2015.11.011

Optimal queue placement in dynamic system optimum solutions for single origin-destination traffic networks

D. Ngoduy[a,1], N.H. Hoang[b], H.L. Vu[b], D. Watling[a]

[a]Institute for Transport Studies, The University of Leeds, Leeds, United Kingdom
[b]Intelligent Transport Systems Lab, The Swinburne University of Technology, Melbourne, Australia

---

Abstract

The Dynamic System Optimum (DSO) traffic assignment problem aims to determine a time-dependent routing pattern of travellers in a network such that the given time-dependent origin-destination demands are satisfied and the total travel time is at a minimum, assuming some model for dynamic network loading. The network kinematic wave model is now widely accepted as such a model, given its realism in reproducing phenomena such as transient queues and spillback to upstream links. An attractive solution strategy for DSO based on such a model is to reformulate as a set of side constraints apply a standard solver, and to this end two methods have been previously proposed, one based on the discretisation scheme known as the Cell Transmission Model (CTM), and the other based on the Link Transmission Model (LTM) derived from variational theory. In the present paper we aim to combine the advantages of CTM (in tracking time-dependent congestion formation within a link) with those of LTM (avoiding cell discretisation, providing a more computationally attractive with much fewer constraints). The motivation for our work is the previously-reported possibility for DSO to have multiple solutions, which differ in where queues are formed and dissipated in the network. Our aim is to find DSO solutions that optimally distribute the congestion over links inside the network which essentially eliminate avoidable queue spillbacks. In order to do so, we require more information than the LTM can offer, but wish to avoid the computational burden of CTM for DSO. We thus adopt an extension of the LTM called the Two-regime Transmission Model (TTM), which is consistent with LTM at link entries and exits but which is additionally able to accurately track the spatial and temporal formation of the congestion boundary within a link (which we later show to be a critical element, relative to LTM). We set out the theoretical background necessary for the formulation of the network-level TTM as a set of linear side constraints. Numerical experiments are used to illustrate the application of the method to determine DSO solutions avoiding spillbacks, reduce/eliminate the congestion and to show the distinctive elements of adopting TTM over LTM. Furthermore, in comparison to a fine-level CTM-based DSO method, our formulation is seen to significantly reduce the number of linear constraints while maintaining a reasonable accuracy.

Keywords: kinematic wave model, cell transmission model, link transmission model, two regime transmission model, double queue model, dynamic system optimum, spillbacks

---

*Corresponding author. email:d.ngoduy@leeds.ac.uk

Accepted for publication in Transportation Research Part B (Nov. 2015) *November 16, 2015*

1. Introduction

The Dynamic System Optimum (DSO) traffic assignment problem predicts the optimal time-dependent routing pattern of travellers in a network such that the given time-dependent origin-destination demands are satisfied and the total system travel time spent by all travellers is minimized, assuming some model for dynamic network loading. It has been shown recently by Shen and Zhang (2014) that DSO may have multiple solutions which share the same DSO objective (i.e. the total travel times) but have different queue lengths at nodes in the network. There are two types of special DSO solutions: the non-holding back solutions and the free-flow solutions. Non-holding back solutions require the system to discharge flow as much as it can so that there might be heavy congestion at the intersections inside the network, which in many cases might cause spillbacks to the upstream links. Here the spillback appears when downstream congestion in a link propagates upstream and reaches the upstream node, and hence possibly restricts the inflow to the link, and subsequently to restrict the exit flow from adjacent upstream links (Qian et al., 2012). The free-flow DSO solutions does not allow vehicles to wait in the queue inside the network, but allows the queue to build up at the origin nodes. Given the contrast in where the queues form in the network, the interesting question to ask is how to optimally distribute queues inside the network in order to achieve a better DSO solution at the end. This paper focuses on the DSO solutions which optimally distribute the (horizontal) congestion over links in the network. We expect that such optimal distribution of the congestion over links will lead to better traffic operation inside the network such as minimal queue spillbacks or even free-flow condition, which consequently reduce the stop-start traffic dynamics in the network and, hence, reduce emissions. To this end, there is a need to explicitly determine the time and space evolution of the queues in the network.

There has been much effort in literature undertaken to formulate the kinematic wave model (KWM) of Lighthill and Whitham (1955) and Richards (1956), based on an analogy between traffic flow and certain types of wave motion in fluids, for the DSO problem. Especially, a discrete version of the KWM, called the cell transmission model (CTM) (Daganzo, 1994, 1995) has been formulated as side constraints in the DSO problem (Carey and Ge, 2012, Carey and Watling, 2012, Gentile et al., 2005, 2007, Lo and Szeto, 2002, Nie and Zhang, 2005, Shen and Zhang, 2014, Szeto et al., 2011, Szeto and Lo, 2006, Ukkusuri et al., 2012, Ukkusuri and Waller, 2008, Ziliaskopoulos, 2000). It is because knowing the time-space dynamics of traffic flow within the link will facilitate a better understanding of the resulting SO solution, for example the level of congestion and spillback location in the network, etc. An important property of the CTM is the possibility to reformulate it as a relaxed set of linear constraints so that a linear programming model for the DSO-DTA problem for a network can be solved efficiently (Beard and Ziliaskopoulos, 2006, Li et al., 2003, Ukkusuri and Waller, 2008, Ziliaskopoulos, 2000). However, the choice to adopt CTM is not without its computational overheads. As noted in Nie and Zhang (2005), if we just consider the issue of DNL (of given route in-flow profiles) then the computational time for the CTM is directly proportional to the number of cells, and hence the computational efficiency is inversely proportional to the accuracy (in recovering the LWR). Thus the choice of discretisation level effectively means a choice/compromise between computational efficiency and the level of agreement with the (continuum) KWM model. If we then consider the wider issue of how the model is integrated within a DSO framework, then we are faced with further computational issues: if the CTM is specified as a set of side constraints as suggested by Peeta and Ziliaskopoulos (2001) then the number of constraints grows with the fineness of the discretisation, whereas if we represent it using a route-based mapping as in Lo and Szeto (2002), then we are led down a path of route enumeration with all the computational difficulties that it is known to bring. As noted in Bar-Gera (2005), the 'computational requirements reduce the attractiveness of this model for large-scale long duration applications'. Such computational issues of CTM based DSO problems are further explored in Shen and Zhang (2008) for different network sizes, where the number of constraints is increased polynomially with the network size.

Recently, a well known Lax-Hopf (LH) formula for the Hamilton-Jacobi (HJ) type partial differential equations has been used to solve the KWM, which can avoid the discretisation, hence enhance the computational efficiency. The LH formula can be used to provide a variational formulation of the HJ equation solutions describing the evolution of cumulative number of vehicles at the two ends of the link. In principle, the LH formula has been derived in various ways including the traffic flow theory by Daganzo (2005, 2006) which actually generalizes the theory of Newell (1993), the viability theory by Aubin et al. (2008, 2011) for given boundary conditions at two ends of the link, which is then extended to include the internal condition (i.e. information of probe vehicles) by Claudel and Bayen (2010a,b), and the technique of calculus of variations by Evans (2010). The Link Tranmission Model (LTM) in either discrete form (Yperman et al., 2005) or continuous form (Han et al., 2015, Jin, 2015) has been developed using the Newells theory which is a special formulation of LH formula above, where the state of the whole link (i.e. either free-flow or congested) will be determined by the entry and exit flow. More specifically, the flow propagation in LTM is based on Newell's cumulative flow curves applied at the entry/exit of each link, with node models used to calculate the transition flows, which are based on conservation of flow between the incoming and outgoing flows. Sending/receiving flows, together with transition flows and other flow constraints, form the basis for updating the cumulative flows at the link boundaries. Osorio and Flotterod (2014), Osorio et al. (2011) have then developed another version of LTM, which is a so-called Double Queue Model (DQM). In the DQM, the link is treated as a set of two queues, referred to as the upstream queue and the downstream queue. Both LTM and DQM can properly capture the free-flow travel time delay when the link is in a free-flow state and the backward shockwave time delay when the link is in a congested state, which make them possible to capture queue spillbacks. The DQM was used in Ma et al. (2014) to find a free-flow DSO solution where spillback is tracked by the traffic state at the link entrance (in free flowing) or at the link exit (in congested) accounting for some time shift. Nevertheless, either LTM or DQM does not determine explicitly the propagation of the front shocks within a link and thus is unsuitable for providing the detailed traffic state within the link.

Given the background described above, we aim to formulate a DSO problem with a KWM as linear constraints. To fulfill our objective (i.e. optimal distribution of the queues over links), we must track the time and space evolution of the queue lengths in the DSO problem. While LTM or DQM cannot describe such evolution of the queue lengths, CTM can do it with high computational cost, especially for the DSO problem. We will address this problem by applying a recently proposed Two-regime Transmission Model (TTM) in Balijepalli et al. (2014) for our DSO problem. The TTM principle is more desirable for our DSO-related problem than using the CTM, LTM or DQM as, on the one hand, it utilizes the entry and exit flows to describe the link state similarly to the LTM or DQM, and, on the other hand, it provides the time and space evolution of the queue lengths. Basically, the TTM assumes that traffic along a link is characterised by two regimes (i.e. either non-congested where the density is below a critical density or congested where the density is above the critical density). In TTM, the (horizontal) length of the congested regime can be determined by the inflow at the upstream node and outflow at the downstream node. With these assumptions, we are only interested in the congestion formation and the propagation of the tail of the congestion under a certain boundary condition (when and where a traffic congestion occurs). We have shown in Balijepalli et al. (2014) that the TTM is able to capture accurately the traffic states (e.g. the evolution of density) both in (horizontal) space and time as a fine-level CTM at a significantly reduced computational cost. The major of this paper, therefore, is to establish a new optimization framework to find DSO solutions which can optimally distribute the queue lengths over links so that the spillbacks can be minimized or free-flow traffic inside the network can be attained.

The organization of this paper follows. Section 3 briefly reviews the development of the TTM for dynamic network loading problems which has been published previously in Balijepalli et al. (2014). Section 4 formulates the TTM as a set of linear constraints for a DSO traffic assignment problem in which the queue lengths are calculated explicitly. Section 5 presents a new optimization framework which finds DSO solutions optimally distributing the queue lengths over links in the network using the linear constraints formulated in Section 4. Section 6 illustrates our numerical studies for a small and a reasonably large network in order to support the advantages of our approach. Finally, we conclude our paper in Section 7.

2. Notation

A traffic network is considered a directed graph, which consists of a set of links connected via a set of nodes. The notation below will define a traffic network being considered in this paper.

- $\mathbb{T}$: a set of discrete time steps, $\mathbb{T} \subset \mathbb{N}$. For continuous time domain, we use notation $\mathbb{T}_{\mathbb{R}}$ ($\mathbb{T}_{\mathbb{R}} \subset \mathbb{R}$). We define $T = |\mathbb{T}|$ as the number of time steps.
- $\mathbb{V}$: Set of nodes. There are two subsets: set of source nodes $\mathbb{V}_R$ and set of sink nodes $\mathbb{V}_S$, such that: $(\mathbb{V}_R \cup \mathbb{V}_S) \subseteq \mathbb{V}$ and, $\mathbb{V}_R \cap \mathbb{V}_S = \varnothing$.
- $\mathbb{E}$: Set of directed links, combined by any two nodes in $\mathbb{V}$, e.g. if $(n, m) \in \mathbb{E}$ then we call n the upstream node, and m the downstream node of this link. There are three subsets of links: normal, source ($\mathbb{E}_R$) and sink ($\mathbb{E}_S$) links. The last two types are the specialized and virtual links to provide features of source nodes and sink nodes. For each link $a \in \mathbb{E}$:
  - l_a: physical length.

- Υ_a^-: set of inflow links to link a.
- Υ_a^+: set of outflow links from link a.
- $S_a(t)$: supply of link a.
- $D_a(t)$: demand from link a.
- $u_a(t)$: incoming traffic flow to link a at time t.
- $v_a(t)$: outgoing traffic flow from link a at time t.
  * Note that: $u_a(t)$, $v_a(t)$, $S_a(t)$ and $D_a(t)$ are continuous in time domain $\mathbb{T}_\mathbb{R}$.

Also for each node $n \in \mathbb{V}$:

- Γ_n^-: set of inflow links to node n.
- Γ_n^+: set of outflow links from node n.

$$
\begin{aligned}
(\Gamma_n^- \cup \Gamma_n^+) &\subseteq \mathbb{E} \\
\Gamma_n^- \cap \Gamma_n^+ &= \varnothing \\
\Upsilon_a^- &= \{e | a \in \Gamma_n^+; e \in \Gamma_n^-\} \\
\Upsilon_a^+ &= \{e | a \in \Gamma_n^-; e \in \Gamma_n^+\}
\end{aligned}
$$

- $f_{ab}(t)$: upstream traffic at link b, coming from downstream traffic at link a.

$$a, b \in \mathbb{E}$$

$$a \cap b \neq \varnothing$$

- (V_a, W_a, K_a): the set of fundamental diagram parameters, e.g. free-flow speed, backward speed and maximum density, for each link $a \in \mathbb{E}$. In this paper, the triangular flow-density relationship is adopted. Accordingly, the maximum flow is $Q_a = \frac{K_a V_a W_a}{V_a + W_a}$, and the critical density is $C_a = \frac{K_a W_a}{V_a + W_a}$.
- $l_a^c(t)$: length of congested regime in link a.
- $l_a^f(t)$: length of non-congested regime in link a.
- Common indexes: a for links in $\mathbb{E}$, n for nodes in $\mathbb{V}$, i for time steps in $\mathbb{T}$, t for continuous time in $\mathbb{T}_\mathbb{R}$.

Network formation. Two rules for source and sink nodes:

- There is only one link from a source node, called source link.
- There is only one link to a sink node, called sink link.

3. Two regime Transmission Model

This section briefly discusses the concept of the TTM for traffic dynamics in a link, which has been previously developed in Balijepalli et al. (2014), and explains the potential advantages of TTM in solving DSO problems, which is our major contribution (Section 5). Define the fundamental relationship according to the LWR model:

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{1}$$

where $\rho = \rho(x,t)$ and $q = q(x,t)$ denote, respectively, the density and flow at continuous location x and continuous time t. Here the link index a is dropped for the sake of simplicity since equation (1) holds for any link. Note, by definition, $q(0,t) \doteq u(t)$ and $q(l,t) \doteq v(t)$. Furthermore, the flow is assumed to be a function of the density, which refers to a so-called fundamental diagram. Thus we may write: $q(x,t) = \phi\left(\rho(x,t)\right)$ for some function $\phi(.)$ that is independent of x and t. We shall specifically focus on the case of a triangular flow-density relationship:

$$\phi(\rho) = \begin{cases} V\rho, & \text{if } 0 \le \rho < C; \\ W\left(K - \rho\right), & \text{if } C \le \rho \le K. \end{cases} \tag{2}$$

where C is the critical density, defined by $C = \frac{KW}{V+W}$. Together, equations (1)-(2) define the LWR model for a link, for given entry and exit flow profiles. To apply the method of characteristics as described by Newell (1993): from any given point (x,t) with density $\rho(x,t)$, we are able to trace along a curve in space-time of traffic density, back in time to a boundary location (in our case, the entry-point to or exit-point from the link). For homogeneous links (where the capacity is unchanged spatially), this space-time curve is linear, such that for our given point (x,t), we determine the density along the space-time points given by tracing back to any earlier time:

$$\rho(x,t) = \begin{cases} \rho(0, t - \frac{x}{V}), & \text{if } 0 \le \rho < C; \\ \rho(l, t - \frac{l-x}{W}), & \text{if } C \le \rho \le K. \end{cases} \tag{3}$$

and we are then able to relate the flows corresponding to the densities in (3) to the boundary conditions:

$$q(x,t) = \phi(\rho) = \begin{cases} \phi\left(\rho(0, t - \frac{x}{V})\right) = u\left(t - \frac{x}{V}\right), & \text{if } 0 \le \rho < C; \\ \phi\left(\rho(l, t - \frac{l-x}{W})\right) = v\left(t - \frac{l-x}{W}\right), & \text{if } C \le \rho \le K. \end{cases} \tag{4}$$

The novel feature of TTM we shall now describe is that it utilises (3) and (4) in a simplified way that neither requires us to treat the state of the link as homogeneous along its length (as is implicitly done in LTM or DQM, for example), nor requires a fixed, fine discretisation scheme to be applied (as in the CTM, for example). It is also worth noticing that the proposed method can capture the traffic states in time and space along the link as CTM while the LTM/DQM cannot. Of course, there are other more advanced models using the LH formula such as in Aubin et al. (2008, 2011) or in Claudel and Bayen (2010a,b) which only use the entry and exit flow to predict the traffic states within a link, but they are not straightforward and efficient to apply in solving DSO problems.

Without loss of generality, we suppose that the initial state of all links (at time $t = 0$) is empty since any initial state can belong to one of the cases below. Returning, then to the definition of our TTM approach, we now go on to consider the different cases that may arise given time instant t:

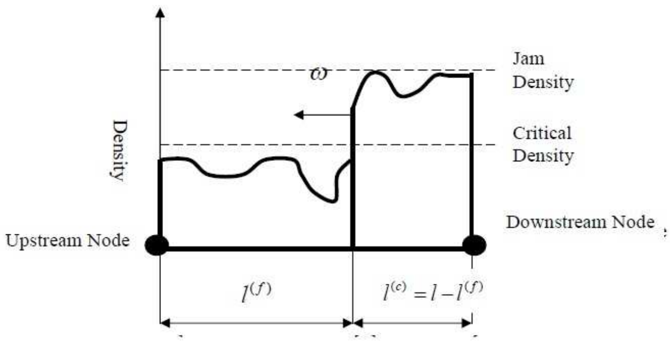


Figure 1: Congestion propagation

i. A downstream section of the link is empty, while there is traffic at the remaining upstream section operating at the free-flow condition (in principle, traffic is at the non-congested conditions in the whole link). This will occur until the first vehicle reaches the downstream end of the link, travelling at free flow speed V for the link length l.

ii. The whole link is free-flowing (front shock reaches the downstream node). This will occur, for example if $v(t) = u(t - l/V)$. This situation is similar to case i and traffic in the whole link still belongs to the non-congested regime.

iii. The whole link is congested (front shock reaches the upstream node). This will occur, for example if $u(t) = v(t - l/W)$. In this situation, traffic in the whole link belongs to the congested regime, and congestion propagates to the entrance of the link which can cause the spillback phenomena. This important phenomena will be illustrated later in Section 6.

iv. A downstream section of the link is congested, while the remaining upstream section is free-flowing (See Figure 1). This will occur, for example if $u(t) > v(t + l/V)$. Hence, there is no interaction between the two boundary conditions nor between upstream and downstream node.

It is clear that if the front shock reaches either upstream or downstream node, there is an interaction between the two nodes or the two boundary conditions, which is a main reason for the queue spillback and gridlock. To this end, both TTM and LTM (or DQM) can handle such queue spillback and gridlock situations (or link shock-wave propagation as shown in Ma et al. (2014)) while TTM can handle how the front shock propagates in time and space as in case (iv). That means with TTM we can compute explicitly the time and space evolution of the length of the congested regime, which can be used for our optimization problems: to evenly distribute the congestion over links so that the front shocks within a link can be refrained as much as possible and the spillbacks can be minimized.

Based on the discussion above, we begin to introduce our pseudo code for the link-model by first describing the time evolution of the 'regime lengths' $l^f(t)$ and $l^c(t)$ defined above, with a dot used to denote time-derivative:

i. **Initialisation** ($t = 0$). Set $l^f(0) = l^c(0)$.

ii. **Warm-up** ($0 < t \leq l/V$). By the argument given earlier, the time evolution of the regime lengths during this time period obeys: $\dot{l}^c(t) = 0$ and $\dot{l}^f(t) = V$. As the link is either free-flowing or empty at all locations, we apply equations (3)-(4) to obtain the flow and density in time and space: In sum, in this case traffic states along the link are determined by the inflow from the upstream node.

iii. **Remaining period** ($t > l/V$). By the start of this period ($t = l/V$) flow will have reached all locations along the link if in-flows are positive ($u(t) > 0$), having all travelled at free-flow speed, and so it should be the case that the length of the non-congested regime at this time is the whole link length, i.e. $l^f(l/V) = l$. During this period of continued positive in-flow we will always have $l^f(t) + l^c(t) = l$. Note that, if $\rho(x,t) \leq C$ then $l^c(t) = 0$. As congestion can only potentially form at the downstream end of the link, and our focus will therefore be on understanding the time-evolution of the length $l^c(t)$ that this congested regime propagates from the exit of the link.

Let us suppose then we are at a time t where congestion has indeed formed at the downstream end of the link (Figure 1). In such a case, the shock will propagate in the upstream direction, or dissipate in the downstream direction, at the prevailing shock-wave speed which we denote $\omega(t)$. The shock-wave speed is influenced by the supply constraints (e.g. control settings or reduced capacity) at the downstream node (in our present case, as reflected in the boundary condition of the given exit flow profile $v(t)$ $(0 < t < +\infty)$) and by the demand entering the link through the upstream node (as reflected, in our present case, in the given entry flow profile $u(t)$ $(0 < t < +\infty)$). In practice, this speed may be computed by applying the Rankie-Hugoniot (RH) condition at the current boundary location $x = l^f(t)$ between the non-congested and congested regimes as below:

$$\omega(t) = \frac{\text{[jump in the flow]}}{\text{[jump in the density]}} = \frac{q^+(t) - q^-(t)}{\rho^+(t) - \rho^-(t)} \tag{5}$$

where $q^+(t)$ and $q^-(t)$ denote, respectively, the flow at the right and left interface between the non-congested and congested regime while $\rho^+(t)$ and $\rho^-(t)$ are the density at the right and left interface between the non-congested and congested regime. By definition, $q^+(t)$ will be determined by the flow at the exit while $q^-(t)$ will be determined by the flow at the entry of the link. To follow equations (3) and (4) we obtain:

$$\omega(t) = \frac{v\left(t - \frac{l^c(t)}{W}\right) - u\left(t - \frac{l^f(t)}{V}\right)}{K - \frac{1}{W}v\left(t - \frac{l^c(t)}{W}\right) - \frac{1}{V}u\left(t - \frac{l^f(t)}{V}\right)} \tag{6}$$

The dynamic equation for the length of the congested regime is computed as:

$$\dot{l}^c(t) = -\omega(t) \tag{7}$$

From equation (7), if $\omega(t) < 0$, that is $q^+(t) < q^-(t)$, the shock is moving upstream, the length of the congested regime is increased, whereas if $\omega(t) > 0$, that is $q^+(t) > q^-(t)$, the shock is dissipating downstream, the length of the congested regime is decreased. Nevertheless, if $\omega(t) = 0$ the length of the congested regime is unchanged (stationary). Note that $\omega(t)$ describes the slope of the line connecting a point in the non-congested branch with a point in the congested branch of the density-flow relationship; it should not be confused with W which is the slope of the congested branch of the flow-density relationship.

We have thus completed the consideration of the different cases that arise during different time periods. In each case we have been able to describe the dynamic evolution of the different regimes across the link length. Knowing these lengths, we are able to know when and where to apply the different cases described via equations (3) and (4), and thereby construct flows and densities across the whole link.

4. Linear constraints formulation of the TTM

This section will show how it is possible to formulate the continuous TTM for a single link equivalently as a system of (discrete) linear constraints and how to determine the time and space evolution of the length of the congested regime which will be used in finding the DSO solutions in Section 5.

4.1. Definition of the discrete time variables

Let us define the following discrete time variables from the continuous TTM described in Section 3:

$$\begin{aligned} u(i) &= \textstyle\int_i^{i+1^-} u(t)dt \\ v(i) &= \textstyle\int_i^{i+1^-} v(t)dt \\ f(i) &= \textstyle\int_i^{i+1^-} f(t)dt \qquad \forall i \in \mathbb{T} \\ D(i) &= \textstyle\int_i^{i+1^-} D(t)dt \\ S(i) &= \textstyle\int_i^{i+1^-} S(t)dt \end{aligned} \tag{8}$$

For the sake of simplicity without loss of generality let us assume that

$$\begin{aligned} \frac{l}{V} &\in \mathbb{N} \\ \frac{l}{W} &\in \mathbb{N} \end{aligned} \qquad \forall a \in \mathbb{E} \tag{9}$$

In general, we could ignore this assumption but the model becomes more complex and may bring numerical issues when working with the non-integer values. As suggested in Ban et al. (2012), the mesh size should be properly chosen to avoid such non-integer problems.

4.2. Definition of the system variables

Basically, we will consider the number of vehicles in a link a major variable for the DNL problem. Nevertheless, the computation of such variable will take into account the dynamics of the length of the congested regimes described by the TTM. First, let us define the the number of vehicles in a link at time instant $t \in \mathbb{T}_\mathbb{R}$ as the number of vehicles currently staying in this link at time t, mathematically represented by:

$$n(t) = \int_0^t (u(h) - v(h))\, \mathrm{d}h \tag{10}$$

In discrete time step, the number of vehicles in a link at time step $i \in \mathbb{T}$, means:

$$n(i) = \lim_{t \to (i+1)^-} \int_0^t (u(h) - v(h))\, \mathrm{d}h = \sum_{k=0}^{i} [u(k) - v(k)] \tag{11}$$

The number of vehicles in a link can also be defined as:

$$n(t) \doteq \int_{0}^{l} \rho(x,t)dx \tag{12}$$

From the definition of the TTM in Section 3, the number of vehicles in a link consists of the total number of vehicles staying in the non-congested regime and in the congested regime at time instant t. Substituting equation (3) into equation (12) gives:

$$\begin{aligned} n(t) &= \underbrace{\frac{1}{V}\int_{0}^{l^f(t)} u\left(t-\frac{x}{V}\right)dx}_{\text{Non-congested regime}} + \underbrace{\int_{l^f(t)}^{l} \left(K - \frac{1}{W}v(t-\frac{l-x}{W})\right)dx}_{\text{Congested regime}} \\ &= \underbrace{\frac{1}{V}\int_{0}^{l^f(t)} u\left(t-\frac{x}{V}\right)dx}_{\text{Non-congested regime}} + \underbrace{\int_{0}^{l^c(t)} \left(K - \frac{1}{W}v\left(t-\frac{x}{W}\right)\right)dx}_{\text{Congested regime}} \end{aligned} \tag{13}$$

This equation indicates that the number of vehicles in a link determined in each traffic regime by equation (3) has been implicitly incorporated and will be used latter to formulate the linear discrete link constraints.

4.3. Linear constraints

For the sake of simplicty, we drop out the link index in this section but the derivation holds for every link. From equation (13), we can compute two specific numbers of vehicles, $n^f(i)$ for $l^f(i) = l$ and $n^c(i)$ for $l^c(i) = l$, to define the states of link (i.e. either fully non-congested or fully congested):

$$n^c(i) = K.l - \sum_{k=i+1-\frac{l}{W}}^{i} v(k) \tag{14}$$

$$n^f(i) = \sum_{k=i-\frac{l}{V}+1}^{i} u(k) \tag{15}$$

These specific numbers of vehicles are the bounds of the number of vehicles in a link as shown below.

Lemma 1 (Upper bound)**.**

$$n^c(i) \geq n(i) = \sum_{k=0}^{i} [u(k) - v(k)]\, ; \forall i \in \mathbb{T} \tag{16}$$

Proof. Since the inflow and outflow are constrained by the link capacity:

$$\begin{aligned} u(t) &\leq \frac{KVW}{V+W} \\ v(t) &\leq \frac{KVW}{V+W} \end{aligned}$$

we infer that:

$$\frac{1}{V}\int_{0}^{l^f(t)} u\left(t-\frac{x}{V}\right)dx + \frac{1}{W}\int_{0}^{l^f(t)} v\left(t-\frac{l-x}{W}\right)dx \leq Kl^f(t)$$

which is then substituted into equation (13) to obtain:

$$\begin{aligned} n(t) &\leq K(l^f(t)+l^c(t)) - \frac{1}{W}\int_0^{l^f(t)} v\left(t-\frac{l-x}{W}\right)dx - \frac{1}{W}\int_0^{l^c(t)} v\left(t-\frac{x}{W}\right)dx \\ &= Kl - \frac{1}{W}\int_{l^c(t)}^{l} v\left(t-\frac{x}{W}\right)dx - \frac{1}{W}\int_0^{l^c(t)} v\left(t-\frac{x}{W}\right)dx \\ &= Kl - \frac{1}{W}\int_0^{l} v\left(t-\frac{x}{W}\right)dx \end{aligned}$$

Let $z = t - x/W$ we have: $\int_0^l v(t-\frac{x}{W})dx = W\int_{t-l/W}^{t} v(z)dz$. The above inequality becomes:

$$n(t) \leq Kl - \int_{t-\frac{l}{W}}^{t} v(z)dz$$

which is then converted into discrete time step as:

$$n(i) = \lim_{t\to(i+1)^-} n(t) \leq Kl - \sum_{k=i+1-\frac{l}{W}}^{i} v(k) = n^c(i)$$

□

Lemma 2 (Lower bound)**.**

$$n^f(i) \quad \leq \quad n(i) = \sum_{k=0}^{i}\left[u(k)-v(k)\right]; \forall i \in \mathbb{T} \tag{17}$$

Proof. From the constraints of the inflow and outflow, we infer that:

$$\frac{1}{V}\int_0^{l^c(t)} u(t-\frac{l-x}{V})dx + \frac{1}{W}\int_0^{l^c(t)} v(t-\frac{x}{W})dx \quad \leq \quad Kl^c(t)$$

which is then substituted into equation (13) to obtain:

$$n(t) \geq \frac{1}{V}\int_0^{l^f(t)} u(t-\frac{x}{V})dx + \frac{1}{V}\int_0^{l^c(t)} u(t-\frac{l-x}{V})dx$$

which can be converted to:

$$n(t) \quad \geq \quad \frac{1}{V}\int_0^{l} u(t-\frac{x}{V})dx = \int_{t-\frac{l}{V}}^{t} u(z)dz$$

or in the discrete time step:

$$n(i) = \lim_{t\to(i+1)^-} n(t) \geq \lim_{t\to(i+1)^-} \int_{t-\frac{l}{V}}^{t} u(z)dz = \sum_{k=i+1-\frac{l}{V}}^{i} u(k) = n^f(i)$$

□

In sum, the number of vehicles in a link has two linear constraints:

$$\sum_{k=i+1-\frac{l}{V}}^{i} u(k) \leq n(i) \leq Kl - \sum_{k=i+1-\frac{l}{W}}^{i} v(k) \tag{18}$$

4.4. Computation of the length of the congested regime

It can be seen from the TTM described in Section 3 that if the length of the congested regime is given, we can compute the phase-space density $\rho(x,t)$ and consequently the number of vehicles per link $n(t)$ via equation (13). In this section, we will solve an inverse problem: if the (discrete) number of vehicles per link $n(i)$ is given, we need to compute the regime length $l^c(i)$ or $l^f(i)$. To this end, we will solve a *fixed point* problem below via a system of linear equations.

Theorem 1. *Given l and a feasible value of $n(i)$, there exists a solution of $l^c(i)$ and $l^f(i)$ satisfying:*

$$\begin{aligned} l &= l^c(i) + l^f(i) \\ n(i) &= \lim_{t\to(i+1)^-} \frac{1}{V} \int_0^{l^f(t)} u\left(t - \frac{x}{V}\right) dx + \int_0^{l^c(t)} \left(K - \frac{1}{W} v\left(t - \frac{x}{W}\right)\right) dx \end{aligned} \tag{19}$$

Proof. We define $F(x)$ as:

$$F(x) = \frac{1}{V} \int_0^{l-x} u\left(t - \frac{y}{V}\right) dy + \int_0^{x} \left(K - \frac{1}{W} v\left(t - \frac{y}{W}\right)\right) dy$$

It is straightforward to obtain:

$$\frac{dF(x)}{dx} = \frac{-1}{V} u\left(t - \frac{l-x}{V}\right) + K - \frac{1}{W} v\left(t - \frac{x}{W}\right)$$

Since

$$u\left(t - \frac{l-x}{V}\right) \leq \frac{KVW}{V+W}$$
$$v\left(t - \frac{x}{W}\right) \leq \frac{KVW}{V+W}$$

hence

$$\frac{dF(x)}{dx} \geq 0$$

Therefore $F(x)$ is monotonically increasing in $(0 \leq x \leq l)$, which leads to:

$$F(0) \leq F(x) \leq F(l)$$

From the definition of number of vehicles in the congested and non-congested regime:

$$n^c(i) = \lim_{t\to(i+1)^-} \int_0^l K - \frac{1}{W} v\left(t - \frac{x}{W}\right) dx = \lim_{t\to(i+1)^-} F(l)$$

$$n^f(i) = \lim_{t\to(i+1)^-} \int_0^l \frac{1}{V} v\left(t - \frac{x}{V}\right) dx = \lim_{t\to(i+1)^-} F(0)$$

Since $F(x)$ is continuous on x, using the *fixed point principle*, for any value $n(i) \in (n^f(i), n^c(i))$ we can find x^* satisfying: $F(x^*) = n(i)$. This means that we can compute the dynamics of the length of the congested regime $l^c(i)$ given the (discrete) number of vehicles per link $n(i)$. □

Theorem 2. *Given l and a feasible value of $n(i)$. The solution of the length of the congested regime in Theorem 1 is unique if the following conditions hold:*

$$\begin{aligned} u(h) &< Q \quad \forall h \in [t - \frac{l}{V}, t] \\ v(h) &< Q \quad \forall h \in [t - \frac{l}{W}, t] \end{aligned}$$

where $Q = \frac{KVW}{V+W}$.

Proof. As in proof of Theorem 1, we use:

$$F(x) = \frac{1}{V}\int_0^{l-x} u(t - \frac{z}{V})dz + \int_0^x (K - \frac{1}{W}v(t - \frac{z}{W}))dz$$

and:

$$\frac{dF(x)}{dx} = K - \left(\frac{1}{V}u\left(t - \frac{l-x}{V}\right) + \frac{1}{W}v\left(t - \frac{x}{W}\right)\right)$$

With the above upper bound constraints on upstream flow and downstream flow, we derive that:

$$\begin{aligned} u\left(t - \frac{l-x}{V}\right) &< Q \quad \forall x \in [0, l] \\ v\left(t - \frac{x}{W}\right) &< Q \quad \forall x \in [0, l] \\ \Longrightarrow \frac{1}{V}u\left(t - \frac{l-x}{V}\right) + \frac{1}{W}v\left(t - \frac{x}{W}\right) &< Q\left(\frac{1}{V} + \frac{1}{W}\right) = K \quad \forall x \in [0, l] \end{aligned}$$

It guarantees the positive value $\frac{dF(x)}{dx} > 0 \quad \forall x \in [0, l]$ at a given time t, meaning that:

$$x < x^* \Leftrightarrow F(x) < F(x^*) \quad \forall x, x^* \in [0, l]$$

This implies that for a feasible value of $n(t)$, there is a unique solution of x so that $F(x) = n(t)$. □

The condition of unique length of the congested regime will avoid the critical point which is related to critical density and maximum flow in the triangular fundamental diagram. If this condition happens, it assures that maximum flow does not appear at any location on the link at a given time and there is unique solution of TTM at that time. A simple way to achieve this condition is that the dynamic maximum flow $q(t)$ is bounded to be strictly less than Q. Furthermore, even if the flow can reach Q at some times, but two bounds of the number of vehicles per link satisfying the strict inequality at given time t, e.g. $n^f(t) < n(t) < n^c(t)$, it also implies the condition in Theorem 2: there is an unique solution of the length of the congested regime at this time.

Theorem 3 (Non-congested condition)**.**

$$l^c(i) = 0 \quad \Leftrightarrow \quad n^f(i) = \sum_{k=0}^{i} [u_a(k) - v_a(k)] \tag{20}$$

Proof.

- Necessary condition: If $l^c(i) = 0$, traffic along the whole link is in the non-congested state, we have:

$$\sum_{k=0}^{i} [u(k) - v(k)] = \lim_{t \to (i+1)^-} \frac{1}{V} \int_0^l u\left(t - \frac{x}{V}\right) dx = \lim_{t \to (i+1)^-} \int_{t-\frac{l}{V}}^{t} u(x)dx$$
$$= \sum_{k=i+1-\frac{l}{V}}^{i} u(k) = n^f(i)$$

- Sufficient condition: If $n^f(i) = \sum_{k=0}^{i} [u(k) - v(k)]$: From Lemma 2, we infer that the equality happens when:
  - $l^c(t) = 0$, or
  - $u(t - \frac{l-x}{V}) = v\left(t - \frac{x}{W}\right) = \frac{KVW}{V+W}; \forall x \in [0, l^c(t)]$. This condition happens when traffic is operating at the critical density, at which $l^c(t) = 0$ by definition .

□

Theorem 4 (Fully-congested condition)**.**

$$l^c(i) = l \quad \Leftrightarrow \quad n^c(i) = \sum_{k=0}^{i} [u(k) - v(k)] \tag{21}$$

Proof. Based on Lemma 1, with similar method in Theorem 3. □

These two Theorems indicate that when the number of vehicles per link is equal to its bounds, the link is either fully non-congested or congested. The algorithm below shows how to compute the dynamics of the length of the congested regime in case the number of vehicles per link is between its bounds.

Let us define,

$$\begin{aligned} -j^f &= \lceil i + 1 - \frac{l^f(i)}{V} \rceil \\ j^c &= \lceil i + 1 - \frac{l^c(i)}{W} \rceil \\ \Delta j^f &= j^f - \left(i + 1 - \frac{l^f(i)}{V}\right) \\ \Delta j^c &= j^c - \left(i + 1 - \frac{l^c(i)}{W}\right) \end{aligned}$$

then we get:

$$\begin{aligned} n(t) &= \frac{1}{V} \int_0^{l^f(t)} u\left(t - \frac{x}{V}\right) dx + \int_0^{l^c(t)} \left(K - \frac{1}{W} v(t - \frac{x}{W})\right) dx \\ &= K.l^c(t) + \int_{t-\frac{l^f(t)}{V}}^{t} u(h)\mathrm{d}h - \int_{t-\frac{l^c(t)}{W}}^{t} v(h)\mathrm{d}h \\ n(i) = \lim_{t \to i+1^-} n(t) &= K.l^c(i) + \sum_{j^f}^{i} u(k) - \sum_{j^c}^{i} v(k) \\ &\quad + \Delta j^f.u(j^f - 1) - \Delta j^c.v(j^c - 1) \end{aligned}$$

It is worth mentioning that the following constraints are imposed for each regime length:

$$\begin{aligned} l^f(t) + l^c(t) &= l \\ l^f(t) &\geq 0 \\ l^c(t) &\geq 0 \end{aligned} \tag{22}$$

From the theories 3 and 4, we adopt the following algorithm to compute the length of the congested regime using the number of vehicles per link.

- If $n(i) = n^f(i)$, then $l^c(i) = 0$ (Theorem 3).
- If $n(i) = n^c(i)$, then $l^c(i) = l$ (Theorem 4).
- If $n^f(i) < n(i) < n^c(i)$, then let's set up $i^c = 1$ and $i^f = 0$.
  - Increase i^c until these conditions happen:

$$\begin{aligned} n(i) &\geq K.W.i^c - \sum_{i+1-i^c}^{i} v(k) + \sum_{j^f}^{i} u(k) + \Delta j^f.u(j^f - 1) \\ n(i) &< K.W.(i^c + 1) - \sum_{i-i^c}^{i} v(k) + \sum_{j^f}^{i} u(k) + \Delta j^f.u(j^f - 1) \end{aligned}$$

  - Similarly, increase i^f until these conditions happen:

$$\begin{aligned} n(i) &\leq K.l^c(i) - \sum_{j^c}^{i} v(k) - \Delta j^c.v(j^c - 1) + \sum_{i+1-i^f}^{i} u(k) \\ n(i) &> K.l^c(i) - \sum_{j^c}^{i} v(k) - \Delta j^c.v(j^c - 1) + \sum_{i-i^f}^{i} u(k) \end{aligned}$$

  - Solve the linear system of equations below to find $l^c(i)$ and $l^f(i)$.

$$\begin{aligned} l^c(i) &= i^c W + \Delta l^c \\ l^f(i) &= i^f V + \Delta l^f \\ \Delta l^c + \Delta l^f &= l - i^f V - i^c W \qquad \text{to satisfy constraint (22)} \\ n(i) &= K.l^c(i) + \sum_{i+1-i^f}^{i} u(k) - \sum_{i+1-i^c}^{i} v(k) \\ &\quad + \frac{\Delta i^f}{V}.u(i - i^f) - \frac{\Delta i^c}{W}.v(i - i^c) \end{aligned}$$

The above method is used as a benchmarking value for the linearized congested regime for the DSO problem presented in Section 5

4.5. Approximation of the congested regime

Let us assume that the time step is selected at at least l/W (Ban et al., 2012) so that $u(i)$ and $v(i)$ are steady in $i \in [0, \frac{l}{W} - 1]$, that is:

$$u(t_0) = u(t_0 + i) = u(i) \tag{23}$$

$$v(t_0) = v(t_0 + i) = v(i) \tag{24}$$

then, based on Theorem 1, we can obtain:

$$n(i) \approx l^c(i)\left(K - \frac{u(i)}{V} - \frac{v(i)}{W}\right) + \frac{u(i)l}{V}$$

So the approximation of l^c reads:

$$l^c(i) \approx \frac{n(i) - \frac{u(i)l}{V}}{K - \frac{u(i)}{V} - \frac{v(i)}{W}} \tag{25}$$

Using equations (23)-(24) for Lemmas 1 and 2 leads to:

$$n^c(i) = Kl - \sum_{k=i+1-\frac{l}{W}}^{i} v(k) \approx Kl - \frac{v(i)l}{W} \tag{26}$$

$$n^f(i) = \sum_{k=i+1-\frac{l}{V}}^{i} u(k) \approx \frac{u(i)l}{V} \tag{27}$$

To substitute equations (26)-(27) to equation (25) we obtain:

$$l^c(i) \approx \frac{n(i) - n^f(i)}{n^c(i) - n^f(i)} l \tag{28}$$

Equation (28) will be used in Section 5 for the new optimization framework, which optimally redistributes the vehicle queues (i.e. the dynamics of l^c) over links.

5. Formulation of DSO-DTA as an optimization problem with TTM based linear constraints

The DSO-DTA problem deals with a directed network $\mathbb{G}(\mathbb{V}, \mathbb{E})$ where $\mathbb{V}$ is the set of nodes and $\mathbb{E}$ the set of the directed links. During each time instant i in the study period $\mathbb{T}$, the single origin node $\mathbb{R}$ has time-dependent demand $D(i)$ heading toward the destination $\mathbb{S}$. The network is assumed to be empty at $i = 0$ and will be cleared at $i = T$. Suppose vehicles travel along each link according to the TTM principle and across any immediate node according to the demand-supply scheme Daganzo (1995). The DSO-DTA problem is to find an optimal dynamic traffic assignment pattern which results in the minimal total system travel time. In this paper, we only consider the DSO-DTA problem for a single origin-destination network. Nevertheless, the modeling and solution methods underpinned by an efficient implicit in-link congestion tracking approach for single origin-destination DSO problems proposed in the ensuing paper will still contribute to the state-of-the-art.

Recent research by Shen and Zhang (2014) have indicated that DSO may have multiple solutions in which all the solutions share the same objective while the difference is where the queues are formed and dissipated in the network. This section formulates a problem to find DSO solutions which can evenly distribute the queues (i.e. the length of the congested regime l^c) over links in the network so that the spillbacks can be avoided. To this end, we present a new optimization problem by adding a new step to the original DSO problem to optimally distribute the congestion regime over links in the network. The new step relies on how the congested length l^c is utilized if the front shock stays within link (note that this step cannot be modelled with the LTM or DQM).

5.1. Objective functions

1. **Stage 1**: find DSO solutions via minimizing the total travel times or the total system cost. The total system cost that all the vehicles spent in the network can be calculated as the total number of vehicles existing in the network for each time interval $i \in \mathbb{T}$, which is defined by the difference between the cumulative departure at the origin and the cumulative arrival at the destination for time interval i as below:
$$TT = \sum_{i\in\mathbb{T}} \left(\sum_{k=0}^{i} U(k) - \sum_{k=0}^{i} \sum_{a\in\mathbb{E}_S} u_a(k) \right) \tag{29}$$
Since the demand $U(i)$ is given $\forall i \in \mathbb{T}$, the objective to minimize the total travel cost TT is equivalent to maximize the total outflow (Shen and Zhang, 2008):
$$F = \sum_{a\in\mathbb{E}_S} \sum_{i\in\mathbb{T}} \sum_{k=0}^{i} u_a(k) = \sum_{a\in\mathbb{E}_S} \sum_{i\in\mathbb{T}} (T+1-i) u_a(i) \tag{30}$$
Let F^* denote the optimal value of F to achieve the DSO solutions. That is:
$$F^* = \max \sum_{a\in\mathbb{E}_S} \sum_{i\in\mathbb{T}} (T+1-i) u_a(i) \tag{31}$$
subject to the node and link constraints in sections 5.4 and 5.3.
2. **Stage 2**: find traffic patterns which optimally distribute the congestion in links.

   Objective: G, defined as the function of the queue length $l_a^c(i)$

   Constraints: $\sum_{a\in\mathbb{E}_S} \sum_{i\in\mathbb{T}} (T+1-i) u_a(i) = F^*$ (i.e. to achieve the DSO solutions).
   and the node and link constraints in Sections 5.4 and 5.3.

   This step restricts the solution domain to system optimal solutions via the first constraint. As reported in Shen and Zhang (2014), there exists a solution without changing the system optimal objective value. It indirectly shows that **Stage 2** provides a feasible solution domain. By optimizing G we can find a DSO solution that optimally distributes the congestion over links which can lead to minimal spillbacks or even free-flow situations as shown latter in our numerical study.

   Remark 1. *To account for the non-holding back issue, the new constraints in Equations* (38) *and* (39) *below can be added into the optimization framework above. Nevertheless, such constraints are nonlinear and we will show later in our example how they affect the nature and solutions of our framework.*

   Remark 2. *In this paper, we consider two specific objective functions G below for the illustration purposes:*

   (a) *Optimizing the heterogeneity of the queue lengths (**MDCL**) in the network*
$$G = \min \sum_{a\in\mathbb{A}} \Big(-\bar{l}^c + \sum_{i\in\mathbb{T}} \frac{l_a^c(i)}{l_a} \Big)^2 \tag{32}$$
   where,
$$\bar{l}^c = \frac{1}{|\mathbb{A}|} \sum_{a\in\mathbb{A}} \sum_{i\in\mathbb{T}} \frac{l_a^c(i)}{l_a} = \frac{1}{|\mathbb{A}|} \sum_{a\in\mathbb{A}} \sum_{i\in\mathbb{T}} \frac{n_a(i) - n_a^f(i)}{n_a^c(i) - n_a^f(i)} \tag{33}$$

(b) *Optimizing the total queue lengths (**MCL**) in the network:*

$$G = \min \sum_{a \in \mathbb{A}} \sum_{i \in \mathbb{T}} \frac{l_a^c(i)}{l_a} \tag{34}$$

5.2. Non-holding back problem

The non-holding back (NHB) solutions require the system to discharge flow as much as it can and have been widely discussed before in the literature (Shen et al., 2007, Shen and Zhang, 2014). For our particular TMM-DSO problem, the non-holding back condition is that, for any direct connection from link a to link b at any time $i \in \mathbb{T}$, at least one of the following conditions must hold:

$$v_a(i) = \underbrace{\min \left\{ Q_a(i); \sum_{k=0}^{i-l_a/V_a} u_a(k) - \sum_{k=0}^{i-1} v_a(k) \right\}}_{D_a(i)} \tag{35}$$

$$u_b(i) = \underbrace{\min \left\{ Q_b(i); K_b l_b + \sum_{k=0}^{i-l_b/W_b} v_b(k) - \sum_{k=0}^{i-1} u_b(k) \right\}}_{S_b(i)} \tag{36}$$

$$\sum_{k \geq i+1} f_{ab}(k) = 0 \tag{37}$$

These conditions indicate that, at any time i, if there is some traffic from link a to link b (i.e. $f_{ab}(i) > 0$), then the downstream flow of link a has to reach link-a demand (equation 35) or the upstream flow of link b has to reach link-b supply (equation 36). Otherwise, there will not be any traffic from link a to link b after time i (equation 37). In fact, conditions Equation (35) and Equation (36) follow the demand and supply principle of traffic flow at nodes.

Based on Lemma 1 and Lemma 2, these non-holding back conditions are represented by the following constraint:

$$\big(n_a(i) - n_a^f(i)\big)\big(n_b^c(i) - n_b(i)\big)\big(Q_a(i) - v_a(i)\big)\big(Q_b(i) - u_b(i)\big) \sum_{k \geq i+1} f_{ab}(k) = 0 \tag{38}$$

$\forall a, b \in \mathbb{E} : a \notin \mathbb{E}_R, b \notin \mathbb{E}_S, b \in \Upsilon_a^+$

$$n_a(i)\big(n_b^c(i) - n_b(i)\big)\big(Q_b(i) - u_b(i)\big) \sum_{k \geq i+1} f_{ab}(k) = 0 \tag{39}$$

$\forall a \in \mathbb{E}_R, b \in \Upsilon_a^+$

$$\big(n_a(i) - n_a^f(i)\big)\big(Q_a(i) - v_a(i)\big) \sum_{k \geq i+1} f_{ab}(k) = 0 \tag{40}$$

$\forall b \in \mathbb{E}_S, a \in \Upsilon_b^-$

Equation (38) guarantees the non-holding back traffic at any pair of normal links in the network (i.e. holding-free at very nodes inside the network), while equations (39) and (40) are used if one of these links is source or sink respectively.

Proposition 1. *System optimal solutions guarantee the holding-free traffic to sinks.*

Proof. Assume that there exists a SO solution such that it does not satisfy non-holding back condition at sinks. It means that there is a flow from normal link a to sink link $b \in \mathbb{E}_S$ at a particular time i so that:

$$\big(n_a(i) - n_a^f(i)\big)\big(Q_a(i) - v_a(i)\big) \sum_{k \geq i+1} f_{ab}(k) > 0$$

Let's consider flow from link a to link b. With the above conditions, there exists time $j > i$ so that $f_{ab}(j) > 0$ and while there is still a space for traffic to move from link a to destination link b because: $n_a(i) > n_a^f(i)$ and $Q_a(i) > v_a(i)$. For this reason, we can extract a small traffic δ at time j to move sooner at time i ($\delta < f_{ab}(j)$) without changing other flows (and also no violating to any other constraints):

$$f_{ab}^{\text{new}}(i) = f_{ab}(i) + \delta \Rightarrow u_b^{\text{new}}(i) = u_b(i) + \delta$$
$$f_{ab}^{\text{new}}(j) = f_{ab}(j) - \delta \Rightarrow u_b^{\text{new}}(j) = u_b(j) - \delta$$

This modification helps to improve the objective F:

$$F^{\text{new}} = F^* + (T + 1 - i)\delta - (T + 1 - j)\delta = F^* + (j - i)\delta > F^*$$

which violates the assumption that F^* is the maximum value of F. For this reason, the traffic to sink is always holding-free. That is equation (40) can be guaranteed by the SO objective. Note that, we still need to take care of holding-free traffic from sources. □

5.3. Link constraints

From the derivations in Section 4, the link constraints for the TTM are summarized as below: $\forall i \in \mathbb{T}$:

$$n_a(i) = \sum_{k=0}^{i} [u_a(k) - v_a(k)] \tag{41}$$

$$n_a^f(i) = \sum_{k=i-\frac{l_a}{V_a}+1}^{i} u_a(k) \tag{42}$$

$$n_a^c(i) = K_a.l_a - \sum_{k=i-\frac{l_a}{W_a}+1}^{i} v_a(k) \tag{43}$$

$$n_a(i) \geq n_a^f(i) \tag{44}$$

$$n_a(i) \leq n_a^c(i) \tag{45}$$

5.4. Node constraints

$\forall i \in \mathbb{T}$:

$$
\begin{aligned}
\sum_{a\in\Gamma_n^-} v_a(i) &= \sum_{a\in\Gamma_n^+} u_a(i) &(46)\\
u_a(i) &= \sum_{b\in\Upsilon_a^-} f_{ba}(i) &(47)\\
v_a(i) &= \sum_{b\in\Upsilon_a^+} f_{ab}(i) &(48)\\
f_{ab}(i) &\geq 0 &(49)
\end{aligned}
$$

Constraint (46) represents the conservation of the vehicles at an immediate node while constraint (47) determines the total number of vehicles entering a node from all upstream links of that node and constraint (48) defines the total number of vehicles exiting a node from all downstream links of that node. In fact, equation (46) and pair of equations (47) and (48) are equivalent, and we only need one of them in the final model. Constraint (49) guarantees the non-negative flow on link. It is worth noticing that no explicit node model is considered in this paper as we only focus on the DSO problem at this stage. We refer to Flötteröd and Rohde (2011), Ngoduy (2010), Ngoduy et al. (2005), Smits et al. (2015), Tampère et al. (2011) for node models in the network.

5.4.1. Source node constraints

$$
\begin{matrix} u_a(i) = U_a(i) \\ \{Q_a(i)\} \to \infty \end{matrix} \quad \forall a \in \mathbb{E}_R, i \in \mathbb{T} \tag{50}
$$

where $U_a(i)$ denotes the demand to source link a at time step i.

5.4.2. Sink node constraints

$$
\begin{matrix} v_a(i) = 0 \\ \{Q_a(i)\} \to \infty \end{matrix} \quad \forall a \in \mathbb{E}_R, i \in \mathbb{T} \tag{51}
$$

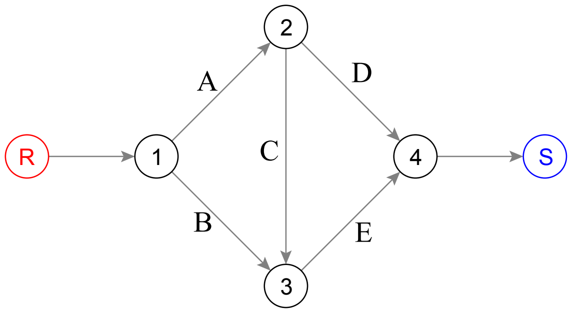
The above linear optimization problem can be solved by any simple optimization solver. In this paper, we use the CBC solver, one of the popular open-source software in solving optimization problems ($http://www.coin-or.org/projects/Cbc.xml$), for our case studies.

6. Numerical results

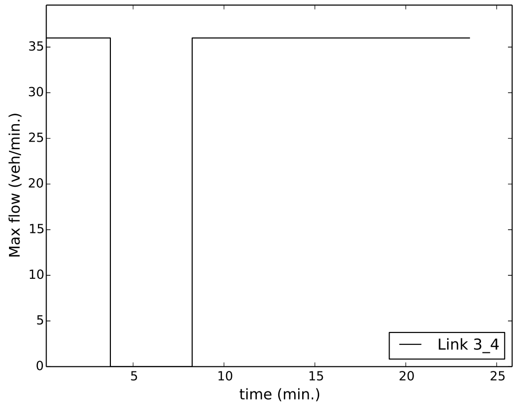
This section provides some numerical results showing the application of our proposed approach for the DSO-DTA problem in terms of accuracy and computational demand. To this end, we carry out two case studies: (1) using a simple Braess network to demonstrate in details the abilities of the TTM-DSO to avoid the queue spill-back problems, and (2) using a medium-sized grid network to compare the computational performance between TTM-DSO and CTM-DSO.

6.1. *Case study 1: Braess network*

Figure 2a shows the Braess network topology with a single origin-destination ($R - S$). Table 1 describes the network parameters for our study. According to the topology in Figure 2a, there are three paths with increasing free-flow travel time: (1,2,3,4), (1,3,4) and (1,2,4). Without disruption of link capacity, people would choose path (1,2,3,4) for their travelling. However, since link (3,4) is closed for some period of time (meaning that traffic is not permitted to enter this link, but its downstream flow is still allowed) as described in Figure 2b, it is essential to choose other paths to reach destination S as soon as possible.



(a) Braess network



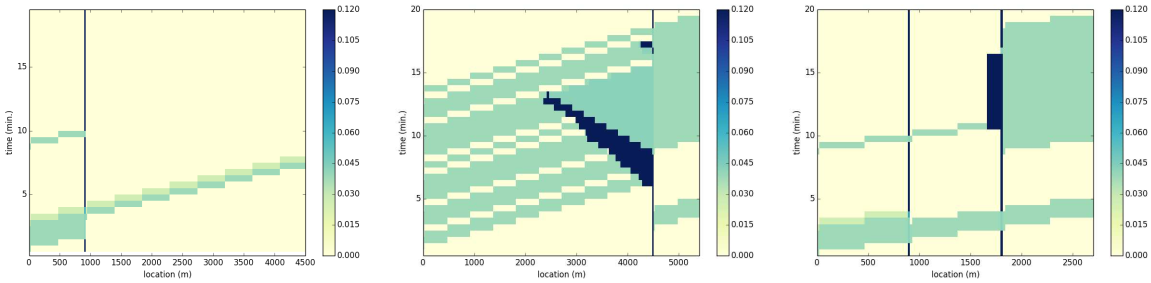
(b) Maximum flow setting

Figure 2: Braess network with the maximum flow setting at link (3,4).

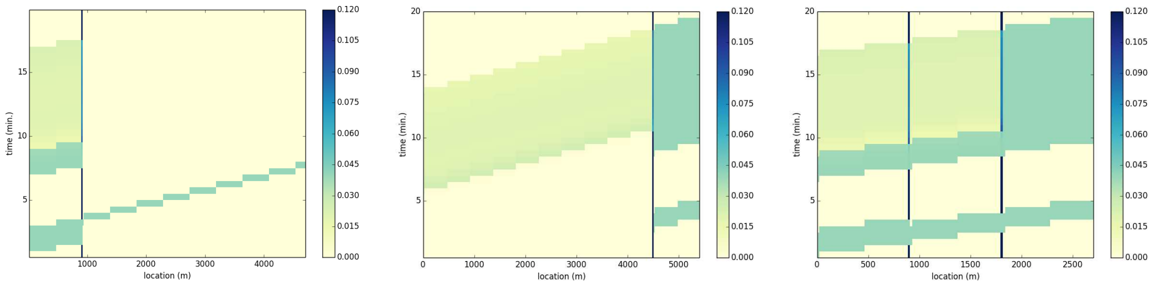
Traffic parameters (applied for all links):		
$V_a = 900$ (m/min) $W_a = 450$ (m/min)		
$K_a = 0.12$ (veh/m)		
Demand: 72 (veh/min), time-step: 0.25 (min).		
Links	l_a (m)	Q_a (veh/min.)
$(1,2)$	900	36
$(1,3)$	4500	36
$(2,3)$	900	36
$(2,4)$	14400	36
$(3,4)$	900	see Figure 2b

Table 1: Networking parameters for the Braess network in Figure 2.

Figure 3 shows different DSO solutions without using the NHB constraints, including the 1-step (original) DSO solution (Figures 3a to 3c) and the proposed 2-step DSO solution (Figures 3d to 3f). It is clear that the 2-step framework finds DSO solution which is free-flow in link 1-3 and link 2-3 at node



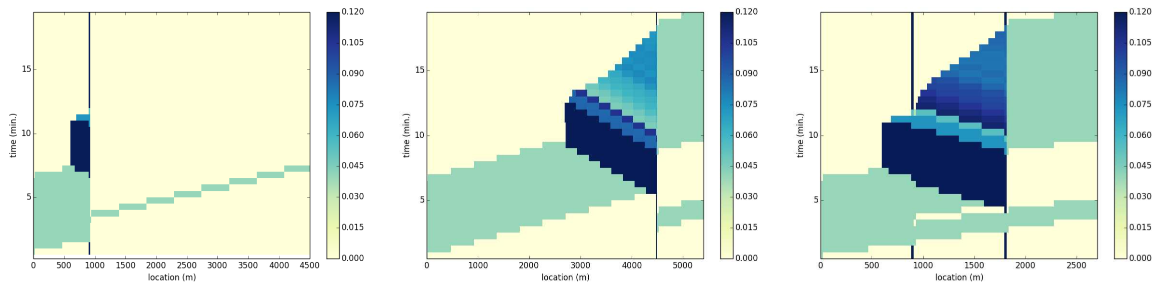
(a) Path (1,2,4) SO (b) Path (1,3,4) SO (c) Path (1,2,3,4) SO



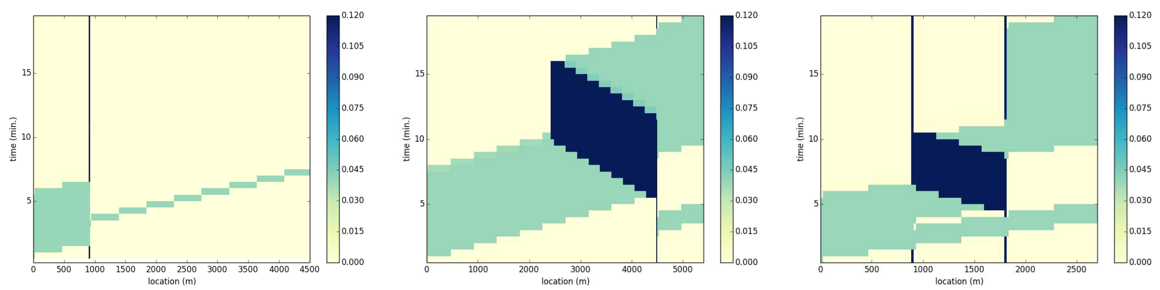
(d) Path (1,2,4) SO-MCL (e) Path (1,3,4) SO-MCL (f) Path (1,2,3,4) SO-MCL

Figure 3: Space-time density at three paths in DSO solutions without NHB constraints.

3 due to the capacity drop in link 3-4. This is obtained by minimizing the total queue lengths (**MCL**). Our 2-step DSO solution means that some vehicles are held at the source so that travellers will not have to wait in congestion inside the network. Obviously delay at the source (i.e. stay home and start the journey later) can be much better than waiting in congestion inside the network.



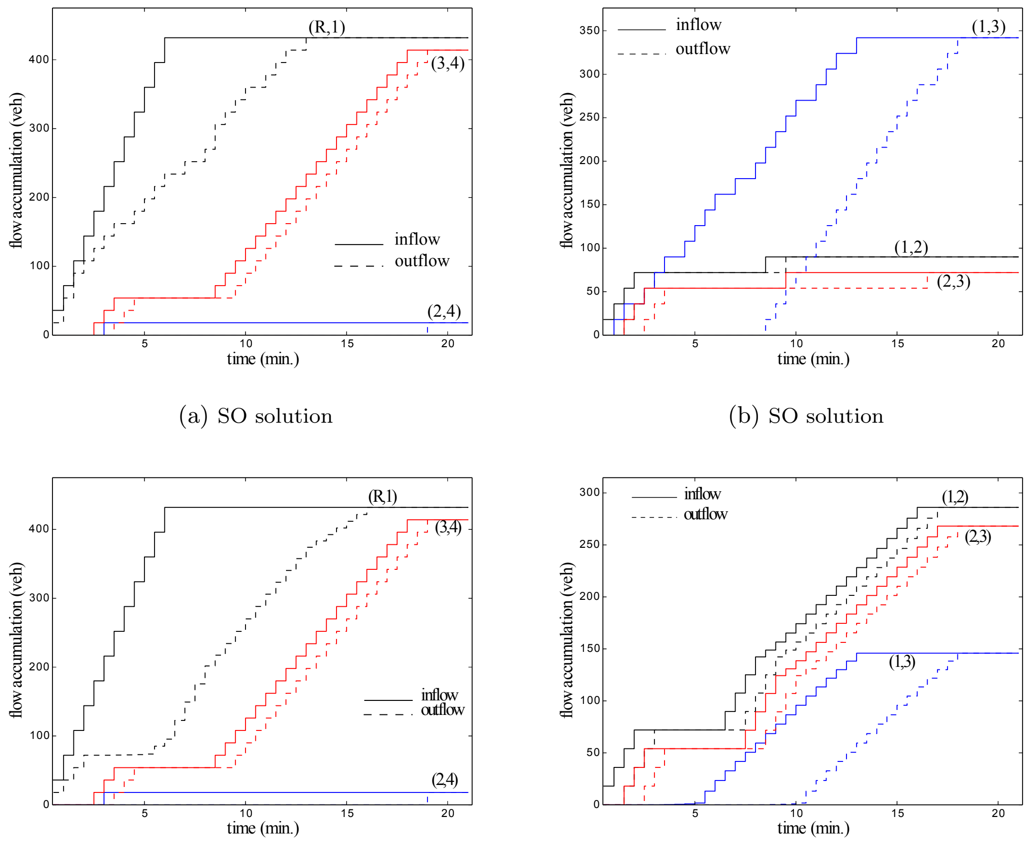
(a) Path (1,2,4) SO-NHB (b) Path (1,3,4) SO-NHB (c) Path (1,2,3,4) SO-NHB



(d) Path (1,2,4) SO-NHB-MDCL (e) Path (1,3,4) SO-NHB-MDCL (f) Path (1,2,3,4) SO-NHB-MDCL

Figure 4: Space-time density at three paths in DSO solutions with NHB constraints.

Figure 4 describes different DSO solutions using the NHB constraints. Since the nature of the NHB constraints is to push traffic going through the network as much as possible so there definitely heavy congestion inside the network, as seen in Figures 4a to 4c where the 1-step framework is used. With such NHB constraints, there are queue spillbacks from link 2-3 to link 1-2. However, our 2-step framework finds a DSO solution which can eliminate such spillback problem by reducing the queue in link 2-3 and increasing the queue in link 1-3 while making link 1-2 free-flowing. (Figures 4d to 4f). This is obtained by minimizing the total difference of the queue lengths (**MDCL**).



(a) SO solution

(b) SO solution

(c) SO-MCL solution

(d) SO-MCL solution

Figure 5: Inflow-outflow profiles for different DSO solutions.

The difference between DSO solutions can be further seen in Figure 5 describing the inflow-outflow profiles for different DSO solutions without NHB conditions. In Figure 5:

- The outflow of link R-1 shows how travellers enter the network while the outflows of link 3-4 and link 2-4 show how they exit the network
- The inflow of link R-1 shows the demand function which is the same in both solutions.

Figure 5 clearly shows the distinct difference of the inflow-outflow profiles between DSO solutions when the congestion inside the network (i.e. in link 1-3 and link 2-3) is eliminated (in the SO-MCL solution).

More specially, given the same demand (i.e. the inflow of link R-1), SO-MCL travellers do not suffer from the congestion in the network except the waiting time at source R, while SO travellers experience the congestion at link 1-3 and link 2-3.

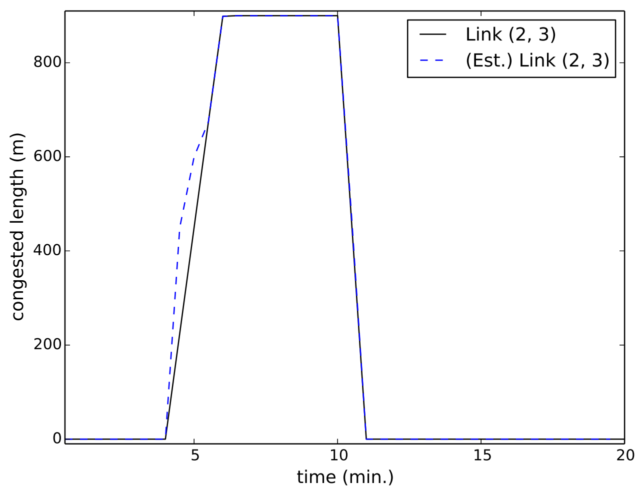


Figure 6: The numerical error of the congested length.

Figure 6 shows the accuracy of our approximation method for the congested length on link 2-3. While the solid line shows the evolution of the length of the congested regime which is computed from algorithm in Section 4.4, the dashed line describes the evolution of the length of the congested regime using an approximation algorithm in Section 4.5. Note that, while the algorithm in Section 4.4 is used after the DSO solution is obtained, the approximation algorithm in Section 4.5 is embedded in the DSO problem.

6.2. Case study 2: Grid network

This section compares the complexity and computational demand of the TTM-DSO against CTM-DSO in a larger network. To this end, we generate a scalable network based on grid topology in Figure 7. Table 2 describes four network configurations for the same physical network with only difference of time units (or the sampling time intervals). Note that smaller time step results in more accuracy of traffic dynamics but requires more computational time (more number of cells are used). The number in configuration name indicates a number of cells per link, for example $C5$ produces CTM topology with 5 cells/link (equivalently 8s time step). There are two types of link: normal and bottleneck links. The number of bottleneck links accounts for 20% of the total links. We examine the impact of time and space domain on CTM-DSO and TTM-DSO models by two different scenarios, both with one O-D pair from top-left node (R) to bottom-right node (S). The first scenario evaluates all configurations, where the time domain also changes to produce equivalent results. The second scenario focuses on some specific configurations with varying sizes of the time domain.

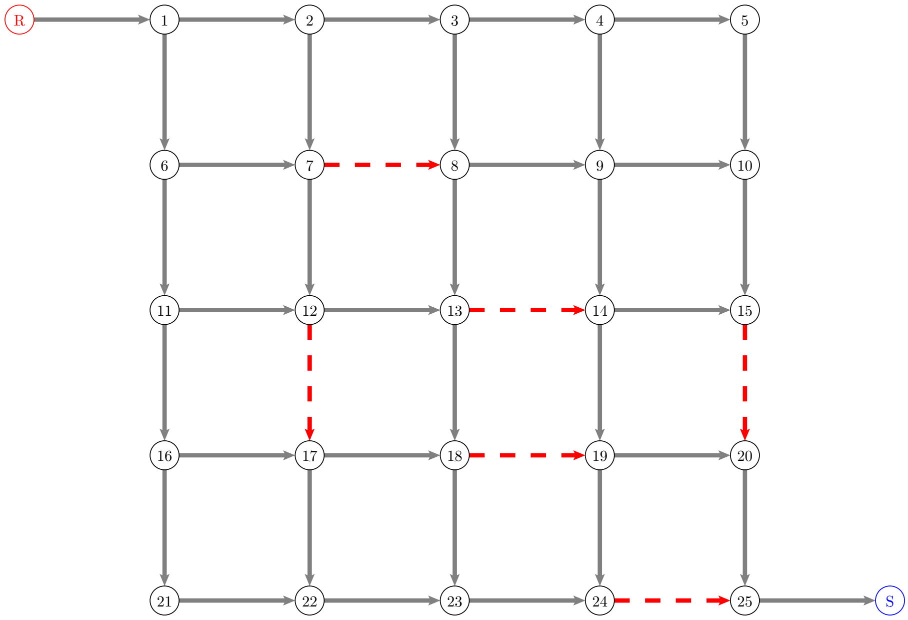


Figure 7: Grid network topology for TTM-DSO (the dashed lines show botttleneck links)

Constant parameters (applied for all links):			
$V_a = 15$ (m/s)		$W_a = 7.5$ (m/s)	
$l_a = 600$ (m), $\omega = \frac{W_a}{V_a} = 0.5$		$D^{rs} = 90$ (veh)	
Time units (sec.): $\{2, 4, 8, 20\} \approx \{C20, C10, C5, C2\}$			
Link types	K_a (veh/m)	Q_a (veh/sec.)	N (per cell) (veh/sec.)
Normal	0.12	0.6	1.8
Bottleneck	0.6	0.3	0.9

Table 2: Grid network with parameter sets for CTM-DSO and TTM-DSO in Figure 7

6.2.1. Estimation of complexity

We consider the complexity of the model in three aspects: i) number of constraints ($\mathcal{C}$), ii) number of variables ($\mathcal{V}$), and iii) computational time ($\mathcal{T}$). For TTM-DSO: $\mathcal{C}_{TTM} = O(|\mathbb{E}.\mathbb{T}|), \mathcal{V}_{TTM} = O(|\mathbb{E}.\mathbb{T}|)$. For CTM-DSO: $\mathcal{C}_{CTM} = O(|\mathbb{C}.\mathbb{T}|), \mathcal{V}_{CTM} = O(|\mathbb{C}.\mathbb{T}|)$, where $\mathbb{C}$ is the set of all cells in the CTM-DSO network.

In TTM-DSO, the variables are the sets of $u_a(i), v_a(i)$ and $f_{ab}(i), \forall a, b \in \mathbb{E}, i \in \mathbb{T}$. In our grid network example, apart from boundary nodes, all nodes have two incoming links and two outgoing links. For this reason, the number of inflow-outflow combinations is $2|\mathbb{E}|$. From that, the estimation of $\mathcal{V}_{TTM}$ is approximately $4|\mathbb{E}\mathbb{T}|$. The number of TTM-DSO constraints is $\mathcal{C}_{TTM} \approx 8|\mathbb{E}\mathbb{T}|$.

To estimate complexity of CTM-DSO, we first take a look at its main constraints:

$$\begin{aligned} n_i(t+1) &= n_i(t) + \sum_{ji\in\mathbb{A}} y_{ji}(t) - \sum_{ij\in\mathbb{A}} y_{ij}(t) \\ \sum_{ji\in\mathbb{A}} y_{ji}(t) &\leq \min\{Q_i(t), \omega_i(N_i(t) - n_i(t))\} \quad \forall i \in \mathbb{C}, t \in \mathbb{T} \\ \sum_{ij\in\mathbb{A}} y_{ij}(t) &\leq \min\{Q_i(t), n_i(t)\} \end{aligned}$$

where $\mathbb{A}$ is the set of links in the cell network. The number of constraints is $\mathcal{C}_{CTM} \approx 5|\mathbb{CT}|$. To evaluate $\mathcal{V}_{CTM}$, we observer that the number of $n_i(t)$ is equal to $|\mathbb{CT}|$, while the number of $y_{ij}(t)$ is equal to $|\mathbb{AT}|$. Assume that each link is divided to k cells, then $|\mathbb{C}| \approx k|\mathbb{E}|$. In the grid topology, the estimation of the number of arcs in cell network is $|\mathbb{A}| \approx (k+1)|\mathbb{E}|$ ($|\mathbb{E}|$ is large enough). This analysis leads to the evaluation of $\mathcal{V}_{CTM} \approx (2k+1)|\mathbb{ET}|$. These estimations, summarized in Table 3, help us to have a quick calculation of the complexity of both models. From Table 3 we expect a good performance of TTM-DSO for any $k \geq 2$.

$\lvert\mathbb{E}\rvert = 40$ (grid network in Figure 7) k cells per link		
	$\mathcal{C}$	$\mathcal{V}$
TTM	$8\lvert\mathbb{ET}\rvert = 320\lvert\mathbb{T}\rvert$	$4\lvert\mathbb{ET}\rvert = 160\lvert\mathbb{T}\rvert$
CTM	$5k\lvert\mathbb{ET}\rvert = 200k\lvert\mathbb{T}\rvert$	$(2k+1)\lvert\mathbb{ET}\rvert = (80k+40)\lvert\mathbb{T}\rvert$

Table 3: Estimation of complexity in terms of constraints and variables (grid topology)

6.2.2. Preparation time

It is the amount of time necessarily for creating a real model of the grid network (case study 2) which is solved by the CBC solver. Since all models are implemented in an abstract language, like AMPL, GAMS or modern programming languages (for example Python), this time is necessary to transform these abstract models into concrete models, which are understood by the solver. It includes processing data input file and generating instance of abstract models with variables, constraints, and parameters. The cost of reading input file is really small because we optimize these files by reducing duplicated values, and no file exceeds 2 MB in our experiments.

Conf.	$\lvert\mathbb{T}\rvert$	$\mathcal{C}$		$\mathcal{V}$		$\mathcal{T}_{prepare}$ (s)		$\mathcal{T}_{solve}(s)$		Obj/$\lvert\mathbb{T}\rvert$ (veh)
C2	28	9104	11690	4204	5455	0.75	2.14	0.57	0.57	41.8
C5	70	22754	71216	10504	30391	2.90	45.59	2.20	5.33	43.7
C10	140	45504	282426	21004	116751	9.60	1186.34	8.00	50.12	44.4
C20	280	91004	1124846	42004	457471	33.61	24297.18	27.48	693.44	44.7

Table 4: Performance comparison between TTM (gray column) and CTM in different traffic details

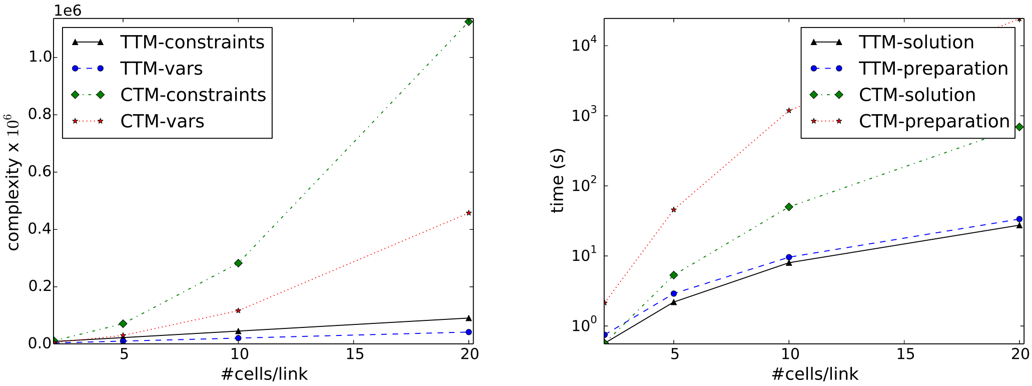


Figure 8: Performance comparison between TTM-DSO and CTM-DSO in different configurations

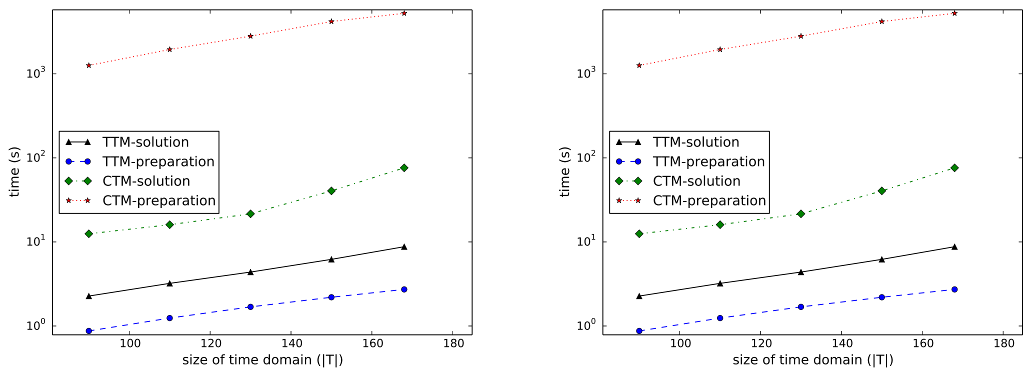


Figure 9: Performance comparison between TTM and CTM with different size of time domain (C20 $\sim$ 20 cells/link)

6.2.3. Comparison results

We compare the performance of TTM-DSO and CTM-DSO with respect to the complexity and preparation time defined above in two following scenarios:

Scenario 1:. Our target is to find the (system optimal) traffic pattern for a fixed demand (90veh) from R to S. In our experiments, all demands need around 400s to travel in the grid. Obviously, decreasing time step leads to the increasing computational time by increasing the time domain for each configurations, see Table 4. Especially, for CTM-DSO, not only time domain but also space domain are increased. We use two cells per link in $C2$ up to 20 cells per link in $C20$. While TTM-DSO model changes only link parameters in new configurations, CTM-DSO model changes the whole network topology including the cell parameters. The results are shown in Table 4 and Figure 8. It is worth noticing that both models produce the same objective values in any configurations (the last column in Table 4). For complexity, the results are consistent with the estimation in Table 3. We observe that the number of constraints and variables are increased linearly in TTM-DSO, while they are increased polynomially in CTM-DSO. Particularly, in the most complex configuration $C20$, TTM-DSO solves the problem 20 times faster than CTM-DSO. Looking at the trend of all complexity aspects in Figure 8, CTM-DSO complexity is always increased at higher rate than TTM-DSO, which means that TTM-DSO can provide more accurate traffic dynamics pattern at the equivalent cost.

Scenario 2:. In stead of increasing the complexity in both time and space as in the first scenario, we use the same configuration for all test cases to ignore the impact of space, and explore the influence of time domain on these models. Accordingly, both $\mathcal{C}$ and $\mathcal{V}$ of TTM-DSO and CTM-DSO are increased linearly in Figure 9. We first use the most complex configuration, $C20$, with $|\mathbb{T}|$ from 160 to 300. Since the fastest travel time is more than 320s, there is no movement of traffic with $|\mathbb{T}|$ less than 160. With initial demand of 90 vehicles, all travellers will finish their trips in about 200 time slots in $C10$ (6.5 minutes equivalently). From Figure 9, TTM-DSO outperforms CTM-DSO where all criteria of complexity show its benefit, especially, their increasing rates are also smaller.

In both scenarios, the preparation time is described in Figure 8 and 9. It shows that TTM-DSO always finishes the preparation period quickly. On the other hand, CTM-DSO preparation time dominates the solution time at all cases. Although we only verify this observation in our programming environment, it shows that TTM-DSO outperforms CTM-DSO in terms of efficiency.

7. Conclusions

The KWM has been widely used in the Dynamic System Optimum (DSO) traffic assignment problem to capture the dynamics of traffic in space and time along links in the network. As DSO may have multiple solutions which can achieve the same DSO objective (i.e. the total travel times) and the difference is where the queues are formed/dissipated in the network, this paper has presented an optimization framework to find DSO solutions which can achieve better traffic operation inside the network. Our new optimization framework contains two steps: i) find DSO solutions via minimizing the total travel times or the total system cost, ii) find traffic patterns which optimally distribute the congestion over links in the network. To fulfill Step 2, there is a need to determine the time and (horizontal) space evolution of the queues in the network. To this end, we have formulated the Two regime Transmission Model (TTM)-an extended version of the Link Transmission Model (LTM)-as a set of side linear constraints for a DSO problem. Basically, TTM models the dynamics of the two regimes of traffic flow namely, non-congested and congested states using a simplified Newell's theory as LTM does and describes the traffic states in time and space using just two regimes of the link as opposed to denser spatial discretisation adopted in CTM. In the proposed optimization framework, the TTM utilizes the same variables at the two ends of the link as the LTM (i.e. the entry and exit flow) in the DSO. However, the evolution of the length of the congested regime within link tracked by the TTM is embedded in Step 2.

We have carried out case studies to illustrate the performance of our framework. On the one hand, we have shown that optimally distributing the congestion over links in the network can help better traffic operation inside the network via minimal spillbacks or free-flow situations. On the other hand, we have shown that the complexity of the TTM-DSO including the number of the side constraints, the number of the variables as well as the computational time is only increased linearly with time steps. The numerical results have indicated that the TTM-DSO model outperforms the CTM-DSO model in terms of the complexity. Although the proposed framework currently only applies for the single origin-destination network, it plays an important role in the state-of-the-art and paves a way for future work covering more generic scenario of networks and traffic demands.

Acknowledgments

This research was financially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) Career Acceleration fellowship grant EP/J002186/1 and the Australian Research Council (ARC) Future fellowship grant FT120100723.

References

Aubin, J. P., Bayen, A. M., Saint-Pierre, P., 2008. Dirichlet problems for some hamilton-jacobi equations with inequality constraints. SIAM Journal on Control and Optimization 47, 2348–2380.

Aubin, J. P., Bayen, A. M., Saint-Pierre, P., 2011. Viability theory: new directions. Springer.

Balijepalli, N., Ngoduy, D., Watling, D., 2014. The two-regime transmission model for network loading in dynamic traffic assignment problems. Transportmetrica A: Transport Science 10 (7), 563–584.

Ban, X., Pang, J. S., Liu, X., Ma, R., 2012. Modeling and solution of continuous-time instantaneous dynamic user equilibria: A differential complementarity systems approach. Transportation Research Part B 46, 389–408.

Bar-Gera, H., 2005. Continuous and discrete trajectory models for dynamic traffic assignment. Networks and Spatial Economics 5, 41–70.

Beard, C., Ziliaskopoulos, A., 2006. A system optimal signal optimization formulation. Transportation Research Record 1978, 102–112.

Carey, M., Ge, Y. E., 2012. Comparison of methods for path flow reassignment for dynamic user equilibrium. Networks and Spatial Economics 12, 337–376.

Carey, M., Watling, D., 2012. Dynamic traffic assignment approximating the kinematic wave model: System optimum, marginal costs, externalities and tolls. Transportation Research Part B 46, 634–648.

Claudel, C. G., Bayen, A. M., 2010a. Lax-hopf based incorporation of internal boundary conditions into hamilton-jacobi equation. part 1: Theory. IEEE Transactions on Automatic Control 55, 1142–1157.

Claudel, C. G., Bayen, A. M., 2010b. Lax-hopf based incorporation of internal boundary conditions into hamilton-jacobi equation. part 2: Computational methods. IEEE Transactions on Automatic Control 55, 1158–1174.

Daganzo, C. F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B 28 (4), 269–287.

Daganzo, C. F., 1995. The cell transmission model, part 2: network traffic. Transportation Research Part B 29 (2), 79–93.

Daganzo, C. F., 2005. A variational formulation of kinematic waves: basic theory and complex boundary conditions. Transportation Research Part B 39, 187–196.

Daganzo, C. F., 2006. On the variational theory of traffic flow: well-posedness, duality and application. Network and heterogeneous media 1, 601–619.

Evans, L. C., 2010. Partial differentital equations: Second Edition, Graduate Studies in Mathematics. Ametican Mathematical Society.

Flötteröd, G., Rohde, J., 2011. Operational macroscopic modeling of complex urban road intersections. Transportation Research Part B 45, 903–922.

Gentile, G., Meschini, L., Papola, N., 2005. Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. Transportation Research Part B 39, 319–338.

Gentile, G., Meschini, L., Papola, N., 2007. Spillback congestion in dynamic traffic assignment: A macroscopic flow model with time-varying bottlenecks. Transportation Research Part B 41, 1114–1138.

Han, K., Piccoli, B., Szeto, W. Y., 2015. Continuous-time link-based kinematic wave model: formulation, solution existence, and well-posedness. Transportmetrica B:Transport Dynamics, "DOI: 10.1080/21680566.2015.1064793".

Jin, W. L., 2015. Continuous formulations and analytical properties of the link transmission model. Transportation Research Part B 74, 88–103.

Li, I. Y., Ziliaskopoulos, A. K., Waller, S. T., 2003. A decomposition scheme for system optimal dynamic traffic assignment models. Networks and Spatial Economics 3 (4), 441–455.

Lighthill, M. H., Whitham, G. B., 1955. On kinematic waves 2: A theory of traffic flow on long, crowded roads. In: Proceedings of the Royal Society of London. A 229. pp. 317–345.

Lo, H. K., Szeto, W. Y., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. Transportation Research Part B 36, 421–443.

Ma, R., Ban, X., Pang, J. S., 2014. Continuous-time dynamic system optimal for single-destination traffic networks with queue spillbacks. Transportation Research Part B 68, 98–122.

Newell, G. F., 1993. A simplified theory of kinematic waves in highway traffic (part i, ii, iii). Transportation Research Part B: Methodological 27 (4), 281–313.

Ngoduy, D., 2010. Multiclass first order modelling of traffic networks using discontinuous flow-density relationships. Transportmetrica 6, 121–141.

Ngoduy, D., Hoogendoorn, S. P., van Lint, J. W. C., 2005. Modeling traffic flow operation in multilane and multiclass urban networks. Transportation Research Record 1923, 73–81.

Nie, X., Zhang, H. M., 2005. A comparative study of some macroscopic link models used in dynamic traffic assignment. Networks and Spatial Economics 5 (1), 89–115.

Osorio, C., Flotterod, G., 2014. Capturing dependency among link boundaries in a stochastic dynamic network loading model. Transportation Science 49, 420–431.

Osorio, C., Flotterod, G., Bierlaire, M., 2011. Dynamic network loading: a stochastic differentiable model that derives link state distributions. Transportation Research Part B 45, 1410–1423.

Peeta, S., Ziliaskopoulos, A. K., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. Networks and Spatial Economics 1 (3-4), 233–265.

Qian, Z. S., Shen, W., Zhang, H. M., 2012. System-optimal dynamic traffic assignment with and without queue spillback: its path-based formulation and solution via approximate path marginal cost. Transportation Research Part B 46, 874–893.

Richards, P. I., 1956. Shock waves on the highway. Operations Research 4, 42–51.

Shen, W., Nie, Y., Zhang, H. M., 2007. A dynamic network simplex method for designing emergency evacuation plans. Transportation Research Record 2022, 83–93.

Shen, W., Zhang, H. M., 2008. What do different traffic flow models mean for system-optimal dynamic traffic assignment in a many-to-one network? Transportation Research Record 2008, 157–166.

Shen, W., Zhang, H. M., 2014. System optimal dynamic traffic assignment: Properties and solution procedures in the case of a many-to-one network. Transportation Research Part B 65 (10), 1–17.

Smits, E. S., Bliemer, M. C. J., Pel, A., van Arem, B., 2015. A family of macroscopic node models. Transportation Research Part B 74, 20–39.

Szeto, W., Jiang, Y., Sumalee, A., 2011. A cell-based model for multi-class doubly stochastic dynamic traffic assignment. Computer-Aided Civil and Infrastructure Engineering 26 (8), 595–611.

Szeto, W. Y., Lo, H. K., 2006. Dynamic traffic assignment: properties and extensions. Transportmetrica 2, 31–52.

Tampère, C. M. J., Corthout, R., Cattrysse, D., Immers, L. H., 2011. A generic class of first order node models for dynamic macroscopic simulation of traffic flows. Transportation Research Part B 45, 289–309.

Ukkusuri, S. V., Han, L., Doan, K., 2012. Dynamic user equilibrium with a path based cell transmission model for general traffic networks. Transportation Research Part B: Methodological 46 (10), 1657–1684.

Ukkusuri, S. V., Waller, S. T., 2008. Linear programming models for the user and system optimal dynamic network design problem: formulations, comparisons and extensions. Networks and Spatial Economics 8 (4), 383–406.

Yperman, I., Logghe, S., Immers, L. H., Tampere, C., 2005. Multicommodity link transmission model for dynamic network loading. In: The 85th Annual Meeting of the Transportation Research Board. Washington, D.C.

Ziliaskopoulos, A. K., 2000. A linear programming model for the single destination system optimum dynamic traffic assignment problem. Transportation science 34 (1), 37–49.