System optimum dynamic traffic assignment with departure time choice on two-terminal networks

Hong Zheng\(^a\), Srinivas Peeta\(^{b,c}\)*

\(^a\) NEXTRANS Center, Purdue University, 3000 Kent Ave., West Lafayette, IN 47906, USA
\(^b\) School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30332, USA
\(^c\) H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Dr NW, Atlanta, GA 30318, USA

* Corresponding author, peeta@gatech.edu, Tel.: +1 404-894-2243; fax: +1 404-894-5418
System optimum dynamic traffic assignment with departure time choice on two-terminal networks

Abstract

This paper addresses the system optimum dynamic traffic assignment (SO-DTA) problem with departure time choice on a two-terminal network, where the travel cost consists of travel time and an early schedule delay cost. Under a certain condition we show that SO-DTA reduces to the latest departure earliest arrival flow (LDEAF), and hence methodologies of LDEAF can be used to study SO-DTA. We show that the optimal departure and arrival schedules are unique, piece-wise linear, and mutually symmetric in the LDEAF. A successive shortest path (SSP) algorithm is used to solve the LDEAF problem. The obtained LDEAF solution provides the simultaneous route and departure time choices in the SO-DTA context. The benefit is that the SSP algorithm involves only shortest path computations on a static network. System marginal costs, externalities and dynamic user equilibrium tolls are analyzed in the LDEAF context, and used to provide a better understanding of the SO-DTA problem characteristics. Further, the study findings can be used for the morning commute problem as it is a special case of the SO-DTA problem on a two-terminal network.

Keywords: dynamic traffic assignment; departure time choice; early schedule delay; system optimum; latest departure earliest arrival flow; morning commute

1. Introduction

This paper addresses the system optimum dynamic traffic assignment (SO-DTA) problem with departure time choice on a two-terminal network, i.e., a network with single origin and single destination. In it, the travel cost consists of the travel time and an early schedule delay cost. In SO-DTA, traffic flow is assigned over time such that the total travel cost of all users in the network is minimized. SO-DTA is often deemed as a useful benchmark for evaluating various transportation policies, including congestion pricing (Yang and Huang, 1997), network access control (Zhang and Shen, 2010), traffic management under evacuation (Hsu and Peeta, 2014), and dynamic parking (Yang, et al. 2013; Ma and Zhang, 2017), etc. In the state-of-the-art, SO-
DTA has been mainly studied in two categories, i.e., with departure time choice (Smeed, 1967; Yang and Meng, 1998; Chow, 2009a; Chow, 2009b; Shen and Zhang, 2009; Doan and Ukkusuri, 2015) and without departure time choice (Carey and Srinivasan, 1993; Ziliaskopoulos, 2000; Munoz and Laval, 2006; Zheng and Chiu, 2011; Carey and Watling, 2012; Shen and Zhang, 2014). From a modeling standpoint, SO-DTA and the congestion tolling schemes have been studied mainly using link-based (Carey and Srinivasan, 1993; Ziliaskopoulos, 2000; Carey and Watling, 2012) and path-based (Shen et al., 2007; Zheng and Chiu, 2011; Qian et al., 2012; Doan and Ukkusuri, 2015) approaches.

In the SO-DTA with departure time choice, commuters are assigned over space and time such that the total system travel time and schedule delay cost are minimized. Schedule delay and departure time choice are essential factors in congestion management during peak periods, as the temporal distribution of traffic is often more critical than the spatial distribution (Arnott et al., 1990). In this domain, the bottleneck model (Vickrey, 1969) has been well-studied in the literature, which describes the congestion mechanism as queuing at bottlenecks. A bottleneck specifies a deterministic capacity $u$ on roadways in the network, and travelers either cross the bottleneck or wait in the queue. The queue is discharged at a constant rate of $u$. Such a bottleneck model has been studied extensively in the literature for time-dependent no-toll equilibrium on either serial or parallel networks (Vickrey, 1969; Hendrickson and Kocur, 1981; Arnott et al., 1990, 1992; Braid, 1996; Ramadurai et al., 2010).

The bottleneck model has been applied for many practical applications in studying SO-DTA. For instance, many transportation researchers study the SO-DTA in a simplified corridor network, where a freeway and a set of arterial routes enable commuting from suburbs to downtown, known as the morning commute problem (Hurdle and Buckley, 1974; Fargier, 1983;
Newell, 1987; Arnott et al., 1990; Shen and Zhang, 2009; Zhang and Shen, 2010; Qian and Zhang, 2013; Yang et al., 2013). In those studies, network topologies have been focused on single route with single bottleneck (Vickrey, 1973; Hendrickson and Kocur, 1981; Fargier, 1983; Smith, 1984; Daganzo, 1985), and parallel routes with multiple bottlenecks (Arnott et al., 1990; Kuwahara, 1990; Arnott et al., 1993). Recently, Munoz and Laval (2006) addressed the graphical method for SO-DTA without departure time choice, and Shen and Zhang (2009) studied a similar problem with multiple bottlenecks and with departure time choice. Later, Zhang and Shen (2010) studied the SO-DTA in a monocentric network – a single destination network without diversion and a unique path between each origin-destination (O-D) pair, which was first studied by Lovell and Daganzo (2000). Ngoduy et al. (2016) investigated SO-DTA which can optimally distribute the queue length over links so that the spillbacks can be minimized or free flow traffic inside the network can be attained. Qian and Zhang (2013) analyzed the user optimal solution for the morning commute problem on a two-terminal network with single route or two routes. Yang et al. (2013) investigated the morning commute problem with both bottleneck congestion and parking space constraints. Ma and Zhang (2017) studied the traffic flow patterns in a single bottleneck corridor with a dynamic ridesharing mode and dynamic parking charges. These morning commute problems typically share a similar network topological structure which consists of a freeway and an abstracted surface street grid, where the grid streets are connected to the freeway through a set of on- and off-ramps. These network-specific topologies are usually restricted to having fixed bottleneck locations either at the end of the freeway, or at on- and off-ramps. Queues at bottlenecks are analyzed by enumerating paths, and departure times are analyzed by manipulating flows along those paths. Hence, nearly all morning commute problems studied in the literature are limited to idealized networks – single corridor where a few
bottlenecks are with known locations – and share similar insights. However, it is difficult to extend these results to networks where the bottleneck locations are unknown.

Yang and Meng (1998) utilized space-time extended network (STEN) to analyze system optimality, externality and congestion toll in SO-DTA with departure time choice with multiple O-D pairs. Later, STEN was also used by Carey and Subrahmanian (2000) and Shen (2008), respectively, to study SO-DTA on a multi-origin single-destination network.

With multiple O-D pairs, SO-DTA is usually solved by an assignment method (e.g., method of successive average (MSA) or gradient projection algorithm) based on computing the path marginal costs (Ghali and Smith, 1995; Peeta and Mahmassani, 1995; Shen et al., 2007; Qian et al., 2012; Doan and Ukkusuri, 2015). In such methods, the dynamic system optimal solution is obtained if the traffic flow is equilibrated with respect to the path marginal costs. However, measuring the path marginal costs is a challenging task because it is difficult to capture traffic flow dynamics and queue spillbacks resulting from multiple bottlenecks and path marginal costs are non-additive and discontinuous in the path-based formulation (Munoz and Laval, 2006; Kuwahara, 2007; Shen et al., 2007; Laval et al., 2015). Hence, mostly heuristics have been used to approximate the path marginal costs.

In this study, we use a different dynamic flow modeling approach (Bookbinder and Sethi, 1980; Aronson, 1989; Powell et al., 1995) rather than STEN to address the SO-DTA problem with departure time choice. Our approach models flows over time on a dynamic network. A dynamic network consists of the same static network, with an additional arc attribute $\pi_{ij}$ measuring a constant traversal time from the tail $i$ of an arc $(i,j)$ to its head $j$. While $\pi_{ij}$ is regarded as constant in a dynamic network, a dynamic flow model specifies both routes and departure schedules of flows entering the network so that flows assigned on arcs vary over time.
Repeatedly sending static flows in different time windows induces dynamic flows over time. The constant arc traversal time in a dynamic network is analogous to a constant flow-independent link travel time in a traffic network. Flows on an arc either travel at a constant speed or wait on a holdover arc (see STEN defined in Section 2.1) that represents waiting in the queue.

In general, a traffic flow model is an essential component of SO-DTA. In this study, we focus on the point queue (P-Q) model. However, the results can be extended to the spatial queue (S-Q) model and the cell-transmission model (CTM) because there are no queues in the network under SO-DTA (see Lemma 2) (Shen and Zhang, 2014; Zheng et al., 2015).

This study shows that when the clearance time (the time horizon such that all traffic reaches the destination) is minimum, the SO-DTA with departure time choice on a two-terminal network reduces to the latest departure earliest arrival flow (LDEAF). Furthermore, we establish the uniqueness in terms of departure schedule at the origin and arrival schedule at the destination in the LDEAF. There are no queues at intermediate nodes in the LDEAF. It is shown that this LDEAF can be solved using the successive shortest path (SSP) algorithm with a run-time that depends on $\log T$, as this algorithm applies static flow techniques to solve for a set of chain flows that induce the LDEAF over time. This SSP-based assignment method is much simpler than using STEN or computing the path marginal costs. As a result of the LDEAF, the cumulative departure and arrival curves in the SO-DTA are shown to exhibit the symmetry property. Additionally, the system marginal costs, externalities and the optimal tolling scheme in the LDEAF solution are analyzed. Using dual variables in the LDEAF, it is shown that there exists a time-dependent link tolling scheme for which the dynamic system optimal (DSO) solution exhibits the dynamic user equilibrium (DUE) property under the tolls. It is noted that path marginal costs, externalities and tolls have been analyzed in previous studies for the SO-DTA,
e.g., (Carey and Srinivasan, 1993; Ziliaskopoulos, 2000; Shen, 2008; Chow, 2009; Carey and Watling, 2012). Different from these studies, our analysis is tailored in the context of LDEAF, rather than SO-DTA directly.

The contributions of the paper are as follows: (1) To the best of our knowledge, this study is the first to show that SO-DTA with departure time choice studied by the traffic flow community reduces to the LDEAF studied by the operations research community under the condition that the clearance time is minimum. This contributes to the theory of SO-DTA. (2) The SSP algorithm, a method used to produce the LDEAF, can be used to determine the simultaneous route and departure time choices. This method has a run time that depends on $\log T$. Unlike Shen (2008), our method does not utilize STEN which has a run-time that depends on $T$, and can thus save computational time. Further, the SSP algorithm can be solved on a large-scale network (Zheng, 2009; Zheng et al., 2014; 2015), as it only involves solving the shortest path problem on a static network and then turning the static flow into dynamic flow over time by repeatedly sending flow in different time windows. (3) The departure schedule at the origin and the arrival schedule at the destination are shown to be mutually symmetric and unique in the LDEAF. (4) The method can be used to study the morning commute problem when the clearance time is minimum. In contrast to most prior morning commute studies, our method does not require restrictions on the number and location of bottlenecks; multiple bottlenecks can be anywhere and can interact with each other as in a general two-terminal network. (5) The path marginal costs, externalities and tolls are analyzed in the LDEAF context, which provides a better understanding of the SO-DTA problem characteristics. All findings are restricted to a two terminal network as the existence of LDEAF is guaranteed only in a two-terminal network. With multiple origins or destinations, the LDEAF may not exist (Fleischer, 2001).
The remainder of the paper is organized as follows. Section 2 provides a preliminary background of the earliest arrival flow and LDEAF problems. Section 3 demonstrates that the SO-DTA with departure time choice on a two terminal network reduces to the LDEAF when clearance time is minimum. Section 4 presents the solution algorithm to solve for the LDEAF. Section 5 discusses the system marginal costs and the dynamic tolling scheme to obtain the DSO solution which is equivalent to the DUE under the tolls. Section 6 discusses numerical analysis performed for a morning commute problem. Section 7 provides some concluding comments.

2. Preliminaries

2.1. Dynamic flows over time

A static network $N = (V, E, u, c)$ consists of a set of nodes $V$ and a set of arcs (or links) $E$. Each arc $(i, j) \in E$ is associated with a positive capacity $u_{ij}$ and cost $c_{ij}$. A subset of nodes $S \subseteq V$ is called terminals; it consists of a set of origins, denoted by $S^+$, and a set of destinations, denoted by $S^-$. The scope of this paper is restricted to a two-terminal network, which consists of the single origin $r$ and the single destination $q$, i.e., $S^+ := \{r\}$ and $S^- := \{q\}$, unless otherwise noted. A static flow defines a function of $f: E \mapsto \mathbb{R}^+$ that maintain the flow conservation condition $\sum_{(i, j) \in E} f_{ij} - \sum_{(j, i) \in E} f_{ji} = 0$ for all $i \in V \setminus (S^+ \cup S^-)$. A static circulation is a static flow that also admits the flow conservation at terminals. $f$ is called feasible if it obeys the capacity constraint $0 \leq f_{ij} \leq u_{ij}$ for all $(i, j) \in E$. The residual network of a static flow $f$, denoted by $N_f$, is an auxiliary network. For each arc $(i, j) \in N$ introduce two arcs $(i, j)$ and $(j, i)$ in $N_f$. Arc $(i, j)$ has capacity $u_{ij} - f_{ij}$ and cost $c_{ij}$; arc $(j, i)$ has capacity $f_{ij}$ and cost $-c_{ij}$. The amount (value) of a static flow $f$ that has been sent from the origin to the destination is denoted by $|f|$.

A dynamic network $\mathcal{N} = (V, E, u, \pi)$ also consists of a set of nodes $V$ and a set of arcs $E$. Each arc $(i, j) \in E$ is associated with a capacity $u_{ij}$ – also called the bottleneck capacity, and a non-negative free flow travel time (or length) $\pi_{ij}$. Given a two-terminal network, if one uses a dummy arc with zero length and infinite capacity to connect $q$ to $r$, the resultant network is called a circulation network. Let interval $[0, T]$ denote a time horizon of interest.
A continuous dynamic flow over time $x_{ij}(\tau)$ is a function that defines the amount of flow per time unit (that is, the rate) that leaves the tail of $(i, j) \in E$ at moment $\tau$, and reaches the head at $\tau + \pi_{ij}$. Assume $x_{ij}(\tau) = 0$ for $\tau < 0$. For a path $P$ which is denoted by a set of ordered arcs, let $\pi(P) := \sum_{(i,j) \in P} \pi_{ij}$ denote the traversal time (or length) of $P$, and $u(P) := \min_{(i,j) \in P} u_{ij}$ denote the capacity of $P$. Given a time horizon $[0, T]$, a flow over time maintains the following flow conservation conditions.

$$
\sum_{(j,i) \in E} \int_0^\theta x_{ji}(\tau - \pi_{ji}) \, d\tau - \sum_{(i,j) \in E} \int_0^\theta x_{ij}(\tau) \, d\tau \geq 0 \quad \forall i \in V \setminus \{S^+ \cup S^\downarrow\}, \quad \forall \theta \in [0, T)
$$

$$
\sum_{(j,i) \in E} \int_0^\theta x_{ji}(\tau - \pi_{ji}) \, d\tau - \sum_{(i,j) \in E} \int_0^\theta x_{ij}(\tau) \, d\tau = 0 \quad \forall i \in V \setminus \{S^+ \cup S^\downarrow\} \quad \forall \theta \in [0, T)
$$

Eq. (1) indicates that the dynamic flow permits holdover at intermediate nodes (i.e., non-terminal nodes). Naturally, vehicles held at an intermediate node imply that a bottleneck exists at that node and queues are built up at the bottleneck without occupying the space.

If holding is not permitted at intermediate nodes, which implies no inside queues, then the flow conservation condition admits (3) instead of (1).

$$
\sum_{(j,i) \in E} \int_0^\theta x_{ji}(\tau - \pi_{ji}) \, d\tau - \sum_{(i,j) \in E} \int_0^\theta x_{ij}(\tau) \, d\tau = 0 \quad \forall i \in V \setminus \{S^+ \cup S^\downarrow\}, \quad \forall \theta \in [0, T)
$$

We state that a dynamic flow $x: = (x_{ij}(\tau))$ is feasible if the capacity constraint (4) holds.

The amount (value) of the dynamic flow $x$, that is, the amount of flow that has reached the destination in $x: = (x_{ij}(\tau))$, is denoted by $|x|$.\n
$$
x_{ij}(\tau) \leq u_{ij} \quad \forall (i, j) \in E, \quad \forall \tau \in [0, T]
$$

The discrete dynamic flow problem can be transformed into a static flow problem on the \textit{STEN}: $\mathcal{N}^\tau = (V^\tau, E^\tau, u, \tau)$ constructed as follows. The node set $V^\tau$ is defined as:
\[ V^\tau := \{i^\tau | i \in \mathcal{N}, \tau = 0,1, ..., T \} \]

The set of arcs \( E^\tau \) is divided into the set of moving arcs \( E^m \) and holdover arcs \( E^h \):

\[ E^m := \{(i^\tau, j^{\tau+\pi}) | (i,j) \in \mathcal{N}, \tau = 0,1, ..., T - \pi_{ij}\} \]

\[ E^h := \{(i^\tau, i^{\tau+1}) | i \in \mathcal{N}, \tau = 0,1, ..., T - 1\} \]

\[ E^\tau := \{E^m \cup E^h\} \]

The traversal time of the holdover arc is 1, and the capacity of the holdover arc is called the holding capacity. P-Q can be modeled in STEN with holdover flow. S-Q requires minor change of STEN, where holdover arcs are associated with links rather than nodes. This can be done by introducing a dummy node for each link. Drissi-Kaitouni and Hameda-Benchekroun (1992) discussed how to build a STEN to model S-Q. When no congestion occurs in the network, CTM can be modeled in a STEN without holdover flow.

2.2. Maximum dynamic flows

The objective of the maximum dynamic flow problem is to send the maximum amount of flow from a source \( r \) to a sink \( q \) by time \( T \). Mathematically, the maximum dynamic flow problem can be stated as:

\[ v_{max}(T) = \max v \]  

Subject to

\[ \sum_{(r,l) \in E} \int_0^T x_{rl}(\tau) d\tau = v \]  

(1), (2), (4).

Note that because Eq. (1) holds in our formulation, it permits holdover flow (or queuing). In fact, the feasibility of holdover flow does not matter in most of the problems discussed in this
study, including the LDEAF discussed later. This is because though the problems permit holdover, the algorithms used in this study do not use holdover flow.

2.3. Earliest arrival flows (EAF)

Given a time horizon \([0, T]\) and a dynamic network \(\mathcal{N}\), a flow over time is called an EAF if the amount of flow that has already reached the sink until time \(\tau\) is maximum simultaneously for all possible time points \(\tau \in [0, T]\). Let \(v_{\text{max}}(\tau)\) be the amount of the maximum dynamic flow that one can send from source \(r\) to sink \(q\) by time \(\tau\); an EAF achieves the value of \(v_{\text{max}}(\tau)\) at any time \(\tau\) such that \(\tau \in [0, T]\). Therefore, an EAF is a multi-dimensional maximum dynamic flow at all possible time moments simultaneously. Let \(y^\tau\) be the arrival (exiting) flow at sink \(q\), i.e., \(y^\tau = \sum_{(j,q) \in E} x_{jq}(\tau)\), and \(B\) the total number of travelers (demand). Then, the EAF problem can be formulated as follows:

\[
\min \int_0^T \tau \cdot y^\tau d\tau
\]

s.t.

\[
\int_0^T y^\tau d\tau = B
\]

(1), (2), (4).

The objective is to minimize the sum of the weighted arrival flows, where the cost of the arrival flow at sink \(q\) is associated with its exiting time – known as the turnstile cost (Hamacher and Tufekci, 1987). Several studies have shown that the objective of the minimum turnstile cost is equivalent to the EAF (Jarvis and Ratliff, 1982; Zheng and Chiu, 2011; Zheng et al., 2015).

Gale (1959) and Philpott (2000) showed, in discrete and continuous time respectively, the existence of an EAF on a two-terminal network. It should be noted that an EAF may not exist on
a network with multiple sinks. Fleischer (2001) presented an instance on a network with two sinks where an EAF does not exist. On a single-sink network, an EAF does exist (Fleischer and Skutella, 2007).

The definition of EAF specified in (7) relies on a given fixed demand $B$. By contrast, the maximum dynamic flow does not rely on the notion of a fixed demand; rather, it relies on a fixed time horizon. To generate a maximum dynamic flow with fixed demand $B$, one needs to first solve the quickest flow problem (Burkard et al., 1993) to obtain the minimum time horizon $T_{\text{min}}$ required to send $B$ to the sink. Thereby, $B$ is the maximum dynamic flow given the horizon $T_{\text{min}}$ (Burkard et al., 1993). One can then solve the EAF problem with demand $B$ and horizon $T_{\text{min}}$. In the solution set, $B$ is guaranteed to be maximum at horizon $T_{\text{min}}$, implying a maximum dynamic flow by definition.

2.4. Latest departure earliest arrival flows (LDEAF)

Another problem closely related to EAF is called the latest departure flow (LDF), which maximizes the amount of flow leaving the origin at $\tau$ and afterward for all $\tau \in [0,T]$ simultaneously. A latest departure earliest arrival flow (LDEAF) is characterized by both latest departure pattern and earliest arrival pattern simultaneously in a maximum dynamic flow. On a two-terminal network, Minieka (1973) proved a symmetry feature between the latest-departure and earliest-arrival schedules through the following theorem.

**Theorem 1** (symmetry feature of LDEAF) (Minieka, 1973): *On a two-terminal network, if the earliest arrival schedule for a maximum dynamic flow consists of $y^\tau$ units arriving at time $\tau$ such that $\tau \in [0,T]$, then the latest departure schedule for a maximum dynamic flow consists of $y^{T-\tau}$ units departing at time $T - \tau$. 
Let $b^r$ be the departure flow, i.e., $b^r = \sum_{(r,j) \in E} x_{rj}(\tau)$; based on Theorem 1, the LDEAF problem can be stated as:

$$\min \int_0^T [(T - \tau) \cdot b^r + \tau \cdot y^r] d\tau$$

(8a)

s.t.

$$\int_0^T b^r d\tau = B$$

(8b)

(1), (2), (4)

In the objective function, the departure flow $b^r$ at source $r$ is associated with cost $T - \tau$; it tells the later the flow departs, the smaller the objective function is. This yields the latest departure flow pattern. The second term is to achieve the earliest arrival flow pattern by minimizing turnstile cost.

3. SO-DTA with departure time choice

3.1 Equivalence of SO-DTA and LDEAF

Departure time choice is an important dimension for SO-DTA. This study considers homogenous travelers, that is, all travelers are assumed to have the same desired arrival time at the destination, denoted by $t^*$. This is true especially for the morning commute problem in which travelers typically have the same work start time. Travel cost consists of two parts: travel time and schedule delay. For simplicity, we consider early schedule delay only, i.e., no late schedule delay is allowed – the same assumption is used in many other DTA studies (Arnott et al., 1993; Shen and Zhang, 2009; Arnott and DePalma, 2011; DePalma and Arnott, 2012). Suppose the starting time is always zero. Then, $t^*$ is the time when the last traveler arrives at the destination, denoted by time horizon $T$. The travel cost of a trip $i$ departing at time $\theta$ is then specified by:
\[
C_i^\theta = \begin{cases} 
\omega_i + \beta(t^* - \theta - \omega_i) & \text{if } t^* \geq \theta + \omega_i \\
+\infty & \text{otherwise} 
\end{cases}
\]  \tag{9}

where \( \omega_i \) specifies the travel time of trip \( i \) departing at time \( \theta \). \( \omega_i \) is composed of two components, free flow travel time and waiting time in the queue. The cost of one unit of travel time is 1 and the cost of arriving one unit of time early is \( \beta \). From empirical data (Small, 1982), the cost of travel time is more than the cost of early schedule delay; so \( \beta < 1 \).

Note that this study considers continuous time in the SO-DTA analysis in general. In this section, however, we use discrete time for convenience.

The objective function of SO-DTA is:

\[
\min \left( \sum_{\tau=0}^{T} \tau \cdot y^\tau - \sum_{\tau=0}^{T} \tau \cdot b^\tau \right) + \beta \cdot \sum_{\tau=0}^{T} (t^* - \tau) y^\tau \tag{10}
\]

or

\[
\min \left( \sum_{\tau=0}^{T} \sum_{(i,j) \in E} \pi_{ij} \cdot x_{ij}(\tau) \right) + \beta \cdot \sum_{\tau=0}^{T} (t^* - \tau) y^\tau \tag{11}
\]

The term \( \sum_{\tau=0}^{T} \tau \cdot y^\tau - \sum_{\tau=0}^{T} \tau \cdot b^\tau \) in Eq. (10), or \( \sum_{\tau=0}^{T} \sum_{(i,j) \in A} \pi_{ij} \cdot x_{ij}(\tau) \) in Eq. (11), represent the travel time cost, and the term \( \beta \cdot \sum_{\tau=0}^{T} (t^* - \tau) y^\tau \) represents the schedule delay cost.

Next, we show the equivalence of SO-DTA and LDEAF when the clearance time is minimum.

**Theorem 2** (Minieka 1973): If there exists a maximal dynamic flow \( f_1 \) with \( y^\tau \) units leaving the source at each time \( \tau \), and if there exists a maximal dynamic flow \( f_2 \) with \( b^\tau \) units entering the sink at each time \( \tau \), then there exists a maximal dynamic flow \( f_3 \) with both \( y^\tau \) units leaving the source and \( b^\tau \) units entering the sink.

**Proof.** See Theorem 1 in Minieka (1973).
Lemma 1: Let $D(\tau)$ and $A(\tau)$ be the cumulative departure and arrival curves of the LDEAF $f$, respectively. Let $D'(\tau)$ and $A'(\tau)$ be the cumulative departure and arrival curves of any dynamic flow $f'$. Let $T$ and $T'$ be the clearance times of $f$ and $f'$, respectively. When $T'$ is minimum, $D(\tau) \leq D'(\tau)$ for all $0 \leq \tau \leq T'$.

Proof. We first show that $T' = T$ when $T'$ is minimum. $T'$ cannot be more than $T$ as it violates our assumption that $T'$ is minimum. Suppose $T' < T$, it then implies that at least at time $T'$, $A'(T') > A(T')$, because $A'(T')$ equals to the total demand and $A(T')$ is less than the total demand. It then violates that $f$ is the earliest arrival flow because $A(T')$ is not maximum at $T'$. Hence, we have $T' = T$. Since $T'$ is minimum, it indicates that both $f$ and $f'$ are maximal dynamic flows.

Suppose there is $D(\tau) > D'(\tau)$ at time $\tau$, then we can produce a maximal dynamic flow $f''$ with departure pattern of $D'(\tau)$ and arrival pattern of $A(\tau)$ due to Theorem 2. Comparing the departure patterns of $f$ and $f''$, it indicates that $f$ is not the latest departure at time $\tau$; a contradiction. □

Theorem 3: When the clearance time is minimum, SO-DTA reduces to LDEAF.

Proof. LDEAF $\Rightarrow$ SO-DTA

Let $(b^T)^*$ and $(y^T)^*$ be the departure and arrival flows for the LDEAF, and $b^\tau$ and $y^\tau$ be the departure and arrival flows for SO-DTA. Since $(y^T)^*$ is the earliest arrival, we have $\sum_{\tau=0}^{T}(y^\tau)^* \geq \sum_{\tau=0}^{T} y^\tau$ for $\theta = 0,1, ..., T$. Performing a summation for $\theta = 0,1, ..., T$, we have:

\[
\sum_{\tau=0}^{T} (T + 1 - \tau)(y^\tau)^* \geq \sum_{\tau=0}^{T} (T + 1 - \tau) y^\tau
\] (12)

Because $\sum_{\tau=0}^{T}(y^\tau)^* = \sum_{\tau=0}^{T} y^\tau = B$ ($B$ is the total demand), we have:

\[
\sum_{\tau=0}^{T} \tau \cdot (y^\tau)^* \leq \sum_{\tau=0}^{T} \tau \cdot (y^\tau)
\] (13)
\[(1 - \beta) \sum_{\tau=0}^{T} \tau \cdot (y^\tau)^* \leq (1 - \beta) \sum_{\tau=0}^{T} \tau \cdot (y^\tau) \]  \hspace{1cm} (14)

Since \((b^\tau)^*\) is the latest departure flow, there is \(\sum_{\tau=0}^{\theta} (b^\tau)^* \leq \sum_{\tau=0}^{\theta} b^\tau\) for \(\theta = 0,1,...,T\).

Performing a summation for \(\theta = 0,1,...,T\), we have

\[
\sum_{\tau=0}^{T} (T + 1 - \tau) (b^\tau)^* \leq \sum_{\tau=0}^{T} (T + 1 - \tau) b^\tau \]  \hspace{1cm} (15)

Because \(\sum_{\tau=0}^{T} (b^\tau)^* = \sum_{\tau=0}^{T} b^\tau = B\), there is

\[
\sum_{\tau=0}^{T} -\tau \cdot (b^\tau)^* \leq \sum_{\tau=0}^{T} -\tau \cdot (b^\tau) \]  \hspace{1cm} (16)

Because

\[
t^* \cdot \beta \cdot \sum_{\tau=0}^{T} (y^\tau)^* = t^* \cdot \beta \cdot \sum_{\tau=0}^{T} y^\tau = t^* \beta B \]  \hspace{1cm} (17)

Combining (14), (16) and (19) we have

\[
\left( \sum_{\tau=0}^{T} \tau \cdot (y^\tau)^* - \sum_{\tau=0}^{T} \tau \cdot (b^\tau)^* \right) + \beta \cdot \sum_{\tau=0}^{T} (t^* - \tau) (y^\tau)^* \\
\leq \left( \sum_{\tau=0}^{T} \tau \cdot y^\tau - \sum_{\tau=0}^{T} \tau \cdot b^\tau \right) + \beta \cdot \sum_{\tau=0}^{T} (t^* - \tau) y^\tau \]  \hspace{1cm} (18)

Eq. (18) indicates that the total system cost of LDEAF is minimum; therefore it is SO-DTA.

SO-DTA ⇒ LDEAF

Figure 1 shows the cumulative arrival and departure curves for SO-DTA (dash line) and LDEAF (solid line), respectively. The total travel cost consists of two components, where TT refers to the travel time and ED refers to the early schedule delay. The departure curve of SO-DTA must be higher than that of LDEAF due to Lemma 1. The arrival curve of SO-DTA must
be lower than that of LDEAF as otherwise it violates that the arrival curve of LDEAF is the earliest arrival. The total cost of the SO-DTA is $C = TT + \beta ED$.

The upper area (TT) is associated with a weight (equal to 1) higher than the weight (equal to $\beta$) associated with the lower area (ED). The total weighted area is minimum only if the upper area is the smallest possible, that is, at any time point the cumulative departure curve is the minimum possible, and at any time point the cumulative arrival flow is the maximum possible, which is the LDEAF by its definition. □

![Diagram showing cumulative flow (N) vs time (t)](image)

Figure 1. Plots of travel time and schedule delay for proof of Theorem 2.

### 3.2 A special case

In the above, we show that SO-DTA reduces to LDEAF if the clearance time is minimum. An interesting question is the condition under which the minimum clearance time can be achieved. In fact, the clearance time in a SO-DTA is related to the value of $\beta$, because $\beta$ represents the
tradeoff between the travel time and schedule delay components in the total travel cost. To achieve a minimum clearance time, it is desirable for $\beta$ to be large, otherwise if $\beta$ is rather small, travelers tend to depart evenly along the departure window to travel on the uncongested road to reduce travel time since schedule delay is negligible. This travel pattern is different from the LDEAF which requires vehicles to depart as late and arrival as early as possible.

We observe that there exists a threshold such that when $\beta$ is more than the threshold the clearance time is minimum, and thus the equivalence of SO-DTA and LDEAF holds. Therefore, one condition to achieve the minimum clearance time is $\beta$ approaching 1 from below, i.e., $\beta \uparrow 1$.

To show that the SO-DTA is equivalent to LDEAF when $\beta \uparrow 1$, we first show that the departure curves of SO-DTA and LDEAF are identical and correspond to the latest departure. When $\beta \uparrow 1$, Eq. (10) implies that the total cost $C \approx t^* \cdot \sum_{r=0}^{T} y^r - \sum_{r=0}^{T} \tau \cdot b^r = t^* \cdot B - \sum_{r=0}^{T} \tau \cdot b^r$. The total cost $C$ equals the shaded area in Figure 2. Clearly, the shaded area is minimum only if at any time point the cumulative departure flow is the minimum possible, that is, the departure curve is lowest possible, which is a LDF.

Since the departure curves in SO-DTA and LDEAF are identical, we can repeat the proof in Theorem 3 to show SO-DTA is equivalent to LDEAF.

When the clearance time is minimum, which implies $\beta$ and the schedule delay component must be significant, the equivalence between SO-DTA and LDEAF could be understood as follows. SO-DTA tends to depart flows as late as possible, which leads to the latest departure schedule. If there are travelers who depart later than the latest departure schedule, then it is impossible for all travelers to arrive no later than $t^*$, regardless of the routing. The latest departure schedule governs the departure time choice of SO-DTA, and the earliest arrival schedule governs the routing of SO-DTA. Travelers must travel on the shortest paths, and
hence arrive earlier to reduce en route travel time though it may increase schedule delay. Since the weight of schedule delay is less than that of travel time, traveling on the shortest path reduces the total system cost. Therefore, SO-DTA requires that the latest departure and earliest arrival schedules must hold simultaneously, which is the LDEAF. The solution of SO-DTA determines the departure time choice and the corresponding routing in the LDEAF.

3.3 Symmetry feature of SO-DTA

**Corollary 1**: On a two-terminal network, when clearance time is minimum, the arrival curve of SO-DTA is unique, piece-wise linear and convex.

**Proof.** As the arrival flow at any time point is the maximum possible, the curve is unique. Let \( v_{\text{max}}(\theta) \) be the maximal dynamic can be sent within time \( \theta \). Due to the algorithm by Ford and Fulkerson (1958), \( v_{\text{max}}(\theta) \) can be obtained by one static min-cost flow computation. Create a circulation network \( N' \) by connecting \( q \) to \( r \) using an arc with traversal time \( \tau_{qr} = -\theta \) and infinite capacity. Interpreting the traversal time as the cost variable, and solve the min-cost

![Figure 2. Plot of total cost for proof of Theorem 2](image-url)
circulation flow \( f_c = f_{ij} \). The value \( v_{\text{max}}(\theta) \) equals the negative cost of \( f_c \) (c.f. Ford and Fulkerson’s algorithm in Section 4.2). It indicates that the function \( \theta \mapsto v_{\text{max}}(\theta) \) is the negative cost of a parametric min-cost flow problem, which is known to be piecewise linear and convex. □

Due to the symmetry feature demonstrated in Theorem 1, one can construct the LDEAF which admits the following equation.

\[
D(t^* - \tau) = N - A(\tau)
\]

where \( N \) denotes the total number of trips.

Eq. (19) indicates that the departure curve \( D(\tau) \) at the origin and arrival curve \( A(\tau) \) at the destination are symmetric in a LDEAF at the central point \( \left( \frac{t^*}{2}, \frac{N}{2} \right) \), as illustrated in Figure 3. The following proposition is an immediate result of Eq. (19).

**Corollary 2:** On a two-terminal network, when clearance time is minimum, the departure curve of SO-DTA is unique, piece-wise linear and concave.

**Proof.** From the symmetry feature in Theorem 1 and Eq. (19), the departure and arrival curves are symmetric at the point \( \left( \frac{t^*}{2}, \frac{N}{2} \right) \) in the LDEAF. As the arrival curve is piecewise linear and convex due to Corollary 1, the departure curve is also piecewise linear but concave. □

**4. Solution algorithm to solve the LDEAF**

Although the proposed SO-DTA and LDEAF models are link-based formulations, we use a path-based solution algorithm to solve for the LDEAF – the SSP algorithm. A key idea of this algorithm is Ford and Fulkerson’s temporally repeated flow (Ford and Fulkerson, 1958). The concept of temporally repeated flows is first introduced, followed by a description of how temporally repeated flows can be used to solve the maximum dynamic flow problem. Then, the
concept of chain flows is introduced and the SSP algorithm that uses chain flows to solve for the LDEAF is presented.

![Cumulative flow](image)

Figure 3. The latest departure curve and earliest arrival curve are symmetric at central point \((t^*/2, N/2)\).

### 4.1. Temporally repeated flows

Given a static flow \(f = (f_{ij})\) from \(r\) to \(q\), decompose \(f\) into \(f_P\) on paths \(P \in \mathcal{P}\) such that \(f_{ij} = \sum_{P \in \mathcal{P}} \sum_{(i,j) \in P} f_P\). Send flows at constant rate \(f_P\) along paths \(P \in \mathcal{P}\), starting from time zero until \(T - \pi(P)\) which is the latest time from which \(f_P\) is able to reach sink \(t\) by \(T\). The dynamic flow constructed (induced) in such a manner is called a *temporally repeated flow*. The value of the temporally repeated flow can be computed by:

\[
\sum_{P \in \mathcal{P}} (T - \pi(P))f_P = T|f| - \sum_{P \in \mathcal{P}} \pi_P f_P = T|f| - \sum_{(i,j) \in A} \pi_{ij} f_{ij}
\]  

(20)

### 4.2. Maximum dynamic flows

Ford and Fulkerson (1958) showed that there is always a maximum dynamic flow featured in the
class of temporally repeated flows. Maximizing the value of a temporally repeated flow is equivalent to minimizing \( \sum_{(i,j) \in P} \pi_{ij} f_{ij} - T\|f\| \), according to Eq.(20). This implies finding a minimum cost circulation flow \( f \) on a circulated network, where traversal times are interpreted as costs, and the sink \( q \) is connected to the source \( r \) by a dummy arc with cost \(-T\) and infinite capacity. The maximum value of the temporally repeated flow is equal to the negative cost of min-cost circulation flow on the circulated network. Then, the min-cost circulation flow induces the temporally repeated flow to produce the maximum dynamic flow. This method is due to Ford and Fulkerson (1958).

The temporally repeated flow concept illustrates a simple way to turn static flows to dynamic flows. Hoppe and Tardos (2000) generalized this concept to chain flows.

### 4.3. Chain flows

A chain flow \( \gamma = (|\gamma|, P) \) is a static flow of value \(|\gamma| \geq 0\) along path (or cycle) \( P \) which starts from source \( r \) and ends in sink \( q \). The traversal time/length of the chain flow \( \pi(\gamma) \) is the traversal time/length of path \( P \). Given a time \( T \) no less than \( \pi(\gamma) \), a chain flow \( \gamma \) induces a dynamic flow \( [\gamma]^T \) by sending flows at the same rate \(|\gamma|\) along \( P \) at every moment \( \tau \in [0, T - \pi(\gamma)] \). The value of \( [\gamma]^T \) is computed as:

\[
| [\gamma]^T | = \int_0^{T - \pi(\gamma)} |\gamma| = (T - \pi(\gamma)) \cdot |\gamma|
\]

Let \( \Gamma = \{\gamma_1, \gamma_2, ..., \gamma_n\} \) be a set of chain flows, which is a decomposition of static flow \( f \) if \( \sum_{i=1}^{n} \gamma_i = f \). The value of \( \Gamma \) is denoted by \( |\Gamma| := \sum_{\gamma_i \in \Gamma} |\gamma_i| \). Given a time \( T \), if every chain flow in \( \Gamma \) has a length no more than \( T \), \( \Gamma \) induces a dynamic flow by summing up each dynamic flow induced by \( \gamma_i \in \Gamma \). The dynamic flow induced in this manner is denoted by \( [\Gamma]^T \). The value
of $[Γ]^T$ is:

$$|[Γ]^T| = \sum_{y_i \in Γ} |[y_i]^T| = \sum_{y_i \in Γ} (T - \pi(y_i)) \cdot |y_i| \quad \text{(22)}$$

### 4.4. Successive shortest path algorithm solving the LDEAF

Consider solving the min-cost flow (MCF) problem from $r$ to $q$ on the static network. Interpret the traversal time of an arc as the cost, and solve the MCF using the successive shortest path (SSP) algorithm. First, determine the chain flow $γ_1$ with the shortest path length; send $|γ_1|$ units of flow along its path $P_1$; then update the residual network. Second, determine the second chain flow $γ_2$ with the shortest path length in the residual network; send $|γ_2|$ units of flow along its path $P_2$. Repeat this procedure until there is no chain flow whose length is no more than the time span $T$. Finally, induce a dynamic flow $[Γ]^T$ by summing up each dynamic flow induced by $γ_i \in Γ$. The dynamic flow $[Γ]^T$ is a LDEAF. This SSP algorithm to solve the LDEAF problem is due to Minieka (1973) and Wilkinson (1971).

---

**Algorithm 1:** SSP-based Algorithm for the LDEAF

1. **Quickest flow module:** determine the minimum horizon $[0,T]$ that is feasible to transship supply/demand (Burkard et al., 1993)
   
   $Γ \leftarrow \emptyset, γ \leftarrow \emptyset$

2. **While** $τ(γ) < T$, **do**

   Select a flow $γ$ with the minimum length $π(γ)$ on the residual network $N_f$
   
   Augment maximum amount of flow $|γ|$ along $γ$, and update residual network $N_f$

   $Γ \leftarrow Γ \cup γ$

---

23
3 Return $[\Gamma]^T$.

While the SSP algorithm runs in a sufficiently efficient manner in practice, it is not polynomial as the number of paths involved in Step 2 could be exponential (Zadeh, 1973). To the best of our knowledge, no known polynomial algorithm is available to solve the LDEAF problem on a general two-terminal network in the current practice.

The following lemma indicates that in any optimal solution for SO-DTA, there must be no queues at intermediate nodes; otherwise such a flow must be suboptimal.

**Lemma 2:** In an optimal solution of SO-DTA with departure time choice when the clearance time is minimum, there are no queues at intermediate nodes.

**Proof.** We can produce an optimal LDEAF (or SO-DTA) solution using Algorithm 1, with no queues at intermediate nodes in this solution. So, there exists an optimal solution without queues at intermediate nodes. Next, we show that it is impossible to have a LDEAF solution with queues at intermediate nodes. Let $f$ be the LDEAF produced by Algorithm 1, and let $f'$ be another LDEAF which assigns positive flow on holdover arcs. Constructing $f$ and $f'$ on the STEN $\mathcal{N}^\tau$ discussed in Section 2.1, the difference of $f$ and $f'$ can be decomposed into a set of cycles. In another word, we can produce $f'$ by sending flows on a set of cycles in the residual network with respect to $f$. In the path-based flow decomposition, these operations are equivalent to shift flows from a set of paths to another set of paths which involves flow holdover. Consider flows are shifted from a time-space path $p$ to another time-space path $p'$ involving holdover on $\mathcal{N}^\tau$. Let $\varphi$ and $\varphi'$ be the paths in the original network $\mathcal{N}$ corresponding to $p$ and $p'$ respectively. Since $f$ is produced by Algorithm 1, there is no holdover on $p$. However, $p'$ involve holdover for at least one time step. Thus, the free flow travel time of path $\varphi'$ is strictly shorter than $\varphi$. According to Step 2 in the SSP algorithm, such a shorter path $\varphi'$ must be saturated in $f$. Hence, it is not
possible to shift flows to a path involving holdover in the LDEAF. This completes the proof. □

5. System marginal costs, externalities and tolls in the LDEAF

This section analyzes path marginal costs, externalities and tolls in the LDEAF. We show that the dual variable associated with the capacity constraint is the link congestion externality. It is shown that time-dependent link-based system optimal tolls exist that completely eliminate congestion within the network, and the LDEAF solution with the extra toll costs exhibits temporal equilibrium. We note here that prior studies in this research line draw a similar conclusion. Specifically, Hendrickson and Kocur (1981) investigated both DUE and DSO with deterministic queuing model on a single OD network with a single route and single bottleneck, with both early and late schedule delays allowed. They showed a tolling strategy in DSO in which the average cost including tolls turns out to be equal to that of DUE and there is no queue developed in such a case. Later, Heydecker and Addison (2005) extended this model to multiple routes and time-specific cost function, where they also showed that the travel cost plus time-dependent toll is equal to the delays that would be incurred in queues in the network in an untolled equilibrium. Arnott et al. (1990) showed how to develop a toll strategy in dynamic system optimum assignment on a single OD network with parallel routes with deterministic queuing model. They showed that the toll eliminates queues in the network and the toll plus travel cost is equal to the untolled equilibrium. Externalities, optimal DSO tolls, and marginal costs are systematically investigated in Carey and Srinivasan (1993) and Carey and Watling (2012), where they showed that if tolls computed from the DSO solution are imposed on users, then the DSO solution plus toll would also satisfy the criteria for a DUE in a single-destination network with approximated CTM as traffic flow model. The system marginal cost, externalities and optimal tolls analysis are similar to those in this paper, whereas we for the first time seek to
derive tolls and verify those theoretical results in the LDEAF context. Compared to path
marginal cost analyses in previous SO-DTA studies (Carey and Srinivasan, 1993; Shen, 2008;
Chow, 2009; Carey and Watling, 2012), our analysis is tailored to the LDEAF on a two-terminal
network, and we present a different link-based tolling scheme. The LDEAF can be formulated as
follows. Since there are no queues at intermediate nodes, the constraint set consists of the flow
mass balance constraints in Eq.(24a) and Eq.(24b), and the capacity constraint in Eq.(24c).

\[
\begin{align*}
\min & \sum_{\tau=0}^{T} (T - \tau) \cdot b^\tau + \sum_{\tau=0}^{T} \tau \cdot y^\tau \\
\text{s.t.} & \\
(\mu) & \quad \sum_{\tau=0}^{T} b^\tau = B \quad (24a) \\
(a^\theta_i) & \quad - \sum_{(j,l) \in E} x_{ij}(\theta - \pi_{jl}) + \sum_{(i,l) \in E} x_{lj}(\theta) = 0 \quad \forall i \in V \setminus \{r, q\}, \forall \theta \quad (24b) \\
(\lambda^\theta_{ij} \geq 0) & \quad x_{ij}(\theta) \leq u_{ij} \quad \forall (i,j) \in E, \theta \in [0, T] \quad (24c)
\end{align*}
\]

Let \( \mu \) (unrestricted), \( a^\theta_i \) (unrestricted) and \( \lambda^\theta_{ij} \geq 0 \) be the dual variables associated with
constraints (24a)-(24c), respectively. The Karush-Kuhn-Tucker (KKT) conditions are the
complementarity conditions for the pairs of inequalities in (25), plus the feasibility constraints
(24a) – (24c).

\[
\begin{align*}
b^\tau & \geq 0 \quad \mu - a^\theta_j + \pi_{ij} - \lambda^\theta_{rj} \leq T - \theta_1 \quad \forall \theta, \forall (r,j) \in E \quad (25a) \\
y^\tau & \geq 0 \quad a^\theta_j - \lambda^\theta_{jq} \leq \theta_2 \quad \forall \theta = 0, 1, \ldots, T, \forall (j,q) \in E \quad (25b) \\
x_{ij}(\theta) & \geq 0 \quad a^\theta_i - a^\theta_j + \pi_{ij} - \lambda^\theta_{ij} \leq 0 \quad \forall \theta, \forall (i,j) \in E, i \neq r, j \neq q \quad (25c)
\end{align*}
\]
where $\theta_1$ denotes the departure time leaving source $r$, and $\theta_2$ represents the arrival time reaching destination $q$.

The system marginal cost (SMC) is defined as the Lagrangian multiplier or dual variable associated with the constraints. The SMC measures how the optimal value of the objective function will change with the value of the constraint. We interpret $\mu$ as the SMC of increasing the demand $B$. $\mu$ measures how the objective function value changes if there is a small perturbation of the demand from the source. $\alpha_i^\theta$ can be interpreted as the SMC of traveling from node $i$ at time $\theta$ to the destination. $\lambda_{ij}^\theta$ can be interpreted as the SMC incurred by the capacity constraint for link $(i,j)$ at entrance time $\theta$.

**Lemma 3.** Let $\mathbb{P}(i, \theta)$ be the set of time-space paths from node $i$ departing at time $\theta$ reaching the sink. In the LDEAF, the SMC of all utilized paths in $\mathbb{P}(i, \theta)$ is equal to $\alpha_i^\theta$, and less than or equal to the SMC of any unutilized path in $\mathbb{P}(i, \theta)$.

**Proof.** Along any utilized path $P \in \mathbb{P}(i, \theta)$, there is $x_{ij}(\tau) > 0$; hence, by the complementarity conditions Eqs.(25a) – (25c) satisfy equality strictly. In such a case, the equality hold for each link along the utilized time-space path $P$, and by substitution there is $\alpha_i^\theta = \sum_{(l,j) \in P} \lambda_{ij}^{\theta'} + \theta_2$, where $\theta'$ denotes the entrance time of link $(i,j)$ in $P$ and $\theta_2$ denotes the arrival time reaching destination $q$. Because $\alpha_i^\theta$ is independent of the path, the SMC is $\sum_{(l,j) \in P} \lambda_{ij}^{\theta'} + \theta_2$ and is the same for all paths utilized in $\mathbb{P}$.

In an unutilized path $P' \in \mathbb{P}$, $x_{ij}(\theta) = 0$ for at least one link; otherwise the path will be utilized. According to the complementarity conditions, at least one equation in (25a) – (25c) satisfies the inequality; hence, there is $\alpha_i^\theta \leq \sum_{(l,j) \in P} \lambda_{ij}^{\theta'} + \theta_2$, which implies $\alpha_i^\theta$ is less than or equal to the SMC of the unutilized path in $\mathbb{P}$. □
In summary, the SMC of a utilized path \( P \) in \( \mathbb{P} \) (also called the path marginal cost) has:

\[
\alpha_i^0 = \sum_{(i,j) \in P} \lambda_{ij}^0 + \theta_2
\]  

(26)

Next, we analyze the value of \( \mu \). \( \mu \) represents the SMC due to one more traveler departing at source \( p \), which is equal to the extra travel time in the network plus the extra schedule delay experienced by all travelers. From equations (25), we sum up the links traversed by the path of a flow \( \gamma = (|\gamma|, P) \), and then have:

\[
\mu = \sum_{(i,j) \in P} \lambda_{ij}^0 + t^* + \theta_2 - \theta_1
\]  

(27)

\( \theta_1 \) is the departure time of \( \gamma \) and \( \theta_2 \) is the arrival time of \( \gamma \); therefore \( \theta_2 - \theta_1 \) represents the travel time of \( \gamma \). On a congested traffic network, an additional user can cause an increase in travel times for other users and this increase of travel time is referred to as the congestion externality. It is well known that the SMC is equal to the travel cost experienced by the additional user plus the congestion externality caused by the additional user. Thus, the externality equals \( \mu \) minus the path travel cost, that is

\[
\mu - \left( \sum_{(i,j) \in P} \pi_{ij} + \beta(t^* - \theta_2) \right) = \sum_{(i,j) \in P} \lambda_{ij}^0 + t^* - \beta(t^* - \theta_2)
\]  

(28)

If a toll exactly equal to the externality is imposed on each user, then the total generalized cost experienced by each user is the same, and equal to the SMC. Then, the DSO solution is also the DUE solution under the tolls.

**Lemma 4.** Define the link-based time-dependent tolls as follows:

1. For each link such that \( x_{ij}(\theta) > 0 \), let \( \text{toll}_{ij}^\theta = \lambda_{ij}^0 \); otherwise \( x_{ij}(\theta) = 0 \), let \( \text{toll}_{ij}^\theta = 0 \).
2. At the destination node \( q \), at each time \( \tau \), impose a time-dependent toll \( \beta \tau \).

Then, users traveling in the LDEAF exhibit the DUE pattern.
Proof. The generalized user cost in the LDEAF for each utilized path $P$ is equal to the travel cost plus the toll, which is:

$$\sum_{(i,j) \in P} \pi_{ij} + \beta (t^* - \theta_2) + \sum_{(i,j) \in P} \lambda_{ij}^{\theta_1} + \beta \theta_2 = \sum_{(i,j) \in P} \lambda_{ij}^{\theta} + \theta_2 - \theta_1 + \beta t^* = \mu - (1 - \beta) t^*$$

Thus, the generalized cost for each utilized path is a constant number $\mu - (1 - \beta) t^*$, and for each unutilized path is more than this number. Hence, the LDEAF exhibits the DUE pattern under the toll. □

In Lemma 4 we propose a link-based toll, mainly because this pricing scheme is convenient to calculate and easy to understand due to Eq. (28). Note that there can be multiple equilibrium tolling schemes. For example, at the origin $r$ or destination $q$ we can add (or reduce) toll by a constant number, which would affect the generalized cost by a constant value for all paths and all departure times. Such a toll scheme is still an equilibrium toll.

6. Numerical example

6.1. Example and flow solution

In this section, we consider the morning commute network (illustrated in Figure 4) which is a special case of the two-terminal network. The morning commute network considers a corridor consisting of a freeway with a set of on- and off-ramps. A single destination is at the end of the freeway. A surface street grid connects through on- and off- ramps. Each arc is associated with two terms, free flow travel time and capacity. The capacity of the freeway is 200, with a bottleneck of capacity 50 near the destination. The capacity of an on-ramp or off-ramp is 60, and the capacity a surface street is 500. There are 5,000 travelers traveling from the source $r$ to the sink $q$. The desired work start time (arrival time) is 9:00AM. In this example, we specify $\beta = 0.8$. We numerically verified that in this example $\beta = 0.8$ is large enough that the SO-DTA
equals the LDEAF. The system optimal flow pattern and dynamic tolls discussed above are numerically verified in this morning commute problem.

The minimum time span (Burkard et al., 1993) to send 5,000 travelers from \( r \) to \( q \) in this network is 55.62. The LDEAF is then solved using the SSP algorithm as presented in Section 4 (Algorithm 1). We coded Algorithm 1 in Python, and it takes less than 1 second to solve for the optimal solution. The chain flows and their repeating time windows are tabulated in Table 1. Table 1 shows that 5 chain flows have been identified. The first chain flow is along path \( \gamma_1 = \{r, 1, 4, 5, 6, 7, 8, 9, q\} \); the length of \( \gamma_1 \) is 23 and the value of \( \gamma_1 \) is 50. Flows are sent at a rate of 50 units per time during the time window [0,32.62]. These flows reach the sink node \( q \) during the time window [23,55.62] at a rate of 50. The other chain flows are similarly characterized. Then, the LDEAF in Figure 5 can be plotted, which clearly exhibits the symmetry property. It is also numerically verified that the arrival curve is piece-wise linear and convex, while the departure curve is piece-wise linear and concave. In this example, there are four bottlenecks: on-ramps (1,4) and (2,6), off-ramp (9,12), and arc (9, q) at the end of freeway.

Table 1. LDEAF solution.

| Route          | \( \gamma_i \) | \( \tau(\gamma) \) | \( |\gamma| \) | Repeating time window |
|----------------|----------------|--------------------|--------------|----------------------|
| Freeway        |                |                    |              |                      |
| On-ramp or off-ramp |            |                    |              |                      |
| Surface street  |                |                    |              |                      |

Figure 4. Morning commute network.
<table>
<thead>
<tr>
<th></th>
<th>Set</th>
<th>Departure</th>
<th>Arrival</th>
<th>Time Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>{r, 1, 4, 5, 6, 7, 8, 9, q}</td>
<td>23</td>
<td>50</td>
<td>[0, 32.62]</td>
</tr>
<tr>
<td>2</td>
<td>{r, 1, 4, 5, 6, 7, 8, 9, 12, q}</td>
<td>32</td>
<td>10</td>
<td>[0, 23.62]</td>
</tr>
<tr>
<td>3</td>
<td>{r, 1, 2, 6, 7, 8, 9, 12, q}</td>
<td>38</td>
<td>50</td>
<td>[0, 17.62]</td>
</tr>
<tr>
<td>4</td>
<td>{r, 1, 2, 6, 7, 11, 12, q}</td>
<td>44</td>
<td>10</td>
<td>[0, 11.62]</td>
</tr>
<tr>
<td>5</td>
<td>{r, 1, 2, 3, q}</td>
<td>50</td>
<td>380</td>
<td>[0, 5.62]</td>
</tr>
</tbody>
</table>

Figure 5. Departure and arrival curves of the LDEAF.

(a) Link (r, 1); (b) On-ramps;
Figure 6. Link flows in the LDEAF.

Figure 7. Route flows in the LDEAF

The dynamic link flows in the LDEAF are plotted in Figure 6, and route flows are plotted in Figure 7. The inflow rates for routes are constant in the LDEAF, and the resulting link flows exhibit the stair-case pattern. There is no queue within the network in the optimal solution. We plot the total travel cost, including the travel time and schedule delay on the five routes in the LDEAF in Figure 8. The travel time of each chain flow with respect to different departure times is the same while the schedule delay is different. In all five chain flows, the later the departure
time, the lesser the travel cost, due to the reduced schedule delay.

### 6.2. Tolls

We solve for $\lambda^\theta_{ij}$ and specify the link-based tolls as $toll^\theta_{ij} = \lambda^\theta_{ij}$, as tabulated in Table 2 and plotted in Figure 9. At the destination node, a toll $toll^\theta_q = \beta \theta$ is imposed as shown in Table 2 and plotted in Figure 9. The $toll^\theta_q$ increases as the departure times increase, implying less tolls for the earlier arrival and more tolls for the on-time arrival. The time-dependent tolls are computed for the five chain flows and plotted in Figure 10. These route tolls also represent the route externalities. As can be seen, the externalities increase with later departure times, and are in the exact opposite direction of the travel costs illustrated in Figure 8. It can be verified that the generalized costs due to the addition of the travel cost and tolls are equal to the same number (94.496) for the five chain flows at any departure time. This confirms that all utilized routes have the same generalized cost, as plotted in Figure 11. Hence, the LDEAF also exhibits the DUE feature with the toll.

### 6.3. Sensitivity result

In this section we change the demand to examine changes in the chain flows. Consider the following scenario:

*Scenario 1:* demand increases to 6,000;

We apply SSP algorithm to solve LDEAF for this scenario. The resulting chain flow solution is shown in Table 3.
Table 2. Optimal time-dependent link tolls.

<table>
<thead>
<tr>
<th>Time $(\theta)$</th>
<th>$\lambda_{ij}^0$</th>
<th>$\lambda_{ij}^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta)$</td>
<td>(1,4)</td>
<td>(2,6)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>0.8</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>1.6</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>2.4</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>3.2</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>4.8</td>
<td>37</td>
</tr>
<tr>
<td>8</td>
<td>5.6</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td>6.4</td>
<td>39</td>
</tr>
<tr>
<td>10</td>
<td>7.2</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>41</td>
</tr>
<tr>
<td>12</td>
<td>8.8</td>
<td>42</td>
</tr>
<tr>
<td>13</td>
<td>9.6</td>
<td>43</td>
</tr>
<tr>
<td>14</td>
<td>10.4</td>
<td>44</td>
</tr>
<tr>
<td>15</td>
<td>11.2</td>
<td>45</td>
</tr>
<tr>
<td>16</td>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td>17</td>
<td>12</td>
<td>47</td>
</tr>
<tr>
<td>18</td>
<td>12.8</td>
<td>48</td>
</tr>
<tr>
<td>19</td>
<td>13.6</td>
<td>49</td>
</tr>
<tr>
<td>20</td>
<td>14.4</td>
<td>50</td>
</tr>
<tr>
<td>21</td>
<td>15.2</td>
<td>51</td>
</tr>
<tr>
<td>22</td>
<td>16.8</td>
<td>52</td>
</tr>
<tr>
<td>23</td>
<td>17.6</td>
<td>53</td>
</tr>
<tr>
<td>24</td>
<td>18.4</td>
<td>54</td>
</tr>
<tr>
<td>25</td>
<td>19.2</td>
<td>55</td>
</tr>
<tr>
<td>26</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>28</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>29</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td>15</td>
</tr>
</tbody>
</table>
Figure 8. Travel costs in the LDEAF.

Figure 9. Time-dependent tolls on links.
Figure 10. Externalities (tolls) on routes.

Figure 11. Total generalized costs (travel costs + tolls) on routes.
Table 3: LDEAF solution for Scenario 1.

<table>
<thead>
<tr>
<th>Route</th>
<th>Route</th>
<th>Repeating time window</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>{r, 1,4,5,6,7,8,9,q}</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>{r, 1,4,5,6,7,8,9,12, q}</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>{r, 1,2,6,7,8,9,12, q}</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>{r, 1,2,6,7,11,12, q}</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>{r, 1,2,3, q}</td>
<td>50</td>
</tr>
</tbody>
</table>

In this scenario, clearance time increases 2 units due to increased demand. However, the chain flow does not change. That is, it reaches the maximum number of shortest paths in the network. This indicates that in the morning commute problem, the number of utilized paths is rather limited and all of them are saturated when demand is fairly large, as in our scenario. Also, with increased demand, the increase in clearance time is rather small, because traffic can be rerouted effectively through arterials whose capacity is rather large – 500 in our example.

The optimal tolls are shown in Table 4. The italicized numbers illustrate the difference in the tolls compared with Table 2. The results show that the tolls are extended for three time steps, due to the extended clearance time to accommodate extra demand.

Table 4: Optimal time-dependent link tolls for Scenario 1.
Through the SSP algorithm, we can verify the following general results for the dynamic flow pattern in the LDEAF for a morning commute problem when clearance time is minimum:

1. in the SO-DTA with departure time choice, there is no queue in the network;
2. the off-ramps downstream of any utilized off-ramp must be capacitated as they are strictly shorter, and thus this shorter path must be saturated in the SSP algorithm;
3. for a similar reason, the on-ramps upstream of any utilized on-ramp must be capacitated.

7. Concluding comments

This paper investigates the properties of dynamic system optimal assignment with departure time choice and its solution method on a two-terminal network. We show that the SO-DTA with departure time choice reduces to the LDEAF when clearance time is minimum. There is no queue in the network in a LDEAF. The solution is not unique, but the departure and arrival schedules are unique and piecewise linear. Furthermore, the departure and arrival schedules in a
LDEAF exhibit the symmetry property. To solve the LDEAF, a set of chain flows are determined based on the SSP algorithm, and then the chain flows are used to induce dynamic flows over time. Using this method, the LDEAF can be solved at the link level at a run-time that depends on $\log T$. As this method only involves solving chain flows on a static network, the method can be applied on a large-scale network efficiently. This would reduce the run time of SO-DTA substantially. Zheng et al. (2014; 2015) used this method to solve for a medium-sized network (Dallas Fort-Worth network, which consists of 180 nodes and 445 links) with a run time of only 0.6 seconds. Analytical DTA is usually hard to apply to a practical problem because of the difficulty of maintaining tractability in the solution algorithm. In this context, the LDEAF methodology illustrates our contribution relative to SO-DTA studies. The LDEAF technique (chain flows induce dynamic flows over time) is particularly simple and could be applied to large-scale networks.

The system marginal costs and congestion externalities in a LDEAF are analyzed. The dual variable associated with the capacity constraint is the link congestion externality. It is shown that time-dependent link-based system optimal tolls exist, and the LDEAF solution with the extra toll costs exhibits temporal equilibrium. These observations are similar to prior bottleneck model studies in the SO-DTA context (Hendrickson and Kocur, 1981; Arnott et al. 1990; Heydecker and Addison, 2005). Nevertheless, we verify those theoretical results in the LDEAF context. Furthermore, these characteristics are verified in the numerical experiments on a morning commute problem.

Note that although we study the SO-DTA using the P-Q model, the research results can be extended to S-Q and CTM as well. This is because since there are no queues at intermediate nodes in the SO-DTA and LDEAF (c.f. Lemma 2), there is no difference in implementing P-Q,
The method suggested in this study provides another perspective to understanding SO-DTA. It is anticipated to be useful in a wide range of situations where analytical SO-DTA applies, for instance, the morning commute problem, because the morning commute network is a special case of the two-terminal network. A solution method is proposed to determine the LDEAF in a computationally efficient manner. The computational advantage of the proposed method offers promise to address the real-time or online control needs for real-time route guidance on real-sized networks. In terms of applications, the LDEAF solution is helpful to analyze shared bottlenecks, ramp metering and congestion pricing to manage the morning commute trips.

There are some potential future research directions. First, this study considers early schedule delay only. A future direction is to relax this assumption to include both early and late schedule delays. Second, the study scope is restricted to a two-terminal network; a more general network topology is worthy of further investigation.

Acknowledgements
This work is based on funding provided by the U.S. Department of Transportation through the NEXTRANS Center, the USDOT Region 5 University Transportation Center. The authors are solely responsible for the contents of this paper.

References