ORIGINAL PAPER

Discrete-time dynamic traffic assignment models with periodic planning horizon: system optimum

Artyom Nahapetyan · Siriphong Lawphongpanich

Abstract This paper proposes a system optimal dynamic traffic assignment model that does not require the network to be empty at the beginning or at the end of the planning horizon. The model assumes that link travel times depend on traffic densities and uses a discretized planning horizon. The resulting formulation is a nonlinear program with binary variables and a time-expanded network structure. Under a relatively mild condition, the nonlinear program has a feasible solution. When necessary, constraints can be added to ensure that the solution satisfies the First-In-First-Out condition. Also included are approximation schemes based on linear integer programs that can provide solutions arbitrarily close to that of the original nonlinear problem.

Keywords Dynamic traffic assignment · Time-varying network flows · System optimum

1 Introduction

Since Merchant and Nemhauser (see [29, 30]) first proposed their model in 1978, there have been a number of papers (see, e.g. [5, 12, 13, 16, 22–24, 26, 41, 42, 46]) discussing the variational inequalities or mathematical programming formulations for the dynamic traffic assignment problem with the assumption that the planning horizon is a set of discrete points instead of a continuous interval. Many of these papers use a dynamic or time-expanded network (see, e.g. [1]) to simultaneously capture the topology of the transportation network and the evolution of traffic over time. Implicitly or

A. Nahapetyan (✉) · S. Lawphongpanich
Center for Applied Optimization, Industrial and Systems Engineering Department,
University of Florida, Gainesville, FL 32611-6595, USA
e-mail: artyom@ufl.edu

S. Lawphongpanich
e-mail: Lawphong@ise.ufl.edu

otherwise, these papers typically assume that there is no traffic at the beginning of the planning horizon (or at time zero) and that all trips must exit the network prior to the end. When there are cars at the time zero, the times at which these cars enter the network must be known in order to determine when they will exit the arcs on which they were travelling. In practice, data with such details do not generally exist.

There are two main factors that distinguish the models in papers referenced above. First, some (e.g. [4, 5, 12, 17, 20, 22, 24, 29, 46]) seek a system optimal solution and others (e.g. [2, 13, 16, 18, 23, 36, 37, 42]) compute a user equilibrium instead. The other factor is the travel cost function used by these models. Among other parameters (physical or otherwise), a travel time or cost function may depend on the number of cars on the link and the input and output rates. Many (e.g. [6–8, 10, 15, 21, 28]) have analyzed the effects of travel cost functions on various models. Some (e.g. [21, 28]) have shown that some travel cost functions are not consistent with the models that use them.

Similar to Carey and Srinivasan [11], Carey and Subrahmanian [12], Carey [5], Chen and Hsueh [13] and Koufman et al. [24], the model in this paper is based on the time-expanded network. However, instead of assuming that the network is empty at the beginning or at the end, this paper treats the planning horizon as a circular interval instead of linear. For example, consider the interval [0, 24], i.e. a 24-h planning horizon. When viewed in a linear fashion, it is typically assumed that there is no car in the network at times 0 and 24. In turn, this implies there is no travel demand after time $k < 24$. Otherwise, cars that enter the network after time k cannot reach their destinations by time 24, thereby leaving cars in the network at the end of the horizon. On the other hand, if there is a car entering a street at 23:55 h (11:55 p.m.) and exiting at 24:06 h (12:06 a.m., the next day) in a circular planning horizon, the exit time of this car would be treated as 00:06 h instead. When accounted in this manner, it is possible to determine the exit time for every car that is in the network at time zero without requiring any additional data. Additionally, models that view the planning horizon in a circular fashion are more general in that they include those with a linear planning horizon. By setting the travel demands and other variables during an appropriate time interval to zero, models with a circular planning horizon effectively reduce to ones with a linear horizon.

It is often argued that the number of cars at the beginning and the end of the horizon are small and solutions to DTA are not drastically affected by setting them to zero. When the paths that these cars use do not overlap, the argument is valid. However, when these cars have to traverse the same arc in reaching their destinations, the number of cars on the arc may be significant and ignoring it may lead to a solution significantly different from the one that accounts for all cars.

This paper makes two main assumptions. One requires the link travel time at time t to be a function of only the number of cars on the link at that time. Carey and Ge [10] show that the solutions of models using functions of this type converge to the solution of the Lighthill–Whitham–Richards model (see [27, 38]) as the discretization of links into smaller segments is refined. Because minimizing the total travel time or delay mitigates its occurrence, models discussed herein do not explicitly address spillback. On the other hand, the models can be extended to handle spillback using a technique similar to the one in [25] or an alternative travel time function that includes the effect of spillback (see, e.g. [33]). However, as indicated in the reference, using such a function may not lead to a model with a solution.

For the remainder, Sect. 2 defines the concept of periodic planning horizon. Section 3 formulates the system version of the discrete-time dynamic traffic assignment problem with periodic planning horizon or discrete-time dynamic traffic assignment (DTDTA) and prove that a feasible solution exists under a relatively mild condition. To our knowledge, there are only four papers [3, 40, 42, 45] that address the existence issue and some (see, e.g. [40, 45]) consider this samll number to be lacking. All four deal with user equilibrium problems instead of system optimal. Section 4 describes two linear integer programs that provide bounds for DTDTA. Section 5 presents numerical results for small test problems (we discuss algorithms for solving larger problems in a separate paper) and, finally, Sect. 6 concludes the paper.

2 Periodic planning horizon

The models in this paper assume that the planning horizon is a half-open interval of length T, i.e. $[0, T)$. Instead of viewing this interval in a linear fashion, the interval is treated in a circular manner as shown in Fig. 1. In doing so, time 0 and T are the same instant. For example, time 0:00 and 24:00 h (or midnight) are the same instant in a 24 h day. For this reason, T is excluded and the planning horizon is half-open. To make the discussion herein more intuitive, we often refer to the planning horizon as a 24 h day, i.e., $T = 24$. In theory, the planning horizon can be of any length as long as events occur in a periodic fashion. If an event (e.g., five cars enter a street) occurs at time t, then the same event also occurs at time $t + kT$, for all integer $k \geq 1$.

Because the planning horizon is circular, events occurring tomorrow are assumed to occur in the same interval that represents today. For example, consider a car that enters a street at $t_1 = 23{:}00$ h (or 11 p.m.) today and traverses the street until it leaves at $t_2 = 01{:}00$ h (or 1 a.m.) tomorrow. (see Fig. 2). In a circular planning horizon, these two events, a car entering and leaving a street, occur at time 23:00 h and 01:00 h in the same interval $[0, 24)$. In general, if a car enters a street at time $t_1 < T$ and takes $\tau < T$ units of time to traverse, then the two events are assumed to occur at t_1 and $\mathrm{mod}\{t_1 + \tau, T\}$ on the interval $[0, T)$.

3 Discrete-time dynamic traffic assignment problem with a periodic time horizon

Although, it is possible to formulate the dynamic traffic assignment problem with a periodic time horizon as an optimal control problem, solving it is typically troublesome (see, e.g. [32]). This section presents a discrete-time version of the problem in which
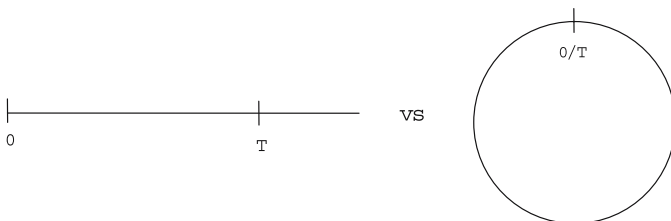

Fig. 1 Linear versus circular intervals

the interval $[0, T)$ is represented as a set of discrete points, i.e., $\Delta = \{0, \delta, 2\delta, \ldots, T-\delta\}$, where $\delta = \frac{T}{N}$ and N is a positive integer. (In general, the subdivision of the planning horizon need not be uniform. For example, the subdivision during the period between 22:00 and 06:00 h may be coarser than the one for the period between 06:00 and 22:00 h.) In order to avoid using fractional numbers in the set of indexes and to simplify our presentation, we typically assume that $\delta = 1$.

To formulate the problem, let $G(N, A)$ represent the underlying transportation network where N and A denote the set of nodes and arcs in the network, respectively. It is convenient to refer to elements of A either as a single index a or a pair of indices (i, j). The latter is used when it is necessary to reference the two ends of an arc explicitly. Furthermore, C is a set of origin-destination (OD) pairs and the travel demand for OD pair k during the time interval $[t, t + \delta], t \in \Delta$, is h_t^k.

There is also a travel time function associated with each arc in the network. In the literature (see, e.g. [9, 35, 43]), these functions can depend on a number of factors such as in-flow and out-flow rates and traffic densities. We assume in this formulation that ϕ_a, the travel time associated with arc a, depends only on the number of cars on the arc. Furthermore, ϕ_a is continuous, nondecreasing and bounded by T, i.e., $0 < \phi_a(w) < T, \forall w \in [0, M_a]$, where M_a is a sufficiently large upper bound for the range of $\phi_a(w)$ and there is no feasible solution whose flow on arc a can exceed M_a. In particular, $\phi_a(0)$ represents the free-flow travel time on arc a.

We use the dynamic or time-expanded (TE) network (see, e.g. Sect. 19.6 in [1]) to determine the state of vehicular traffic in the system at each time $t \in \Delta$. To illustrate the concept of time expansion, consider the static network with three nodes shown in Fig. 3 or the three-node network. In this network, all arcs have the same upper bound value, $M_a = M$, and there is only one OD pair $(1, 3)$. Let planning horizon be the interval $[0, 5)$ and $\delta = 1$. Thus, $\Delta = \{0, 1, 2, 3, 4\}$. The travel time function for every arc is ϕ and $1.5 \leq \phi(w) \leq 4, \forall w \in [0, M]$. To construct the TE network, the travel time also needs to be discretized. In general, the set of possible discrete travel times of arc a is $\Gamma_a = \{s : s = \lceil \frac{\phi_a(w)}{\delta} \rceil, 0 \leq w \leq M_a\}$. For our example, the set of possible discrete travel times for each arc is $\Gamma_a = \{2, 3, 4\}, \forall a$.
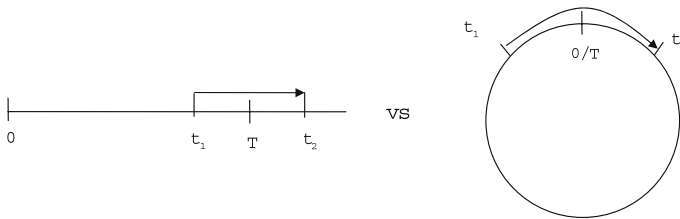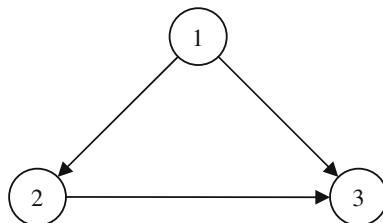

Fig. 2 Events occurring in two consecutive planning horizons

Fig. 3 Three-node network

To incorporate the time component in the TE network, every node in the static network (or static node) is 'expanded' or replicated once for each $t \in \Delta$. For the three-node network, static node 1 is transformed into five TE nodes, one for each $t \in \Delta$, in the TE network. For example, node 1 is expanded into nodes $1_0, 1_1, 1_2, 1_3$, and 1_4 in the TE network (see Fig. 4). Similarly, each arc (i,j) in the static network (or static arc) is replicated once for each pair of (t,s), where $t \in \Delta$ and $s \in \Gamma_{(i,j)}$. Consider arc $(1,2)$ in the three-node network. Cars that enter this arc at time 1 can take 2, 3 or 4 units of time to traverse depending (as assumed earlier) on the number of cars on the arc at $t = 1$. To allow all possibilities, arc $(1,2)$ is expanded into three TE arcs $(1_1, 2_3)$, $(1_1, 2_4)$, and $(1_1, 2_0)$. The latter represents a car that enters arc $(1,2)$ at time 1, takes four units of time to traverse, and leaves the arc at time 5 or 0 (or mod $(1 + 4, 5)$) of the following day. Similar expansion applies to each $t \in \Delta$. In general, each static arc (i,j) expands into $|\Delta| \times |\Gamma_{(i,j)}|$ TE arcs of the form $(i_t, j_{\text{mod} (t+s,T)}), \forall t \in \Delta, s \in \Gamma_{(i,j)}$.

Figure 5 displays the complete time expansion of the three-node network. In addition to the time-expanded nodes and arcs, the figure also displays the travel demand at the origin TE nodes (i.e., node $1_t, \forall t \in \Delta$) and decision variables $g^k_{d(k)_t}$ representing number of cars arriving at the destination node $d(k)$ of OD pair k at time t, i.e., at node $3_t, \forall t \in \Delta$.

To reference flows on TE arcs, let $y^k_{a(t,s)}$ denote the amount of flow for commodity k that enters static arc a at time $t \in \Delta$, takes $s \in \Gamma_a$ units of time to traverse it, and then exits the arc at time mod$\{t + s, T\}$. In particular, if $a = (i,j)$, then the subscript $a(t,s)$ refers to TE arcs of the form $(i_t, j_{\text{mod} (t+s,T)}), \forall \, t \in \Delta, s \in \Gamma_{(i,j)}$. In addition, $Y_{a(t,s)} = \sum_{k \in C} y^k_{a(t,s)}$ represents the total flow on arc $a(t,s)$.

To compute the time to traverse a static arc at time t, let

$$\Omega_{a(t)} = \{(\tau,s) : \tau = [t-1]_T, [t-2]_T, \ldots, [t-s]_T, s \in \Gamma_a\},$$

where

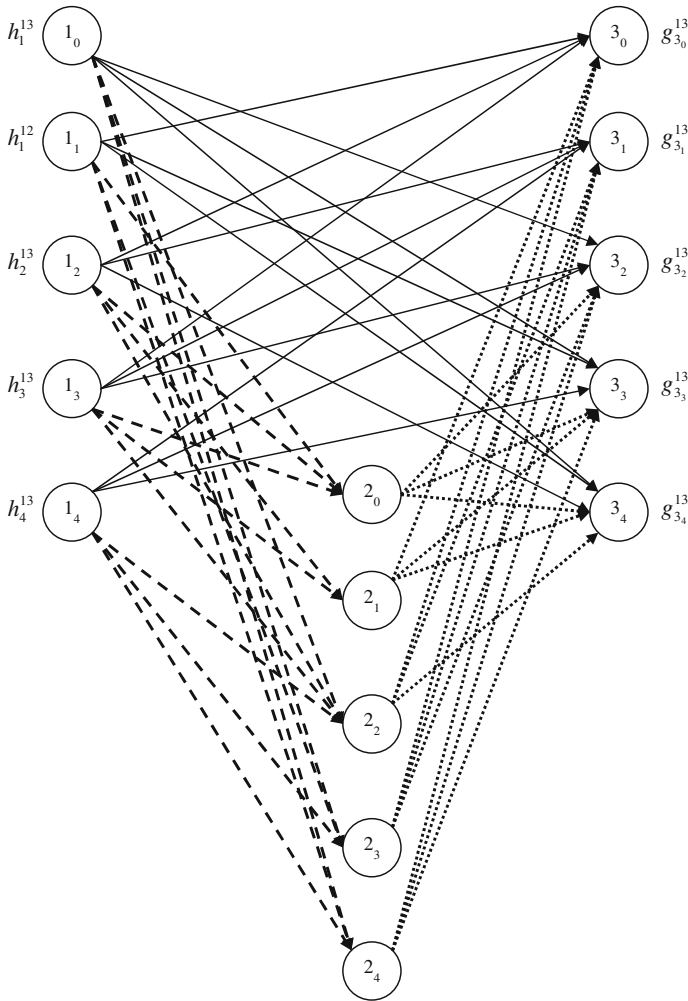**Fig. 4** Time expansion of arc $(1,2)$ at $t = 1$
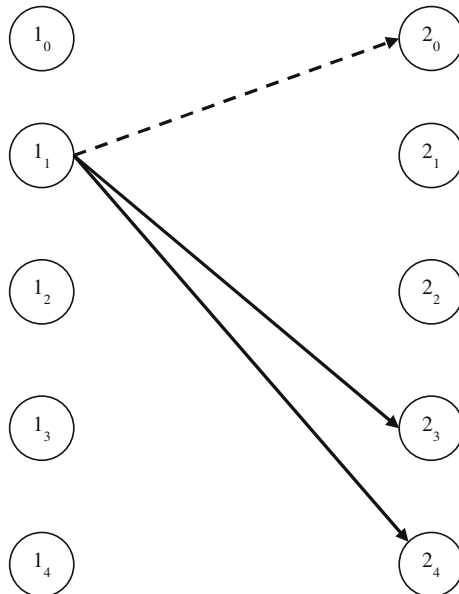
Fig. 5 Time-expansion of the three-node network

$$[q]_T = \begin{cases} q, & \text{if } q \geq 0, \\ T + q, & \text{if } q < 0. \end{cases}$$

In words, $\Omega_{a(t)}$ contains pairs of entrance, τ, and travel times, s, for static arc a such that, if a car enters static arc a at time τ and takes s time units to traverse it, the car will still be on the arc at time t. For example, if $t = 11{:}00$ h and the time to traverse arc a is 5 h for the previous five consecutive time periods, then cars entering arc a at time $\tau = 10{:}00$, $9{:}00$, $8{:}00$, $7{:}00$, and $6{:}00$ h will be on the arc at $11{:}00$ h. (We assume here that cars entering arc a at, e.g., $6{:}00$ h are still on the arc at $11{:}00$ h even though it is scheduled or expected to leave at $11{:}00$ h.) When t is relatively near the beginning of the planning horizon, the notation $[\cdot]_T$ accounts for cars on the arc at time t that enter it from the previous day. Continuing with the foregoing example, let $t = 3{:}00$ h instead. Then, cars entering arc a at time $\tau = 2{:}00$, $1{:}00$, $0{:}00$, $23{:}00$, and $22{:}00$ h are

still on the arc at 3:00 h. Using the set $\Omega_{a(t)}$, the total amount of flow on static arc a at time t or $x_{a(t)}$ is $\sum_{(\tau,s)\in\Omega_{a(t)}} Y_{a(\tau,s)}$.

There are two additional sets of decision variables. One set consists of $z_{a(t,s)}$, a binary variable that equals one if it takes between $(s-\delta)$ and s units of time to traverse arc a at time t. In the formulation below, the value of $z_{a(t,s)}$ depends on $x_{a(t)}$ and, for each t, $z_{a(t,s)} = 1$ for only one $s \in \Gamma_a$. The other set consists of g^k, a vector with a component for each node in the TE network. Component i_t of g^k is set to zero if i is not the destination node of OD pair k. Otherwise, $g^k_{d(k)_t}$, where $d(k)$ denotes the destination node of OD pair k, is a decision variable that represents the amount of flow for commodity k that reaches its destination, $d(k)$, at time t.

Below is a mathematical formulation of the discrete-time dynamic traffic assignment problem with periodic planning horizon (DTDTA).

$$\min_{(x,y,z,g)} \sum_{t\in\Delta}\sum_{a\in A} \left[\phi_a(x_{a(t)}) \sum_{s\in\Gamma_a} Y_{a(t,s)} \right],$$

subject to:

$$By^k + g^k = b^k, \quad \forall k \in C, \tag{1}$$

$$\sum_{t\in\Delta} g^k_{d(k)_t} = \sum_{t\in\Delta} h^k_t, \quad \forall k \in C, \tag{2}$$

$$Y_{a(t,s)} = \sum_{k\in C} y^k_{a(t,s)}, \quad \forall t \in \Delta, a \in A \text{ and } s \in \Gamma_a, \tag{3}$$

$$x_{a(t)} = \sum_{(\tau,s)\in\Omega_{a(t)}} Y_{a(\tau,s)}, \quad \forall t \in \Delta \text{ and } a \in A, \tag{4}$$

$$\sum_{s\in\Gamma_a} z_{a(t,s)} = 1, \quad \forall t \in \Delta \text{ and } a \in A, \tag{5}$$

$$\sum_{s\in\Gamma_a} (s-\delta)z_{a(t,s)} \le \phi_a(x_{a(t)}) \le \sum_{s\in\Gamma_a} sz_{a(t,s)}, \quad \forall t \in \Delta \text{ and } a \in A, \tag{6}$$

$$Y_{a(t,s)} \le M_a z_{a(t,s)}, \quad \forall t \in \Delta, a \in A \text{ and } a \in \Gamma_a, \tag{7}$$

$$y^k_{a(t,s)} \ge 0, g^k_{d(k)_t} \ge 0, x_{a(t)} \ge 0, z_{a(t,s)} \in \{0,1\}, \quad \forall t \in \Delta, a \in A, s \in \Gamma_a \text{ and } k \in C. \tag{8}$$

In the objective function, $\sum_{s\in\Gamma_a} Y_{a(t,s)}$ represents the number of cars that enter arc a at time t and, based on our assumption, these cars experience the same travel time, $\phi_a(x_{a(t)})$. Thus, the goal of this problem is to minimize the total travel time or delay. Using constraint (4), the objective function can be equivalently written as

$$\min_{(x,y,z,g)} \sum_{t\in\Delta}\sum_{a\in A} \left[\phi_a\left(\sum_{(\tau,s)\in\Omega_{a(t)}} Y_{a(\tau,s)} \right) \sum_{s\in\Gamma_a} Y_{a(t,s)} \right]$$

or, more compactly, as

$$\min \Phi(Y)^T Y,$$

where Y and $\Phi(Y)$ are vectors of arc flows $(Y_{a(t,s)})$ and travel times $(\phi_a(\sum_{(\tau,s)\in\Omega_{a(t)}} Y_{a(\tau,s)}))$ whose components are defined so that their inner product is consistent with the summations.

Constraint (1) ensures that flows are balanced at each node in the TE network. In this constraint, B denotes the node-arc incidence matrix of the TE network and b^k is a constant vector with a component for each TE node and defined as follows:

$$b_{i_t}^k = \begin{cases} 0, & \text{if } i \neq o(k), \\ h_t^k, & \text{if } i = o(k), \end{cases}$$

where $o(k)$ denotes the origin node of OD pair k. Constraint (2) guarantees that the number of cars arriving at the destination node $d(k)$ equals the total travel demand of OD pair k during the planning horizon. Then, constraint (3) computes the total flow on each TE arc and (4) determines the number of cars that are still on static arc a at time t.

In combination, the next three constraints, i.e., constraints (5)–(7), compute the travel time for the cars that enter arc a at time t and only allow flows to traverse the corresponding arc in the TE network. In particular, constraint (5), in conjunction with (6), chooses one (discretized) travel time $s \in \Gamma_a$ that best approximates $\phi_a(x_{a(t)})$, i.e., $\phi_a(x_{a(t)}) \in (s - \delta, s]$. When a represents arc (i, j), constraint (7) only allows arc $(i_t, i_{\text{mod}(t+s,T)})$ to have a positive flow. Otherwise, (7) forces flows on arc $(i_t, i_{\text{mod}(t+\tau,T)})$, for $\tau \in \Gamma_a$ and $\tau \neq s$, to be zero. Finally, constraint (8) makes sure that appropriate decision variables are either nonnegative or binary.

As formulated above, the travel time associated with $z_{a(t,s)}$ in Eq. 6 can only take on discrete values from the set Γ_a while the travel time in the objective function varies continuously. Although it may be more consistent to use discrete values of travel times in the objective function, the above model would provide a better solution because the true travel time is used to calculate the total delay. The model also has interesting properties discussed in Sect. 4. In addition, the treatments of travel times in both the objective function and constraints can be made consistent by solving the (approximation) refinement problem also discussed in the same section.

Under a relatively mild sufficient condition, we show below that DTDTA has a solution by constructing a feasible solution. In fact, the solution we construct below is generally far from being optimal. However, it suffices for the purpose of proving existence. Let $R_{a(t)}$ be a set of discrete times at which a car enters arc a and still remains on the arc at time t. Below, we refer to $R_{a(t)}$ as the *enter-remain* set. Given $x_{a(t)}$, $R_{a(t)} \subseteq \Delta$ is a union of two sets, i.e.,

$$R_{a(t)} = \left\{ w \in \Delta : w \leq (t-1), w + \left\lceil \frac{\phi(x_{a(w)})}{\delta} \right\rceil \geq t \right\} \cup \left\{ w \in \Delta : w \geq (t+1), w \right.$$
$$\left. + \left\lceil \frac{\phi(x_{a(w)})}{\delta} \right\rceil - T \geq t \right\}.$$

In addition, let $u_{a(t)}$ denote the total flow into arc a at time t. When $u_{a(t)}$ is given for each $t \in \Delta$, the lemma below shows that a set of $x_{a(t)}$, $Y_{a(t,s)}$, and $z_{a(t,s)}$ consistent with constraints (4)–(7) and relevant conditions in (8) exists when M_a is sufficiently large.

Lemma 1 *Assume that $u_{a(t)}$ is known for a given $a \in A$ and all $t \in \Delta$. If M_a is sufficiently large, then there exists a set of $x_{a(t)}$, $Y_{a(t,s)}$, and $z_{a(t,s)}$ that satisfies constraints* (4)–(7) *and the relevant conditions in* (8).

Proof (Because a is given, we discard the subscript a in places where there is no confusion in order to simplify our notation.) Below, we construct sequences $\{s_t^m\}$, $\{z_{a(t,s)}^m\}$, $\{Y_{a(t,s)}^m\}$, $\{x_{a(t)}^m\}$, and $\{R_t^m\}$ whose limits yield the set of decision variables feasible to constraints stated above.

For $m = 1$, let

- $s_t^1 = \lceil \phi_a(0)/\delta \rceil$, i.e., s_t^1 is the discretized free flow travel time for arc a;
- $z_{a(t,s_t^1)}^1 = 1$ and $z_{a(t,s)}^1 = 0, \forall s \in \Delta, s \neq s_t^1$;
- $Y_{a(t,s_t^1)}^1 = u_{a(t)}$ and $Y_{a(t,s)}^1 = 0, \forall s \in \Delta, s \neq s_t^1$;
- $R_t^1 = \{\omega \in \Delta : \omega \leq (t-1), \omega + s_\omega^1 \geq t\} \cup \{\omega \in \Delta : \omega \geq (t+1), \omega + s_\omega^1 - T \geq t\}$;
- $x_{a(t)}^1 = \sum_{\omega \in R_t^1} Y_{a(\omega,s_\omega^1)}^1$.

As defined above, R_t^1 is the enter-remain set based on the travel time s^1, a vector of $s_\omega^1, \forall \omega \in \Delta$. For $m \geq 2$, let

- $s_t^m = \lceil \phi_a(x_{a(t)}^{m-1})/\delta \rceil$;
- $z_{a(t,s_t^m)}^m = 1$ and $z_{a(t,s)}^m = 0, \forall s \in \Delta, s \neq s_t^m$;
- $Y_{a(t,s_t^m)}^m = u_{a(t)}$ and $Y_{a(t,s)}^m = 0, \forall s \in \Delta, s \neq s_t^m$;
- $R_t^m = \{\omega \in \Delta : \omega \leq (t-1), \omega + s_\omega^m \geq t\} \cup \{\omega \in \Delta : \omega \geq (t+1), \omega + s_\omega^m - T \geq t\}$;
- $x_{a(t)}^m = \sum_{\omega \in R_t^m} Y_{a(\omega,s_\omega^m)}^m$.

Sequences $\{s_t^m\}$, $\{x_{a(t)}^m\}$, and $\{R_t^m\}$ constructed above are monotonically nondecreasing. Consider the sequence $\{s_t^m\}$. Observe that $s_t^2 \geq s_t^1, \forall t \in \Delta$ because $x_{a(t)}^1 \geq 0, \forall t \in \Delta$, and as assumed earlier $\phi_a(\cdot)$ is nondecreasing. It follows that, for any $t \in \Delta$, $\omega + s_\omega^2 \geq \omega + s_\omega^1 \geq t$, and $\omega + s_\omega^2 - T \geq \omega + s_\omega^1 - T \geq t$. Thus, $\omega \in R_t^1$ implies that $\omega \in R_t^2$, i.e., $R_t^1 \subseteq R_t^2$ for all $t \in \Delta$. The latter and the fact that $u_{a(t)}$ is nonnegative imply that $x_{a(t)}^2 = \sum_{\omega \in R_t^2} Y_{a(\omega,s_\omega^2)}^2 \geq \sum_{\omega \in R_t^1} Y_{a(\omega,s_\omega^1)}^1 = x_{a(t)}^1, \forall t \in \Delta$.

Assume that the claim is true up to some fixed m. For all $t \in \Delta, s_t^{m+1} = \lceil \phi_a(x_{a(t)}^m)/\delta \rceil \geq \lceil \phi_a(x_{a(t)}^{m-1})/\delta \rceil = s_t^m$, because $x_{a(t)}^m \geq x_{a(t)}^{m-1}$ and $\phi_a(\cdot)$ is nondecreasing. Using an argument similar to above, $R_t^m \subseteq R_t^{m+1}$ and $x_{a(t)}^{m+1} \geq x_{a(t)}^m$. Thus, the three sequences are monotonically nondecreasing. In addition, all three sequences are bounded, i.e., $s_t^m < T$, $R_t^m \subseteq \Delta$, and $x_{a(t)}^m \leq \sum_{t \in \Delta} u_{a(t)}$ and, therefore, convergent. Let s_t^∞, R_t^∞, and $x_{a(t)}^\infty$ be their limits. Based on our construction, $s_t^\infty = \lceil \phi_a(x_{a(t)}^\infty)/\delta \rceil$, $z_{a(t,s_t^\infty)}^\infty = 1$, and $z_{a(t,s)}^\infty = 0, \forall s \in \Delta, s \neq s_t^\infty$. In combination, these ensure that constraints (5) and (6) are satisfied. Our construction also implies that $Y_{a(t,s_t^\infty)}^\infty = u_{a(t)}$ and $Y_{a(t,s)}^\infty = 0$, $\forall s \in \Delta, s \neq s_t^\infty$. Because M_a is sufficiently large, $Y_{a(t,s)}^\infty$ satisfies constraint (7).

In the limit, $R_t^\infty = \{\omega \in \Delta : \omega \leq (t-1), \omega + s_\omega^\infty \geq t\} \cup \{\omega \in \Delta : \omega \geq (t+1), \omega + s_\omega^\infty - T \geq t\}$. Thus, R_t^∞ is consistent with s^∞ and $x_{a(t)}^\infty = \sum_{\omega \in R_t^\infty} Y_{a(\omega,s_\omega^\infty)}^\infty = \sum_{(\tau,s) \in \Omega_{a(t)}} Y_{a(\tau,s)}^\infty$ because $Y_{a(t,s)}^\infty = 0, \forall s \in \Delta, s \neq s_t^\infty$. So, $x_{a(t)}^\infty$ satisfies (4). Furthermore, $x_{a(t)}^\infty$ and $Y_{a(t,s)}^\infty$ are both nonnegative and $z_{a(t,s)}^\infty$ is binary. Thus, the proof is complete. □

In the above proof, if, for some m, s_t^m is larger than the maximum travel time for arc a, i.e., $\max\{s : s \in \Gamma_a\}$ (or, equivalently, $x_{a(t)}^m > M_a$), then $u_{a(t)}$ is infeasible or not compatible with the upper bound M_a.

To establish the existence of a feasible solution to DTDTA, recall that $G(N, A)$ denotes the (static) transportation network. For the theorem below, assume without loss of generality that each node in N can be either an origin or destination, but not both. If node i is both an origin and a destination, then we create a dummy node i' and use node i as the origin node and i' as a destination. For example, consider OD pairs (i, j) and (j, i). In this case, i and j are both origins and destinations. When the dummy nodes are added, the two OD pairs become (i, j') and (j, i'). Let p^k denote a path in $G(N, A)$ connecting the OD pair k, i.e., $p^k \in P^k$. The set of these paths, $\Psi = \{p^k : k \in C\}$, induces a subgraph $G(\widehat{N}, \widehat{A})$, where $\widehat{N} \subseteq N$ and $\widehat{A} \subseteq A$ denote the sets of nodes and arcs, respectively, belonging to the paths in Ψ. For each $i \in \widehat{N}$, define $[i^+] = \{(i, j) : (i, j) \in \widehat{A}\}$ and $[i^-] = \{(j, i) : (j, i) \in \widehat{A}\}$. In words, $[i^+]$ and $[i^-]$ are the sets of arcs in $G(\widehat{N}, \widehat{A})$ that emanate from and terminate at node i, respectively. Also, let $order(i)$ denote a topological order of node i (see [1]). If $(i, j) \in \widehat{A}$ and $G(\widehat{N}, \widehat{A})$ can be topologically ordered, then $order(i) < order(j)$.

Theorem 1 *Assume that M_a is sufficiently large for all $a \in \widehat{A}$ and a node can be either an origin or a destination, but not both. Then, DTDTA has a feasible solution, if there exists a path p^k for each $k \in C$ such that the subgraph they induce is acyclic.*

Proof Let Ψ be a set of paths, one per OD pair, such that the subgraph, $G(\widehat{N}, \widehat{A})$, it induces has no cycle. Thus, $\widehat{N}$ can be ordered topologically (see [1]). Below, we construct a feasible solution one arc at a time and in a topological order using Lemma 1 and only the paths in Ψ. The latter implies that $Y_{a(t,s)} = y^k_{a(t,s)} = x_{a(t)} = 0$ for all $a \notin \widehat{A}$.

Let node $i \in \widehat{N}$ be of (topological) order 1 and, for each arc a in $[i^+]$, define $Q(a)$ to be the set of paths in Ψ that contain or use arc a, i.e., $Q(a) = \{k : a \in p^k, k \in C\}$. (It is not necessary to index $Q(a)$ with i because each arc a can belong to only one $[i^+]$.) For each $k \in Q(a)$, arc a must be the first arc in path p^k because node i is of order 1. Let $u_{a(t)} = \sum_{k \in Q(a)} h^k_t$. Because M_a is sufficiently large, Lemma 1 ensures that there exist $x_{a(t)}$, $Y_{a(t,s)}$, and $z_{a(t,s)}$ feasible to (4)–(7) and the relevant conditions in (8). Let $y^k_{a(t, s^\infty_{a(t)})} = h^k_t$ and $y^k_{a(t,s)} = 0, \forall s \in \Delta, s \neq s^\infty_{a(t)}$. So constructed, these $y^k_{a(t,s)}$'s are consistent with $Y_{a(t,s)}$ and satisfy the flow balance equation (1) for node i.

To construct the variables $x_{a(t)}$, $Y_{a(t,s)}$, $z_{a(t,s)}$, and $y^k_{a(t,s)}$ for arcs emanating from nodes of higher order, assume that the decision variables for arcs emanating from nodes with order m or less have been constructed and let node i be of order $(m + 1)$.

Case 1 The set $[i^+]$ is empty. Then, i must be a destination node for some commodity k, i.e., $i = d(k)$. For $\widehat{a} \in [i^-]$, $k \in Q(\widehat{a})$ and $t \in \Delta$, set

$$g^k_{d(k)_t} = \sum_{\{t:\ t + s^\infty_{a(t)} = t\}} y^k_{a(t, s^\infty_{a(t)})} + \sum_{\{t:\ t + s^\infty_{a(t)} - T = t\}} y^k_{a(t, s^\infty_{a(t)})}.$$

For each $k \in Q(\widehat{a})$, every demand h^k_t uses arc $\widehat{a}$. Thus, $g^k_{d(k)_t}$ as constructed must satisfy the appropriate constraints in (1) and (2).

Case 2 The set $[i^+]$ is not empty. Let $\widehat{a} \in [i^-]$, also a nonempty set. Assume that $\widehat{a} = (q, i)$. Then, $order(q) < order(i)$ and, by the above assumption, $x_{a(t)}$, $Y_{a(t,s)}$, $z_{a(t,s)}$, and $y^k_{a(t,s)}$ are available.

Consider an arc $a \in [i^+]$. For each $\widehat{a} \in [i^-]$, define $Q(\widehat{a}, a) = \{k : \widehat{a} \in p^k, a \in p^k, k \in C\}$ and, for each $k \in Q(\widehat{a}, a)$, let $u^k_{a(t)}$ denote the flow into arc a at time t for OD pair

k. Then,

$$u_{a(t)}^k = \sum_{\{t:\ t+s_{a(t)}^\infty=t\}} y_{a(t,s_{a(t)}^\infty)}^k + \sum_{\{t:\ t+s_{a(t)}^\infty-T=t\}} y_{a(t,s_{a(t)}^\infty)}^k$$

and the total flow into arc a at time t is $u_{a(t)} = \sum_{k\in Q(a,a)} u_{a(t)}^k$. Because M_a is sufficiently large, Lemma 1 ensures that $x_{a(t)}$, $Y_{a(t,s)}$, $y_{a(t,s)}^k$, and $z_{a(t,s)}$ feasible to relevant constraints exist.

Thus, when carried out in the topological order for every arc in $\widehat{A}$, the above process must produce a feasible solution to DTDTA.   □

The theorem above assumes that each M_a is sufficiently large so that it is feasible to send the entire flow for each OD pair along a single path. Although this assumption appears to be stringent, it can be made less so by allowing the flow for each OD pair to traverse over several paths as long as they do not induce cycles in $G(\widehat{N},\widehat{A})$. With more cumbersome notation, the above argument can be extended to the case with multiple paths per OD pair as well.

When applied to the above example in which the OD pairs (i,j) and (j,i) become (i,j') and (j,i'), the acyclic subgraph assumption implies that the paths from i to j' and from j to i' cannot form a cycle. Intuitively, this means that there must exist two routes between the original nodes i (e.g., home) and j (e.g., work) with no road in common. These routes need not be optimal and there is no requirement in our formulation or algorithms to use them. They are used only to establish the existence in Theorem 1.

The First-In-First-Out (FIFO) condition requires that cars entering an arc at time t must leave the arc before those entering after time t. In the literature, many (see, e.g. [33, 34, 45]) assume that the travel cost function satisfied certain conditions to ensure FIFO. To avoid making additional assumptions, we ensure FIFO by adding the following constraints to DTDTA instead. Doing so may make the problem harder to solve because of the additional constraints.

$$t + \sum_{s\in\Gamma_a} sz_{a(t,s)} \leq \bar{t} + \sum_{s\in\Gamma_a} sz_{a(\bar{t},s)}, \qquad \forall a \in A_1 \text{ and } t,\bar{t} \in \Delta: (t+\delta) \leq \bar{t},$$

$$\bar{t} + \sum_{s\in\Gamma_a} sz_{a(\bar{t},s)} \leq (t+T) + \sum_{s\in\Gamma_a} sz_{a(t,s)}, \quad \forall a \in A_1 \text{ and } t,\bar{t} \in \Delta: (t+\delta) \leq \bar{t}.$$

When t and $\bar{t}$ represent two instances of time on the same day, the first inequality ensures that cars entering arc a at time t leave the arc before those that enter at time $\bar{t} > t$. On the other hand, t and $\bar{t}$ may refer to times on consecutive days, e.g., $t = 08:00$ h today and $\bar{t} = 09:00$ h yesterday. Because of our periodic assumption, these two times are on the same interval [00:00 and 24:00 h] and t (incorrectly) appears to be an earlier time than $\bar{t}$. To distinguish times on consecutive days and preserve FIFO, the second equation represents today's time t (e.g., 08:00 h of today) as $(t+T)$ (e.g., as 08:00 h of yesterday plus T) and forces cars entering the arc at this time to depart after those that enter at yesterday's time $\bar{t}$ (e.g., 09:00 h yesterday).

4 Bounds for discrete-time dynamic traffic assignment

As formulated in the previous section, DTDTA is a nonlinear optimization problem with binary decision variables, a difficult class of problems to solve. This section

describes mixed integer programs for obtaining an approximate solution to DTDTA as well as bounds for the optimal delay.

Except for constraint (6), the constraints for DTDTA are linear. To develop a linear version of (6), assume that the travel time function, ϕ_a, is invertible for all $a \in A$. For example, if ϕ_a is a continuous and increasing function, then ϕ_a^{-1} exists on the interval $[\phi_a(0), \phi_a(M_a)]$ (see Fig. 6). Under this assumption, $\phi_a(x_{a(t)}) \in (s - \delta, s]$ if and only if $x_{a(t)} \in (\phi_a^{-1}(s - \delta), \phi_a^{-1}(\delta)]$. Thus, the requirement $(s - \delta)z_{a(t,s)} < \phi_a(x_{a(t)}) \leq s z_{a(t,s)}$ is equivalent to $\phi_a^{-1}(s - \delta)z_{a(t,s)} < x_{a(t)} \leq \phi_a^{-1}(s)z_{a(t,s)}$. Recall that $\Gamma_a = \{s: s = \lceil \frac{\phi_a(w)}{\delta} \rceil, 0 \leq w \leq M_a\}$. Let $s_1 = \lceil \frac{\phi_a(0)}{\delta} \rceil$. Then, $(s_1 - \delta) \notin [\phi_a(0), \phi_a(M_a)]$ and $\phi_a^{-1}(s_1 - \delta)$ is not well defined. (In Fig. 6, $\phi_a^{-1}(s_1 - \delta) = \phi_a^{-1}(1)$ is not well defined.) In this paper, we set $\phi_a^{-1}(s_1 - \delta) = 0$. Using this convention, constraint (6) can be replaced by the following linear equivalence:

$$\sum_{s \in \Gamma_a} \phi_a^{-1}(s - \delta)z_{a(t,s)} < x_{a(t)} \leq \sum_{s \in \Gamma_a} \phi_a^{-1}(s)z_{a(t,s)}, \quad \forall t \in \Delta \text{ and } a \in A. \quad (9)$$

The following lemma implies that there exist linear functions that approximate the objective function of DTDTA.

Lemma 2 *There exist vectors q_l and q_u such that $q_l^T Y \leq \Phi(Y)^T Y \leq q_u^T Y$ for all Y feasible to DTDTA.*

Proof As defined earlier, $\Phi(Y)^T Y = \sum_{t \in \Delta} \sum_{a \in A} \phi_a(x_{a(t)}) \left[\sum_{s \in \Gamma_a} Y_{a(t,s)} \right]$. From constraint (6), the following hold for any feasible solution to DTDTA:

$$\sum_{t \in \Delta} \sum_{a \in A} \left[\sum_{s \in \Gamma_a} (s - \delta)z_{a(t,s)} \right] \left[\sum_{s \in \Gamma_a} Y_{a(t,s)} \right] \leq \Phi(Y)^T Y$$

$$\leq \sum_{t \in \Delta} \sum_{a \in A} \left[\sum_{s \in \Gamma_a} s z_{a(t,s)} \right] \left[\sum_{s \in \Gamma_a} Y_{a(t,s)} \right].$$
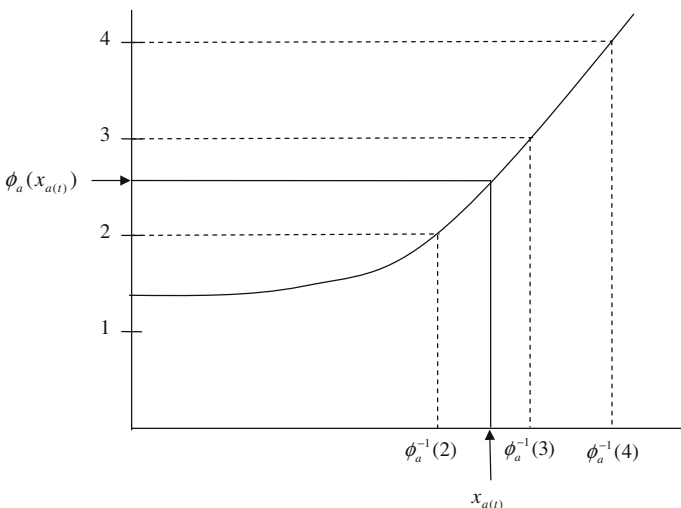

Fig. 6 $\phi_a(x_{a(t)}) \in (s - \delta, s]$ versus $x_{a(t)} \in (\phi_a^{-1}(s - \delta), \phi_a^{-1}(\delta)]$

The summand of the last set of summations (i.e., $\left[\sum_{s\in\Gamma_a} sz_{a(t,s)}\right]\left[\sum_{s\in\Gamma_a} Y_{a(t,s)}\right]$) can be simplified. Constraint (7) implies that $Y_{a(t,s)} > 0$ only if $z_{a(t,s)} = 1$. In addition, constraint (5) ensures that, for each pair (a,t), $z_{a(t,\bar{s})} = 1$ for some $\bar{s} \in \Gamma_a$ and $z_{a(t,s)} = 0, \forall s \in \Gamma_a, s \neq \bar{s}$. This implies that $Y_{a(t,\bar{s})} \geq 0$ and $Y_{a(t,s)} = 0, \forall s \in \Gamma_a, s \neq \bar{s}$ and

$$\left[\sum_{s\in\Gamma_a} sz_{a(t,s)}\right]\left[\sum_{s\in\Gamma_a} Y_{a(t,s)}\right] = \bar{s}Y_{a(t,\bar{s})} = \sum_{s\in\Gamma_a} sY_{a(t,s)}.$$

A similar result holds for the first set of summations. Thus, the above inequalities reduce to the following

$$\sum_{t\in\Delta}\sum_{a\in A}\sum_{s\in\Gamma_a} (s-\delta)Y_{a(t,s)} \leq \Phi(Y)^T Y \leq \sum_{t\in\Delta}\sum_{a\in A}\sum_{s\in\Gamma_a} sY_{a(t,s)}.$$

Let q_l and q_u be two constant vectors with a component for each arc in the TE network such that $[q_l]_{a(t,s)} = (s-\delta)$ and $[q_u]_{a(t,s)} = s$, respectively, for all $a \in A, t \in \Delta$, and $s \in \Gamma_a$. Then,

$$q_l^T Y = \sum_{t\in\Delta}\sum_{a\in A}\sum_{s\in\Gamma_a} (s-\delta)Y_{a(t,s)} \leq \Phi(Y)^T Y \leq \sum_{t\in\Delta}\sum_{a\in A}\sum_{s\in\Gamma_a} sY_{a(t,s)} = q_u^T Y. \qquad \square$$

Let $S(\delta)$ denote the feasible region defined by linear constraints (1)–(5), (7)–(9), and, for convenient, (Y, Z) represents an element in $S(\delta)$. In addition, let (Y^l, Z^l), (Y^*, Z^*), and (Y^u, Z^u) be solutions to the lower-bound problem (or $\min\{q_l^T Y : (Y,Z) \in S(\delta)\}$), the original problem (or $\min\{\Phi(Y)^T Y : (Y,Z) \in S(\delta)\}$), and the upper-bound problem (or $\min\{q_u^T Y : (Y,Z) \in S(\delta)\}$), respectively. Then, the following lemma holds.

Lemma 3 *For any $\delta > 0$, $q_l^T Y^l \leq \Phi(Y^*)^T Y^* \leq q_u^T Y^u \leq q_u^T Y^l$.*

Proof In following sequence of inequalities, the first one holds because Y^* is feasible to the lower-bound problem and the second follows from Lemma 2:

$$q_l^T Y^l \leq q_l^T Y^* \leq \Phi(Y^*)^T Y^*.$$

Similarly, the following sequence also holds

$$\Phi(Y^*)^T Y^* \leq \Phi(Y^u)^T Y^u \leq q_u^T Y^u.$$

Combining the above two sequences yield the first two inequalities in the lemma. Finally, the last inequality holds because Y^l is not necessarily optimal to $\min\{q_u^T Y : (Y,Z) \in S(\delta)\}$. $\qquad \square$

In view of the above lemma, the solutions to the upper and lower-bound problems are approximations of the solution to the original problem. The theorem below states that the approximation can be made arbitrarily close to the original problem by choosing a sufficiently small δ.

Theorem 2 *Given $\epsilon > 0$, there exists $\delta > 0$ such that $q_u^T Y^u - \Phi(Y^*)^T Y^* \leq \epsilon$ and $\Phi(Y^*)^T Y^* - q_l^T Y^l \leq \epsilon$.*

Proof By construction, $q_u = q_l + \delta e$, where e is $(1, 1, \ldots, 1)^T$. Let H^k denote the travel demand for OD pair k during the entire planning horizon, i.e., $H^k = \sum_{t\in\Delta} h_t^k$ and set

$H = \sum_{k \in C} H^k$. For each $t \in \Delta$, $s_a(t)$ is such that $s_a(t) \in \Gamma_a$ and $z^l_{a(t,s_a(t))} = 1$. In words, $s_a(t)$ is the approximate travel time for (static) arc a at time t in the optimal solution (Y^l, Z^l).

Then, the following sequence must hold:

$$
\begin{aligned}
0 &\leq (q_u - q_l)^T Y^l \\
&= \delta e^T Y^l \\
&= \delta \sum_{a \in A} \sum_{k \in C} \sum_{t \in \Delta} \sum_{s \in \Gamma_a} y^k_{a(t,s)} \\
&= \delta \sum_{a \in A} \sum_{k \in C} \sum_{t \in \Delta} y^k_{a(t,s_a(t))} \\
&\leq \delta \sum_{a \in A} \sum_{k \in C} H^k \\
&= \delta H \sum_{a \in A} 1 \\
&= \delta H |A|.
\end{aligned}
$$

The first inequality follows from Lemma 3. Then, the above relationship between q_u and q_l and letting $\sum_{k \in C} y^k_{a(t,s)}$ denote individuals components of Y^l yield the first two equalities. The third equality follows from the definition of $s_a(t)$. Following this, the second inequality holds because the total amount of flow on (static) arc a for OD pair k during the entire planning horizon cannot exceed H^k. The sum of the latter is H, a constant that can be factored out of the summation over A. This validates the penultimate equality. Finally, the last equality follows from the fact that $\sum_{a \in A} 1$ simply denotes the number of elements in the set A. Choosing $\delta = \frac{\varepsilon}{H|A|}$ guarantees that $q^T_u Y^l - q^T_l Y^l \leq \epsilon$. When combined with the results in Lemma 2, the latter implies that $q^T_u Y^u - \Phi(Y^*)^T Y^* \leq \epsilon$ and $\Phi(Y^*)^T Y^* - q^T_l Y^l \leq \epsilon$. □

The approximate solution Y^u can be improved by solving an additional nonlinear program. In particular, consider the approximation refinement problem $\min\{\Phi(Y)^T Y : (Y, Z^u) \in S(\delta)\}$, i.e., this is the original problem with $Z = Z^u$. Doing so makes it possible to remove TE arcs corresponding to $z^u_{a(t,s)} = 0$ from the TE network and discard decision variables $x_{a(t)}$ and constraints (5) and (7) from the problem. In DTDTA, we use $x_{a(t)}$, the number of cars on arc a at time t, to compute the travel time on arc a and, subsequently, to select which TE arc to use or which $z_{a(t,s)}$ to set to one. Thus, when Z is given, $x_{a(t)}$ becomes unnecessary. Additionally, let $s(t)$ be such that $z^u_{a(t,s(t))} = 1$ for each $t \in \Delta$. Then, constraint (9), originally (6), reduces to requiring $\sum_{(\tau,s) \in \Omega_{a(t)}} Y_{a(\tau,s)}$ to be in the interval $(s(t) - \delta, s(t)]$. In other words, the original problem with $Z = Z^u$ is a nonlinear multi-commodity flow problem with the latter as side constraints.

Let $\hat{Y}^u$ be an optimal solution to $\min\{\Phi(Y)^T Y : (Y, Z^u) \in S(\delta)\}$. Then, the following corollary shows that $\hat{Y}^u$ better approximates Y^*.

Corollary 1 $\Phi(Y^*)^T Y^* \leq \Phi(\hat{Y}^u)^T \hat{Y}^u \leq \Phi(Y^u)^T Y^u \leq q^T_u Y^u$

Proof In the above sequence of inequalities, the first one follows because Y^* is optimal to the original problem and $\hat{Y}^u$ is only feasible. The second holds because Y^u is feasible to $\min\{\Phi(Y)^T Y : (Y, Z^u) \in S(\delta)\}$. Finally, the last is due to Lemma 2. □

5 Computational results

We conducted numerical experiments using small test networks to empirically verify our understanding of DTDTA as well as to evaluate the efficiency and effectiveness of the approximation schemes discussed in previous sections.

In all problems, the planning horizon is [0, 10) and the travel cost functions are either linear, i.e., $\phi(w) = 1.5 + 2.5(\frac{w}{100})$, or quadratic, i.e. $\phi(w) = 1.5 + 2.5(\frac{w}{100})^2$, where w is the number of cars on the arc. We consider the three different demand patterns displayed in Table 1.

In all three patterns, travel demands at discrete points increases gradually until time 4, levels off briefly, and then decreases gradually after time 5. The individual demands in the three patterns are different and represent three traffic intensities: low, medium, and high. We used GAMS [19] to implement and solve all problems using NEOS Server of Optimization [31]. In particular, we used SBB [39] to solve our nonlinear integer programming problem, i.e., DTDTA, XPress-XP [44] to solve our linear integer programs, i.e., the lower and upper-bound problems, and CONOPT [14] to solve our linearly constrained optimization problems, i.e., the approximation refinement problems. All CPU times reported herein are from the NEOS server.

To empirically verify that DTDTA problem is not convex, we first consider the two-arc network in Fig. 7 that has one OD pair. We let $\delta = 1$. Thus, the discrete-time planning horizon is $\Delta = \{0, 1, \ldots, 9\}$. We use the above quadratic travel time function for both arcs and the function yields travel times in the interval [1.5, 4.0]. Because $\delta = 1$, the set of discrete travel times is $\Gamma = \{2, 3, 4\}$. Using the low traffic intensity demand pattern in Table 1, we solved DTDTA using SBB and terminated it when the relative optimality gap is less than 0.005 (or 0.5%). There are two optimal solutions (see Table 2) to the two-arc problem with an optimal total delay of 450.

Consider the first solution, labeled 'Solution 1', in the Table 2. At time 0, there are 20 cars to travel from nodes 1 to 2. At this time, there are also 20 cars already on arc a_1. These cars enter the arc at time 9 and have not reached their destination at time 0. Because DTDTA assumes that the time to traverse arc a_1 depends on the number of cars on the arc at the entrance time, the travel time for arc a_1 at time 0 is

Table 1 Demand patterns

Traffic intensity	Time										Total
	0	1	2	3	4	5	6	7	8	9	
Low	20	25	30	35	40	40	35	30	25	20	300
Medium	30	35	40	45	50	50	45	40	35	30	400
High	40	45	50	55	60	60	55	50	45	40	500

Fig. 7 Two-arc network
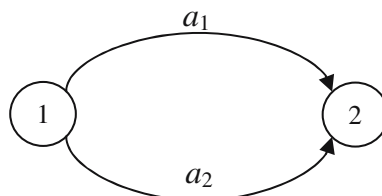
$1.5 + 2.5(\frac{20}{100})^2 = 1.6$. On the other hand, there is no car on a_2 at time 0. Cars that enter the arc at time 8 already left the arc by time 0. Thus, the travel time for a_2 at time 0 is 1.5, the free-flow travel time. To minimize the travel time, all 20 cars entering the network at time 0 must travel on a_2. In fact, every car in Solution 1 travels at the free-flow travel time of 1.5. Thus, there cannot be any solution with less total delay and Solution 1 must be optimal. Because of the symmetry in the data, switching the flows between the two arcs in the network yields Solution 2, another optimal solution. Furthermore, it is easy to verify that every convex combination of these two solutions is feasible to DTDTA and yields, on the other hand, a larger total delay, thereby confirming empirically that the objective function is not convex.

Additionally, the 'extreme' travel behavior displayed in Table 2 may not be intuitive. This is due to the assumption that the system operator is extremely sensitive to the difference in travel times and is willing to switch routes in order to save a minute amount of travel time.

When the network is large, it would be too time-consuming to solve DTDTA optimally or otherwise. In our experiments, we consider four approximate solutions to DTDTA: (Y^l, Z^l), (Y^U, Z^U), $(\hat{Y}^l, Z^l)$, and $(\hat{Y}^U, Z^U)$, where the last two are refinements of the first two. To evaluate the quality and the computation times of these solutions, we consider the four-node network in Fig. 8 with two OD pairs, (1, 4) and (2, 4). In our experiments, both OD pairs have the same demand pattern and all arcs have the same travel cost function, linear or quadratic, as specified above.

First, we solved the lower and upper-bound problems with using two levels of discretization, $\delta = 1$ and $\delta = 0.5$. As before, when $\delta = 1$, the discrete-time planning horizon is $\Delta = \{0, 1, \ldots, 9\}$. On the other hand, when $\delta = 0.5$, Δ becomes $\{0, 0.5, 1, 1.5, \ldots, 9, 9.5\}$. For the comparison below (see Tables 3 and 4), we assume that, when $\delta = 0.5$, there is no demand at fractional times (e.g., at 0.5, 1.5, 2.5, etc.) and the demands at integral times (i.e., 1, 2, 3, etc.) are as shown in Table 1.

For both types of travel cost functions, the size of the optimality gap (i.e., $q_u^T Y^u - q_l^T Y^l$) decreases by approximately 50% as δ decreases from 1 to 0.5. However, the results in Tables 3 and 4 suggest that the reduction in the gap is due mainly to the improvement in the solution, Y^l, of the lower-bound problem. The approximate travel delays as estimated by Y^u change relatively little for the two values of δ.

Table 2 Optimal solutions to the two-arc problem

| Time	Solution 1				Solution 2			
	Inflow		Travel time		Inflow		Travel time	
	a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2
0	0	20	1.600	1.500	20	0	1.500	1.600
1	25	0	1.500	1.600	0	25	1.600	1.500
2	0	30	1.656	1.500	30	0	1.500	1.656
3	35	0	1.500	1.725	0	35	1.725	1.500
4	0	40	1.806	1.500	40	0	1.500	1.806
5	40	0	1.500	1.900	0	40	1.900	1.500
6	0	35	1.900	1.500	35	0	1.500	1.900
7	30	0	1.500	1.806	0	30	1.806	1.500
8	0	25	1.725	1.500	25	0	1.500	1.725
9	20	0	1.500	1.656	0	20	1.656	1.500

Fig. 8 Four-node network
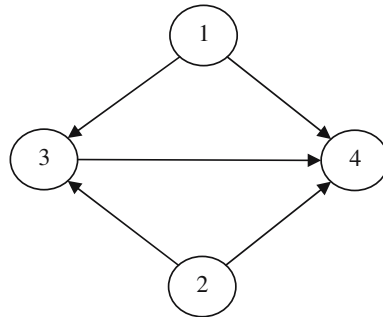
Table 3 Solutions from the lower and upper-bound problems using the linear travel cost function

Traffic intensity	$\delta = 1$			$\delta = 0.5$		
	$q_l^T Y^l$	$q_u^T Y^u$	Gap	$q_l^T Y^l$	$q_u^T Y^u$	Gap
low	820.0	1580.0	760.0	1187.5	1560.0	372.5
medium	1200.0	2230.0	1030.0	1705.0	2230.0	525.0
high	1500.0	2875.0	1375.0	2187.5	2870.0	682.5

Table 4 Solutions from the lower and upper-bound problems using the quadratic travel cost function

Traffic intensity	$\delta = 1$			$\delta = 0.5$		
	$q_l^T Y^l$	$q_u^T Y^u$	Gap	$q_l^T Y^l$	$q_u^T Y^u$	Gap
Low	600.0	1200.0	600.0	900.0	1200.0	300.0
Medium	822.2	1644.5	822.2	1233.3	1644.5	411.1
High	1124.5	2248.9	1124.5	1686.7	2248.9	562.2

Tables 5 and 6 compare the solutions from DTDTA, (Y^*, Z^*), against two approximations, $(\hat{Y}^u, Z^u)$ and $(\hat{Y}^l, Z^l)$. As in the two-node problem, we solve DTDTA using SBB to obtain a (integer) solution (Y^*, Z^*) with less than 0.5% relative optimality gap. To obtain $(\hat{Y}^u, Z^u)$, we first solve the upper-bound problem using XPress-MP to obtain (Y^u, Z^u), a (integer) solution with less than 0.5% optimality gap, and, then, solve the approximation refinement problem (with $Z = Z^u$) using CONOPT to obtain $(\hat{Y}^u, Z^u)$. The solution $(\hat{Y}^l, Z^l)$ are obtained in the same manner. In the two tables, the CPU times for the two approximations are times for solving both bounding and refinement problems.

For both linear and quadratic travel time functions, the two approximation schemes provide solutions with relatively small errors using much less CPU time required to solve DTDTA (see the ratios of the CPU times in Tables 5 and 6). For quadratic travel time functions, the approximate solutions are identical to DTDTA solutions, except for the high traffic intensity case when the approximate solutions are slightly better (by 0.06%).

Table 5 Quality of refined upper and lower-bound solutions: linear travel cost function

Traffic intensity	(Y^*, Z^*)		$(\hat{Y}^u, Z^u)$		$(\hat{Y}^l, Z^l)$		Rel. Err (%)	CPU Ratio	
	Delay	CPU* (s)	Delay	CPU^u (s)	Delay	CPU^l (s)		$\frac{\text{CPU}^*}{\text{CPU}^u}$	$\frac{\text{CPU}^*}{\text{CPU}^l}$
Low	1337.50	27.42	1385.00	2.57	1392.50	2.38	3.55	10.7	11.5
Medium	1800.00	15.92	1866.30	2.66	1815.30	2.90	0.85	6.0	5.5
High	2290.00	95.02	2327.50	4.07	2315.00	1.25	1.09	23.3	76.0

Table 6 Quality of refined upper and lower-bound solutions: quadratic travel cost function.

Traffic intensity	(Y^*, Z^*)		$(\hat{Y}^u, Z^u)$		$(\hat{Y}^l, Z^l)$		Rel. Err (%)	CPU Ratio	
	Delay	CPU* (s)	Delay	CPU^u(s)	Delay	CPU^l(s)		$\frac{\text{CPU}^*}{\text{CPU}^u}$	$\frac{\text{CPU}^*}{\text{CPU}^l}$
Low	1054.50	0.88	1054.50	0.09	1054.50	0.08	0.00	9.8	11.0
Medium	1543.80	6.62	1543.80	0.14	1543.80	0.34	0.00	47.3	19.5
High	2129.80	501.17	2128.60	0.10	2128.60	0.13	−0.06	5011.7	3855.2

6 Conclusion

This paper formulates a DTDTA in which the planning horizon is treated in a circular fashion and events occur periodically. Doing so allows positive flows on the network both at the beginning and at the end of the planning horizon. The structure underlying the formulation is the time-expansion of the (static) network representation of streets and highways. The resulting problem is a nonlinear program with binary variables, a difficult class of problems to solve. Alternatively, two linear integer programs are constructed to obtain approximate solutions and bounds on the total travel delay. It is shown that solutions from the latter can be made arbitrarily close to solutions of DTDTA. Furthermore, numerical results from small test problems suggest that solving linear integer program is more efficient.

References

1. Ahuja, R., Magnanti, T., Orlin, J.: Network Flows. Prentice Hall, Englewood Cliffs, NJ (1993)
2. Boyce, D., Ran, B., LeBlanc, L.: Solving an instantaneous dynamic user-optimal route choice model. Transport. Sci. **29**, 128–142 (1995)
3. Brotcorne, L., De Wolf, D., Gendreau, M., Labbé, M.: A dynamic user equilibrium model for traffic assignment in urban areas. In: Gendreau, M., Marcotte, P. (eds.) Transportation and Network Analysis: Current Trends pp. 49–69. Kluwer Academic Publishers, Norwell, MA (2002)
4. Carey, M.: Optimal time-varying flow on congested network. Oper. Res. **35**(1), 58–69 (1987)
5. Carey, M.: Dynamic traffic assignment with more flexible modelling within links. Netw. Spatial Econ. **1**, 349–375 (2001)

6. Carey, M.: Link travel times I: desirable properties. Netw. Spatial Econ. **4**, 257–268 (2004)
7. Carey, M.: Efficient discretisation for link travel time models. Netw. Spatial Econ. **4**, 269–290 (2004)
8. Carey, M., McCartney, M.: Behaviour of a whole-link travel time model used in dynamic traffic assignment. Transport. Res. B **36**, 83–95 (2002)
9. Carey, M., Ge, Y., McCartney, M.: A whole-link travel-time model with desirable properties. Transport. Sci. **37**(1), 83–96 (2003)
10. Carey, M., Ge, Y.: Convergence of a discretised travel-time model. Transport. Sci. **39**(1), 25–38 (2005)
11. Carey, M., Srinivasan, A.: Solving a class of network models for dynamic flow control. Eur. J. Oper. Res. **75**, 151–170 (1994)
12. Carey, M., Subrahmanian, E.: An approach to modelling time-varing flows on congested networks. Transport. Res. B **34**, 157–183 (2000)
13. Chen, H., Hsueh, C.: A model and an algorithm for the dynamic user-optimal route choice problem. Transport. Res. B **32**(3), 219–234 (1998)
14. CONOPT Solver: ARKI Consulting and Development A/S, Bagsvaerdvej 246 A, DK-2880 Bagsvaerd, Denmark, Phone +45 44 49 03 23, Fax +45 44 49 03 33, E-mail info@arki.dk.
15. Daganzo, C.: Properties of link travel time function under dynamic loading. Transport. Res. B **29**, 95–98 (1995)
16. Drissi-Kaitouni, O., Hameda-Benchekroun, A.: A dynamic traffic assignment model and a solution algorithm. Transport. Sci. **26**(2), 119–128 (1992)
17. Friesz, T., Bernstein, D., Mehta, N., Tobin, R., Ganjalizadeh, S.: Dynamic network traffic assignment considered as a continuous time optimal control problem. Oper. Res. **37**(6), 893–901 (1989)
18. Friesz, T., Bernstein, D., Smith, T., Tobin, R., Wie, B.: A variational inequality formulation of the dynamic network user equilibrium problem. Oper. Res. **41**(1), 178–191 (1993)
19. GAMS Development Corporation, 1217 Potomac Street, NW, Washington, DC 20007, USA, Phone: (202) 342–0180, Fax: (202) 342–0181.
20. Garcia, A., Reaume, D., Smith, R.l.: Fictitious play for finding system optimal routing in dynamic traffic networks. Transport. Res. B **34**, 147–156 (2000)
21. Han, S., Heydecker, B.: Consistent objective and solution of dynamic user equilibrium models. Transport. Res. B, **40**, 16–34 (2006)
22. Ho, J.: A successive linear optimization approach to the dynamic traffic assignment problem. Transport. Sci. **14**, 295–305 (1980)
23. Janson, B.N.: Dynamic traffic assignment for urban road network. Transport. Res. B **25**, 143–161 (1991)
24. Kaufman, D.E., Nonis, J., Smith, R.L.: A mixed integer linear programming model for dynamic route guidance. Transport. Res. B **32**, 431–440 (1998)
25. Lieberman, E.B.: An advance approach to meeting saturated flow requirements. In: Proceeding of the 72nd Annual Transportation Research Board Meeting, Washington, DC (1993)
26. Li, W., Waller, S.T., Ziliaskopoulos, A.K.: A decomposition scheme for system optimal dynamic traffic assignment models. Netw. Spatial Econ. **3**, 441–455 (2003)
27. Lighthill, M., Whitham, G.: On kinematic waves I: flow movement in long rivers and on kinematic waves II: a theory of traffic flow on long crowded roads. Proc. R. Soc. A **229**, 281–345 (1955)
28. Lin, W., Lo, H.: Are the objective and solutions of dynamic user-equilibrium models always consistent? Transport. Res. A **34**, 137–144 (2000)
29. Merchant, D.K., Nemhauser, G.L.: A model and an algorithm for the dynamic traffic assignment problems. Transport. Sci. **12**(3), 183–199 (1978)
30. Merchant, D.K., Nemhauser, G.L.: Optimality conditions for dynamic traffic assigment problem. Transport. Sci. **12**(3), 200–207 (1978)
31. NEOS Server of Optimization, Technical Report 97/04, Optimization Technology Center, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, 1997.
32. Peeta, S., Ziliaskopoulos, A.: Foundations of dynamic traffic assignment: the past, the present and the future. Netw. Spec. Econ. **1**, 233–265 (2001)
33. Perakis, G., Roels, G.: An analytical model for traffic delays and the dynamic user equilibrium models. Techical Report, OR-368-04, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA (2004)
34. Ran, B., Boyce, D., LeBlanc, L.: A new class of instantaneous dynamic user-optimal traffic assignment models. Oper. Res. **41**(1), 192–202 (1993)
35. Ran, B., Boyce, D.: Modeling Dynamic Transportation Networks. Springer, Berlin (1996)

36. Ran, B., Boyce, D.: A link-based variational inequality formulation for ideal dynamic user-optimal route choice problem. Transport. Res. C **4**(1), 1–12 (1996)
37. Ran, B., Hall, R., Boyce, D.: A link-based variational inequality model for dynamic departure time/route choice. Transport. Res. B, **30**(1), 31–46 (1996)
38. Richards, P.: Shock waves on the highway. Oper. Res. **4**, 42–51 (1956)
39. SBB Solver, ARKI Consulting and Development A/S, Bagsvaerdvej 246 A, DK-2880 Bagsvaerd, Denmark, Phone +45 44 49 03 23, Fax +45 44 49 03 33, E-mail info@arki.dk.
40. Smith, M.: A new dynamic traffic model and the existence and calculation of dynamic user equilibria on congested capacity-constrained road network. Transport. Res. B **27**(1), 49–63 (1993)
41. Wie, B., Tobin, R., Friesz, T., Bernstein, D.: A discrete time, nested cost operator approach to the dynamic network user equilibrium problem. Transport. Sci. **29**(1), 79–92 (1995)
42. Wie, B., Tobin, R., Carey, M.: The existence, uniqueness and computation of an arc based dynamic network user equilibrium formulation. Transport. Res. B **36**, 897–928 (2002)
43. Wu, J., Chen, Y., Florian, M.: The continuous dynamic network loading problem: a mathematical fromulation and solutions methods. Transport. Res. B **32**(3), 173–187 (1997)
44. XPress Solver, Dash Optimization Inc, 560 Sylvan Avenue, Englewood Cliffs, NJ 07632, USA, Phone 201 567 9445, Fax 201 567 9443, E-mail info@dashoptimization.com.
45. Zhu, D., Marcotte, P.: On the existence of solutions to the dynamic user equilibrium problem. Transport. Sci. **34**(4), 402–414 (2000)
46. Ziliaskopoulos, A.: A linear programming model for a single destination system optimum dynamic traffic assignment problem. Transport. Sci. **34**(1), 37–49 (2000)