Predicting Corridor-level Travel Time Distributions Based on Stochastic Flow and Capacity Variations

Hao Lei

Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Email: hlei@eng.utah.edu

Xuesong Zhou

Assistant Professor
Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Tel: 801-585-6590 Fax: 801-585-5477
Email: zhou@eng.utah.edu

George List

Professor
Department of Civil, Construction, and Environmental Engineering
North Carolina State University
Raleigh, NC 27695-7908
Email: gflist@ncsu.edu

Jeffrey Taylor

Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Email: jeff.d.taylor@utah.edu

DRAFT

Abstract

Trip travel time variability is an important measure of transportation system performance and a key factor affecting travelers' choices. This paper aims to establish a point-queue based end-to-end travel time prediction method on a corridor with multiple merges and diverges. A set of analytical equations is developed to calculate the number of queued vehicles ahead of the probe vehicle and further capture many important factors affecting end-to-end travel times: the prevailing congestion level, queue discharge rates at bottlenecks, and flow rates associated with merges and diverges. Based on multiple random scenarios and a vector of arrival times, the experienced delay at each bottleneck along a corridor is recursively estimated to produce end-to-end travel time distributions. The proposed model can incorporate stochastic variations of bottleneck capacity and demand, and explain the travel time correlation between sequential links.

A high-fidelity vehicle trajectory data set available from the NGSIM project is used to verify the proposed method, and the sources of prediction error are systematically examined.

Key words: travel time variability, stochastic capacity and demand, point-queue, NGSIM

1. Introduction

Travel time has long been regarded as one of the most important performance measures in transportation systems. Recently, significant attention has been devoted to evaluating and quantifying travel time variability due to its influences on travelers' mode, route and departure time choices. Operating and management agencies have also increased efforts for monitoring and improving the reliability of transportation systems through probe-based data collection, integrated corridor management and advanced traveler information provision. In particular, vigorous data collection efforts have been made to improve measurement quality and uncover the root sources of travel time unreliability. For instance, a wide range of corridor management strategies have been designed to balance traffic between freeway and arterial corridors in response to various non-recurring traffic congestion sources. In addition, advanced traveler information provision systems have been enhanced to provide reliability-related information to enable travelers to meet their limited travel time budget constraints. There also appears to be a growing trend toward incorporating end-to-end trip travel time variability measures, and their related traveler behavior components, into traffic network analysis and management models.

Although noteworthy progress has been made in travel time variability quantification, there are still a number of theoretical challenges to be addressed in the above practically important applications. The first challenge is how to establish a systematic estimation and prediction framework to capture the stochasticity of the traffic flow under both recurrent and non-recurrent congestion conditions. Such a framework is vital for both travelers and operating agencies (e.g., traffic management team) to obtain accurate time-dependent travel time variability information to make informed decisions. Past research on travel time variability usually assumes either stochastic demand or stochastic capacity. Few studies have considered the travel time variability under both stochastic demand and capacity situations. Additionally, merging and diverging cause significant disturbances on both mainline lanes and ramps. Hence, all of these sources of uncertainty must be carefully examined and addressed for corridor-level travel time variability quantification.

The objective of this paper is to develop an integrated travel time variability prediction framework which can characterize the travel time dynamics through time-dependent demand and capacity fluctuations under both recurrent and non-recurrent congestion conditions. Potential applications for our travel time variability prediction methods are investigated, and an illustrated example application is presented. Particularly, we use high-fidelity NGSIM vehicle trajectory data to validate the predicted travel time distributions under different traffic conditions, and investigate lane-by-lane travel time variability.

1.1 Literature review and motivations

This section first reviews deterministic link-based travel time models and examines the underlying distributions of capacity and demand elements. This is followed by a discussion on the various modeling approaches for connecting travel time variability with its root sources.

1.1.1 Travel time models and capacity/demand element distributions

DRAFT

Within the subject of analytical dynamic traffic network analysis, the "whole-link" model is widely adopted to describe link travel time evolution due to its simple description of traffic flow propagation through an analytical form. The link travel time function introduced by Friesz et al. (1993) defines the travel time $\tau(t)$ on a single link at a time t as a linear function of the number of vehicles $x(t)$ on the link at time t:

$$\tau(t) = a + bx(t) \tag{1}$$

where a and b are constants in the above general linear form. A nondecreasing and continuous function is defined to calculate the number of vehicles on the link based on the inflow and outflow rates, $u(t)$ and $v(t)$, at time t:

$$x(t) = x(0) + \int_0^t \big(u(s) - v(s)\big)ds \tag{2}$$

Meanwhile, some more general non-linear travel time functions have been proposed as:

$$\tau(t) = f\big(x(t), u(t), v(t)\big). \tag{3}$$

A special case of this form, introduced by Ran et al. (1993), decomposes the link travel time as two different functions: g_1 accounts for flow-independent travel time and g_2 accounts for the queuing delay. A detailed mathematical representation is shown below.

$$\tau(t) = g_1\big[x(t), u(t)\big] + g_2\big[x(t), v(t)\big] \tag{4}$$

They later showed that, by assuming g_1 and g_2 are separable, i.e., $g_1 = g_{1a}\big[x(t)\big] + g_{1b}\big[u(t)\big]$ and $g_2 = g_{2a}\big[x(t)\big] + g_{2b}\big[v(t)\big]$, and Eq. 4 can be rewritten as

$$\tau(t) = \alpha + f\big(u(t)\big) + g\big(v(t)\big) + h\big(x(t)\big) \tag{5}$$

where α is the free flow travel time, and $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ correspond to the functions of link inflow rate, link outflow rate and the number of vehicles on the link, respectively.

Daganzo (1995) draws attention to problems with the general form in Eq. (3), indicating that either a rapid decline in the inflows $u(t)$ or a rapid increase in outflow $v(t)$ would lead to unrealistic travel time. Thus, he recommended omitting $u(t)$ and $v(t)$ from Eq. 3, reducing the link travel time to a function of the number of vehicles on the link, that is, $\tau(t) = f\big(x(t)\big)$.

Although the link travel time function models provide some degree of simplification on travel time analysis, there is one significant drawback. Traffic congestion usually occurs at some bottleneck, and queues are produced and often grow beyond the bottleneck, which is difficult for any travel time function to capture (Zhang and Nie, 2005).

In dynamic traffic assignment and other applications, the vertical queue or point-queue model (Daganzo, 1995) was widely adopted to describe bottleneck traffic dynamics (Zhang and Nie, 2005). In a queuing-based travel time model, it is important to capture the variations of queue discharge flow rates and incoming demand to a bottleneck.

DRAFT

Conventionally, freeway capacity is viewed as a constant value – the maximum discharge flow rate before failure (HCM, 2000). However, the capacities vary according to different external factors in real life situations. Conceptually, capacity or discharge flow rate can be represented as the reciprocal of the average of vehicle headways. Over the past decades, many researchers have developed a number of headway models to describe its distribution. Representatives of these models include the exponential-distribution by Cowan (1975), normal distribution, gamma-distribution, and lognormal-distribution models by Greenberg (1966).

Incidents are one of the major contributing factors in capacity reductions, and the magnitude and duration of capacity reductions are directly related to the severity and duration of incidents (Kripalari and Scherer, 2007; Guiliano, 1989). In quantifying capacity reduction, the HCM 2000 provides guidance for estimating the remaining freeway capacity during incident conditions. Using over two years of data collected on freeways in the greater Los Angeles area, Golob et al. (1987) found that accident duration fit a lognormal distribution. By extending the research of Golob et al., Guiliano (1989) applied a lognormal distribution when analyzing incident duration for 512 incidents in Los Angeles.

It is commonly observed that travel demand fluctuates significantly within a day. During the morning and evening peak hours, surging demand may overwhelm a roadway's physical capacity and results in delays (FHWA, 2009). Waller and Ziliaskopoulos (2001), Chen et al. (2003) and Lam et al. (2008) have used the normal distribution for modeling travel demand variation. Other researchers have modeled travel demand using the Poisson distribution (Hazelton 2001; Clark and Watling, 2005) and the uniform distribution (Ukkusuri et al. 2005).

1.1.2 Methods for estimating travel time variability

Substantial efforts have been devoted to travel time variability estimation over the last decade, producing several different approaches for estimating travel time variability. Statistical approaches (Richardson 2003; Oh and Chung 2006) have been widely adopted to quantify travel time variability from archived sensor data. In recent studies investigating the different sources of travel time variability, Kwon et al. (2010) proposed a quantile regression model to quantify the 95th percentile travel time based on the congestion source variables, such as incidents, weather. In their multi-state travel time reliability modeling framework, Guo et al. (2010) provided connections between the travel time distributions and the uncertainty associated with the traffic states, e.g., with incidents vs. without incidents. In addition, they (Park et al. 2010) show that a multi-mode model could lead to better representations of real-world observations compared to single-mode models (represented by mean and variance parameters).

A second approach uses numerical approximation methods to characterize travel time variability distributions as a result of stochastic capacity and stochastic demand. Given a stochastic capacity probability distribution function (PDF), a Mellin transforms-based method was adopted by Lo and Tung (2003) to estimate the mean and variance of travel time distributions. Using a sensitivity analysis framework, Clark and Watling (2005) developed a computational procedure to construct a link travel time PDF under stochastic demand conditions. Given various sets of traffic flow assignment results, Ng and Waller (2010) introduced a fast Fourier transformation approach to approximate the travel time PDF from underlying stochastic capacity distributions. Although it can quantify the impacts of demand and capacity variation on the travel times, the steady-state travel time function-based approach is still unable to address the underlying time-dependent traffic dynamics.

In order to account for the inherent time-dependent traffic dynamics, some researchers have incorporated point-queue models into travel time variability estimation techniques.

DRAFT

Assuming lognormal distributions on capacity and demand, Zhou et al. (2010) adopted a point-queue model and a cumulative count curve approach to quantify the day-to-day travel time variability. For single bottlenecks, the travel time variability is analytically derived from the variation parameters in demand and capacity. The challenging issue in extending their model on a corridor-level analysis is how to quantify end-to-end travel time along several corridors where downstream and upstream traffic states are correlated. Using a dynamic traffic assignment simulator, Alibabai (2010) developed an algorithmic framework to investigate the properties of the path travel time function with respect to various path flow variables. While realistic simulation results require significant efforts in simulation/assignment model calibration, this approach is particularly suited for studying the effects of various uncertainty sources and assessing the benefits of traffic management strategies and traffic information systems.

The rest of the paper is organized as follows. Section 2 describes the point-queue based end-to-end travel time estimation framework with deterministic inputs. Monte Carlo simulation is applied in Section 3 to compute the travel time distribution with stochastic inflow, outflow and discharge rates. High-fidelity vehicle trajectory data from the NGSIM project is utilized to validate our methods in Section 4.

2. Computing end-to-end time-dependent delay

2.1 Problem statement

Consider a corridor with *M* bottlenecks, where each node in a node-link structure represents a bottleneck, and the road segments between consecutive bottlenecks are modeled as a link with homogeneous capacity. That is, node *0* is the starting point of the corridor, node *m* corresponds to bottleneck *m*, and each link between bottlenecks is denoted as link (*m-1*, *m*), for $1 \le m \le M$. Link (*m-1*, *m*) can also be described as link *m* interchangeably. Figure 1 illustrates a node-link representation for a corridor with *M* bottlenecks. Possible merge or diverge nodes are connected to bottleneck *m* and are denoted as *m'* or *m''*, respectively, so that the on-ramp before node 1 is denoted as (1', 1), and the off-ramp before node 2 is denoted as (2, 2''). In other words, we assume merge and diverge links are directly connected to the bottleneck.

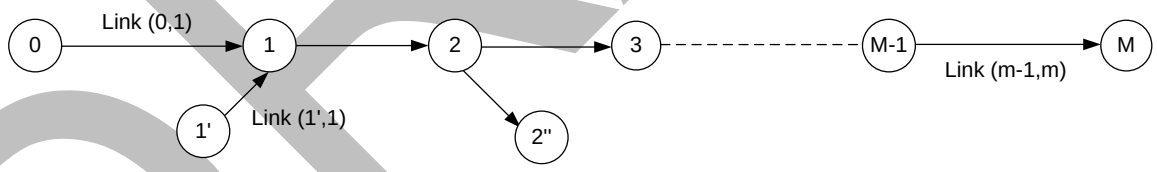


Figure 1: A node-link representation of a corridor with *M* bottlenecks

In this study, we are interested in how to predict the travel time from bottleneck *0* to bottleneck *m* for a probe vehicle *z*, departing at time $t_0 = 0$. The proposed model aims to find the end-to-end travel time, p_m^z , based on the following given conditions: (1) the number of vehicles $x_m(t_0)$ on each link *m* along the path at time t_0, (2) the discharge flow rate for each bottleneck c_m , and (3) the on-ramp or off-ramp flow rates f_m^{net} . For this study, the end-to-end travel time is defined as the difference between the departure time at bottleneck *0* and the departure time at bottleneck *m* for probe vehicle *z*. The departure time at bottleneck *m* is defined as the time the probe vehicle leaves the queue at bottleneck *m*, which is the time when the number of vehicles in the queue before probe vehicle *z* at bottleneck *m* is 0.

DRAFT

The number of vehicles on a link is assumed to be observable from sensors, such as loop detectors, and the discharge flow rates and net flow rates from on-ramps and off-ramps are assumed to be estimated from historical flow patterns or predicted based on prevailing traffic conditions (e.g., capacity reduction due to incidents).

The notation for end-to-end travel time is described below.

Indices:

z: index for identifying a probe vehicle;

k: index of a simulation instance used in Monte Carlo simulation;

m: index of bottlenecks and links along the corridor.

Input:

t_0 : starting time, $t_0 = 0$.

M: number of bottlenecks along the corridor of interest;

$FFTT_m$: free-flow travel time over link (*m-1, m*);

c_m : queue discharge rate of bottleneck *m*;

f_m^{net} : net flow rate at a merge or diverge corresponding to bottleneck *m*, that is, from an on-ramp to the mainline segment or from the mainline to the off-ramp;

$x_m(t_0)$: number of vehicles on link (*m-1, m*) at time t_0 ;

$\mu_m(t)$: arrival rate of link *(m-1, m)* at time *t*;

$v_m(t)$: departure rate of link *(m-1, m)* at time *t*;

Variables to be calculated:

$\tau_m^z(t)$: travel time on link *m* for probe vehicle *z* entering the link at time *t*;

$\lambda_m(t)$: number of vehicles waiting at the bottleneck *m* at time *t*, that is, the number of queued vehicles behind bottleneck *m*;

w_m^z : waiting time in the vertical queue of bottleneck *m* for probe vehicle *z*;

t_m^z : arrival time for probe vehicle *z* at bottleneck *m*;

p_m^z : end-to-end path travel time from node *0* to bottleneck *m*.

DRAFT

2.2 Travel time calculation

In a point-queue model, a link can be considered as two segments: the free-flow segment and the queuing segment. A vehicle can always travel at the free-flow speed on the free-flow segment until reaching the beginning of the queuing segment, where this vehicle joins the end of the queue waiting to be discharged. A queue is only formed if the link demand exceeds the bottleneck capacity, or the link arrival rate exceeds the link departure rate.

To construct a deterministic, numerically tractable form for calculating the end-to-end travel delay along a corridor with multiple bottlenecks, several important assumptions are made in this study.

1) A point-queue model is adopted to calculate the delay on each link. On each link, a FIFO (First-In, First-Out) property is assumed to assure that any vehicles that enter the link before time *t* will exit the link before those entering after time *t*.
2) The link traversal time is composed of free-flow travel time and queuing delay. The free-flow travel time is constant and flow-independent. The queuing delay is determined by the number of vehicles in the queue when the probe vehicle arrives at the bottleneck $\lambda\left(t+FFTT\right)$ and the bottleneck queue discharge rate c_m. Thus, the link traversal time is

$$\tau_m\left(t\right)=FFTT_m+w(t+FFTT_m)=FFTT_m+\frac{\lambda_m\left(t+FFTT_m\right)}{c_m} \tag{6}$$

   where $w\left(t+FFTT\right)$ is the queuing delay when vehicle *z* reaches the vertical queue of bottleneck at time $t+FFTT$.

3) The merge and diverge links are connected to the beginning of the queuing segment for each bottleneck.
4) The bottleneck *m* remains congested in the prediction horizon, which extends from the current time t_0 to the arrival time of the probe vehicle *z* at the bottleneck *m*, t_m^z. The corresponding queue discharge rates c_m and net flow rates f_m^{net} in the prediction horizon are assumed to be constant.

The first two assumptions are widely used in a typical queuing model. The third assumption makes it easy to incorporate the flow rate from a merge/diverge without implicitly considering the driving distance and free-flow travel time from the merge/diverge point to the bottleneck *m*.

Eq. (6) considers the arrival time at the beginning of a link. By considering the arrival time at the queue of bottleneck *m* for vehicle *z*, t_m^z, the link traversal time can be rewritten as

$$\tau_m\left(t_m^z-FFTT_m\right)=FFTT_m+\frac{\lambda_m\left(t_m^z\right)}{c_m} \tag{7}$$

For a general queue with time-dependent arrival and departure rates, a continuous transition model can be used in Eq. (6) to update the number of vehicles in the queue at any given time *t*.

DRAFT

$$\frac{d\lambda_m(t)}{dt} = \mu_m(t - FFTT_m) - v_m(t) \tag{8}$$

The number of queued vehicles $\lambda_m(t_m^z)$ at time t_m^z on bottleneck *m* can be derived from the queue length updating Eq. (8), as shown in Eq. (9).

$$\begin{aligned}\lambda_m(t_m^z) &= \lambda_m(t_0) + \int_{t_0}^{t_m^z} \frac{d\lambda_m(t)}{dt} dt = \lambda_m(t_0) + \int_{t_0}^{t_m^z} \left[\mu_m(t - FFTT_m) - v_m(t)\right] dt \\ &= \lambda_m(t_0) + \int_{t_0}^{t_m^z} \mu_m(t - FFTT_m) dt - \int_{t_0}^{t_m^z} v_m(t) dt\end{aligned} \tag{9}$$

Since the fourth assumption states that the bottleneck remains congested in the prediction horizon, the departure rate is equal to the bottleneck capacity $\int_{t_0}^{t_m^z} v_m(t)dt = c_m \times (t_m^z - t_0)$. The remaining challenge is to estimate the unknown queue length $\lambda_m(t_0)$ at time t_0, and calculate the complex integral of $\int_{t_0}^{t_m^z} \mu_m(t - FFTT_m)dt$.

For illustrative purposes, the following discussions start with the first bottleneck from Figure 1, where *m* = 1. In this case, the number of vehicles $x_1(t_0)$ and the net flow rate f_1^{net} associated with bottleneck 1 are given. For the specific starting time t_0, a probe vehicle *z* enters the vertical queue of bottleneck 1 at time $t_1^z = t_0 + FFTT_1$, and the number of vehicles in the queue at time t_1^z is

$$\lambda_1(t_1^z) = \lambda_1(t_0) + \int_{t0}^{t_1^z} \mu_1(t - FFTT_1)dt - c_1 \times (t_1^z - t_0) \tag{10}$$

Let us consider a simpler case without merge and diverge points, i.e., f_1^{net} = 0. Thanks to the first-in and first-out property, we can now show that $\lambda_1(t_0) + \int_{t0}^{t_1^z} \mu_1(t - FFTT_1)dt = x_1(t_0)$. The left-hand side $\lambda_1(t_0) + \int_{t0}^{t_1^z} \mu_1(t - FFTT_1)dt$ is the total number of vehicles stored in both the free-flow segment and the queuing segment before the probe vehicle *z*. The right-hand side is the actual number of vehicles observed on the *physical* link. One can use Figure 2 to map or "rotate" some of the vehicles from the physical link (shaded) to the vertical stack queue, and the other vehicles on the physical link (not shaded) correspond to the vehicles that will arrive at the vertical queue between time t_0 and time $FFTT_1$ (that is, right before the probe vehicle). Notice that the length of the queue segment in the point-queue model is equal to zero and has unlimited storage capacity. Interested readers are referred to the paper by Hurdle and Son (2001) to examine the connection between physical queues and vertical stack queues.

The individual components in Equation (10) are described visually using the cumulative vehicle count curves shown in Figure 3. The cumulative arrival curve A is equivalent to the integral over the arrival rate, $A(t) = \int \mu_1(t)\,dt$, and the cumulative arrival curve at the vertical stack queue V is the cumulative arrival curve shifted by the free-flow travel time, $V(t) = A(t - FFTT)$, and thus $V(t) = \int \mu_1(t - FFTT_1)dt$. The cumulative departure curve D is

DRAFT

equivalent to the integral over the departure rate, $D(t) = \int v_1(t)dt = c_1 \times (t_1^z - t_0)$. Substituting t with values of t_0 and t_1^z for $V(t)$, Figure 3 shows that $V(t_1^z) - V(t_0) = \int_{t0}^{t_1^z} \mu_1(t - FFTT_1)dt$ and thus $\lambda_1(t_1^z) = \lambda_1(t_0) + \int \mu_1(t - FFTT_1)dt - c_1 \times (t_1^z - t_0)$ and $x_1(t_0) = \lambda_1(t_0) + \int_{t0}^{t_1^z} \mu_1(t - FFTT_1)dt$.

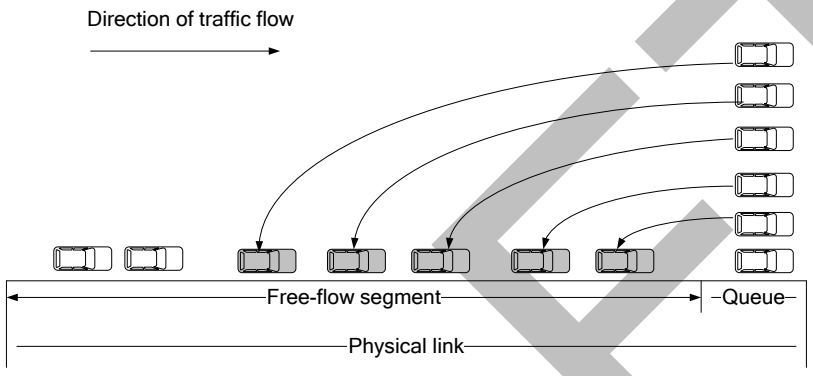


Figure 2: A vertical stack queue

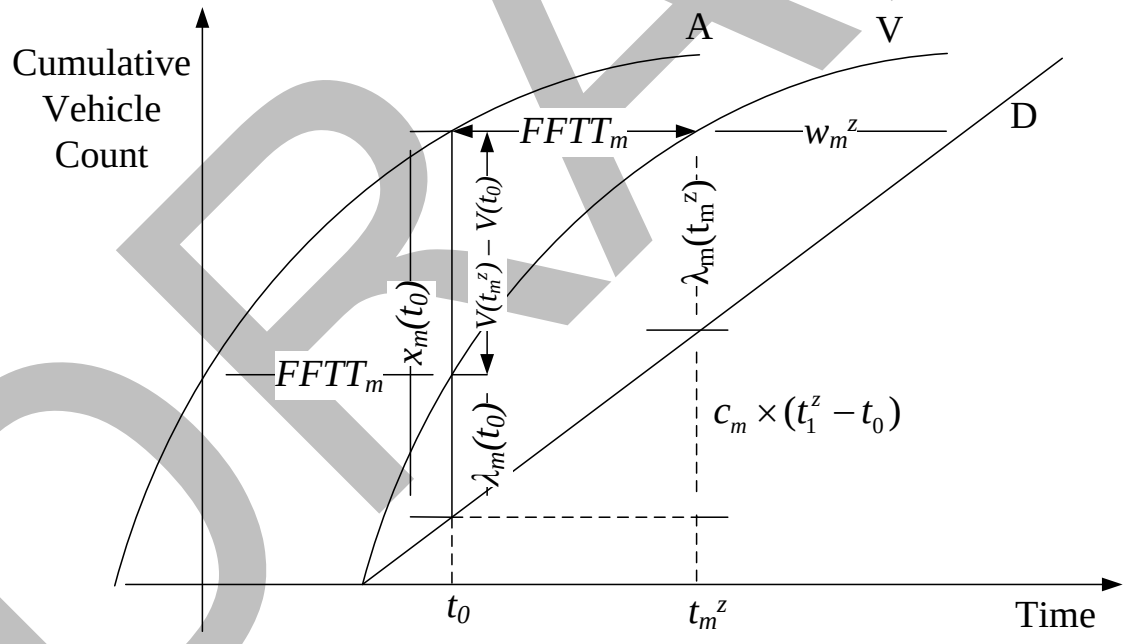


Figure 3: Visual representation for Equation (10)

By further considering the net flow rate from the merge or diverge point connected to the bottleneck, we now have

$$\lambda_1(t_1^z) = \lambda_1(t_0) + \int_{t0}^{t_1^z} [\mu_1(t - FFTT_1) - c_1]dt = x_1(t_0) + f_1^{net} \times t_1^z - c_1 \times t_1^z \quad (11)$$

DRAFT

Continuing to link 2 in Figure 1, the probe vehicle *z* will arrive at the queue of bottleneck 2 at t_2^z. Again, considering the FIFO assumption, the number of vehicles transferring from the first link to the second link before the probe vehicle *z* includes two terms, $\lambda_1\left(t_1^z\right)+c_1\times t_1^z$, which are the number of queued vehicles $\lambda_1\left(t_1^z\right)$ when the probe vehicle arrives at the first bottleneck at time t_1^z, and those vehicles $c_1\times t_1^z$ already entering the second link before time t_1^z. Following the derivation logic for Eq. (11), the number of vehicles waiting in the queue ahead of vehicle *z* when it arrives at the second bottleneck at time t_2^z is

$$\lambda_2\left(t_2^z\right)=\lambda_1\left(t_1^z\right)+c_1\times t_1^z+x_2\left(t_0\right)+f_2^{net}\times t_2^z-c_2\times t_2^z \tag{12}$$

By substituting $\lambda_1\left(t_1^z\right)$ from Eq. (11), Eq. (12) reduces to

$$\begin{aligned}\lambda_2\left(t_2^z\right)&=x_1\left(t_0\right)+f_1^{net}\times t_1^z-c_1\times t_1^z+c_1\times t_1^z+x_2\left(t_0\right)+f_2^{net}\times t_2^z-c_2\times t_2^z\\&=x_1\left(t_0\right)+f_1^{net}\times t_1^z+x_2\left(t_0\right)+f_2^{net}\times t_2^z-c_2\times t_2^z\end{aligned} \tag{13}$$

More generally, for bottleneck *m*:

(1) The number of vehicles waiting at the vertical queue of bottleneck *m* at time t_m^z can be expressed as

$$\lambda_m\left(t_m^z\right)=\sum_{i=1}^{m}x_i\left(t_0\right)+\sum_{i=1}^{m}\left(f_i^{net}\times t_i^z\right)-c_m\times t_m^z \tag{14}$$

(2) The arrival time for the probe vehicle at bottleneck *m* is

$$t_m^z=t_{m-1}^z+w_{m-1}^z+FFTT_m=t_{m-1}^z+\frac{\lambda_{m-1}\left(t_{m-1}^z\right)}{c_{m-1}}+FFTT_m \tag{15}$$

where $w_m^z=\dfrac{\lambda_m\left(t_m^z\right)}{c_m}$.

(3) Finally, the end-to-end travel time from bottleneck 0 to bottleneck *m* is

$$p_m^z=t_m^z+w_m^z=\sum_{i=1}^{m}\left[FFTT_i+w_i^z\right] \tag{16}$$

In a summary, given the number of vehicles on each link, the queue discharge rate and the net flow on each bottleneck, the end-to-end travel time for a vehicle can be calculated by applying Eqs. (14-16) iteratively from link 1 to link *m*. At each iteration, first apply Eq. 14 to obtain the number of queued vehicles at the bottleneck, then compute the queuing delay and update the end-to-end travel time up to the bottleneck of interest.

2.3 Illustrative example

To demonstrate how to use our model to calculate the end-to-end travel time and capture the delay propagation along a corridor, a short corridor with 3 bottlenecks (Figure 4) is used as an

DRAFT

illustrative example. In this example corridor, bottleneck 2 is connected with an on-ramp and bottleneck 3 is connected with an off-ramp. The bottleneck discharge rates for those bottlenecks are 90, 90 and 60 vehicles/min, respectively. The existing numbers of vehicles on each link are 750, 600 and 650, respectively. The inflow rate from the on-ramp to bottleneck 2 is 20 vehicles/min (vpm), and the outflow rate to the off-ramp is 18 vehicles/min (vpm). The free-flow travel time over each link is 5, 4 and 4.5 minutes, respectively.

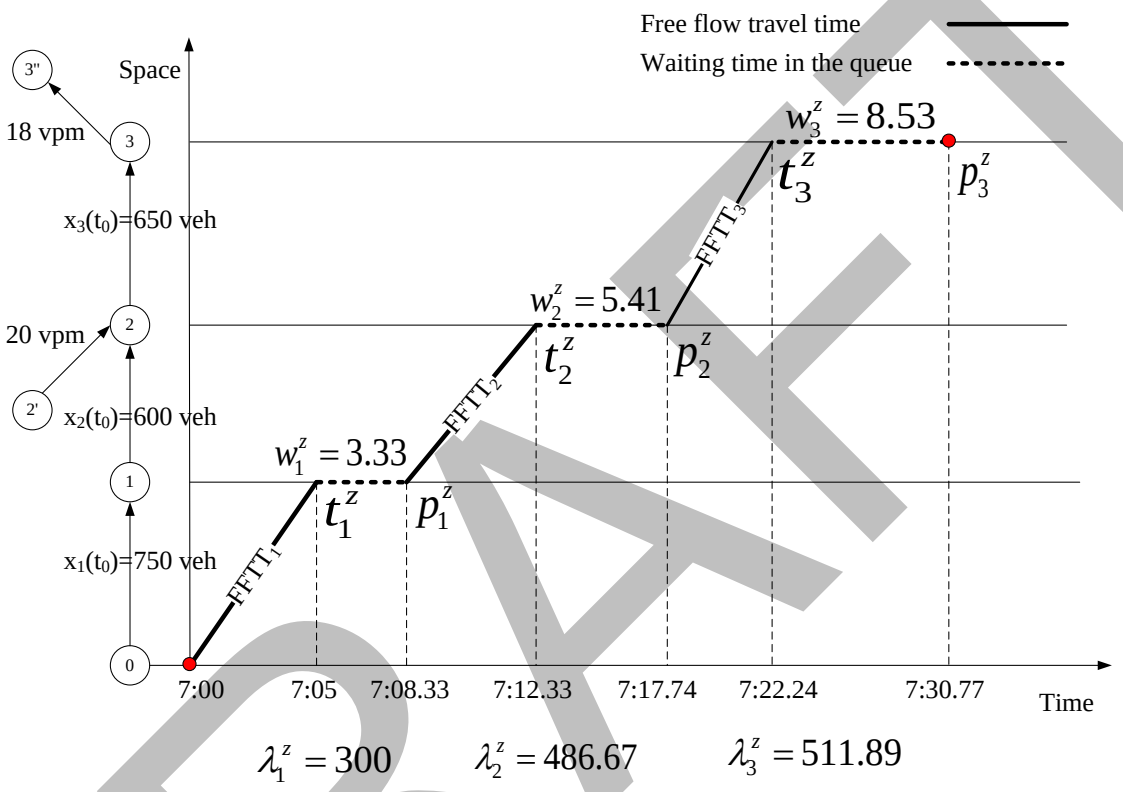


Figure 4: 3-bottleneck example corridor

For the probe vehicle in Figure 4 (starting at time 7:00 AM), we now have the following calculation process for its end-to-end travel time.

1) Departing at 7:00, it takes 5 minutes (free-flow travel time) for this probe vehicle to reach the point-queue of bottleneck 1 at 7:05. At this time instance, the number of vehicles waiting in the queue is $\lambda_1\left(t_1^z\right)$ = 750 – (5 min*90 veh/min) = 300 vehicles. With the discharge rate of 90 vehicles/min, this probe vehicle will spend w_1^z = 3.33 minutes waiting in the queue. Thus, the total travel time for this vehicle is 8.33 minutes at the end of this bottleneck.

2) The probe vehicle enters link 2 at 7:08.33, spends 4 minutes traveling through the free-flow segment, and arrives at the vertical stack queue at t_2^z = 7:12.33. From 7:00 to 7:12.33, there have been 12.33 min*20 veh/min = 246.6 vehicles entering this bottleneck from the on-ramp. The number of vehicles waiting in the queue at this time is $\lambda_2\left(t_2^z\right)$ = (750 + 600) + (12.33*20) – (12.33*90) = 486.67. With the discharge rate of 90

DRAFT

vehicles/min, this vehicle leaves the queue w_2^z =5.41 minutes later. The departure time from the second bottleneck is 7:17.74.

3) Following the same calculation process, the number of vehicles waiting at the queue of bottleneck 3 is $\lambda_3\left(t_3^z\right)$ = (750 + 600 + 650) + (12.33*20) + (-18*22.24) – (60*22.24) = 511.89 vehicles and the waiting time in the queue is w_3^z = 8.53 minutes. This vehicle leaves bottleneck 3 at 7:30.77. The total end-to-end-travel time p_3^z is 30.77 minutes.

2.4 Travel time calculation algorithm with deterministic inputs

The algorithm for calculating the end-to-end path travel time for vehicle *z* entering the corridor with *M* bottlenecks at time t_0 is summarized below.

Input: The specific starting time t_0, the number of vehicles on each link $x_m\left(t_0\right)$, the net flow rate on each bottleneck f_m^{net}, and the bottleneck discharge rate c_m, at time t_0

End-to-end travel time calculation

For *m* = 1 to *M*

1. Calculate the arrival time at bottleneck *m*
$$t_m^z = t_{m-1}^z + w_{m-1}^z + FFTT_m, \text{ where } t_0^z = 0,\ w_0^z = 0.$$
2. Use Eq. (14) to calculate the number of vehicles ahead of the probe vehicle *z* in the vertical stack queue of bottleneck *m*, $\lambda_m^z\left(t_m^z\right)$, when the probe vehicle *z* reaches the beginning of the queue at time t_m^z.
3. Use Eq. (15) to calculate the delay experienced by the probe vehicle on bottleneck *m*, w_m^z.
4. Use Eq. (16) to update end-to-end travel time over *m*, p_m^z

End For

Output: The end-to-end travel time p_M^z from bottleneck 1 to bottleneck *M*.

2.5 Discussions

To consider complex real-life conditions, the proposed analytical framework must further use the following approximation methods for calculating the end-to-end path travel time along a corridor with multiple bottlenecks.

2.5.1 Approximating time-dependent flow rate with average flow rate

In Eq. (9) to (10), we use the maximum bottleneck discharge rate to approximate the actual discharge rate. In reality, the traffic flow rates (including the queue discharge flow rates and net flow rates from and to ramps) are highly dynamic and could fluctuate significantly even in a short time interval, as shown in Figure 5. In this situation, one needs to use the average flow rate (i.e. the dashed line in Figure 5) during the interval from t_0 to the arrival time t_m^z to approximate the time-dependent volume. Although this approximation ignores the traffic

DRAFT

dynamics, it can still give a reasonable estimate about the total number of vehicles leaving or entering the bottleneck before the probe vehicles in Eq. (14). Section 4 will examine the possible impact from using this approximation method for representing travel time distributions.

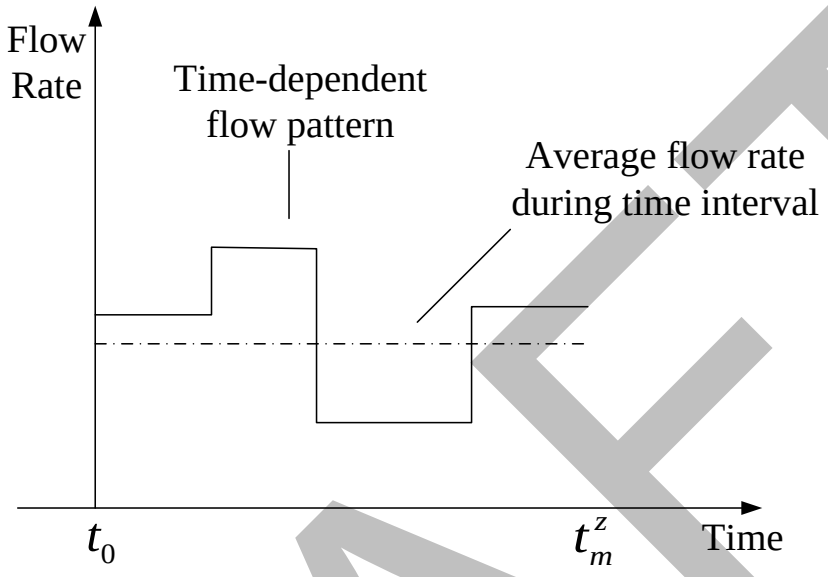


Figure 5: Time-dependent flow rate to average flow rate

2.5.2 Considering further reduced bottleneck discharge flow rate due to queue spillback

The proposed point-queue based model needs to capture the effects of queue spillback from a downstream bottleneck. Essentially, when a queue spillback occurs, the discharge capacity from the upstream bottleneck is then constrained by the discharge rates at the downstream bottleneck. In this case, the proposed method should first detect the possible queue spillback, and then use the reduced queue discharge rate for calculating the waiting time at the bottleneck with queue spillback.

As illustrated in Figure 6, the physical queue on bottleneck *m* spills back to bottleneck (*m-1*) between time t_1 and t_5 through the backward wave lines. Interested readers are referred to the paper by Newell, 1993. Due to the queue spillback from bottleneck *m*, the actual discharge rate c'_{m-1} of bottleneck (*m-1*) between time t_1 and t_5 is constrained by the discharge rate of bottleneck *m*, c_m, rather than the original discharge rate c_{m-1}. Suppose at time t_2 (where $t_2 > t_1$), a probe vehicle arrives at bottleneck (*m-1*), if the effect of queue spillback is not taken into account, this probe vehicle in the model will leave bottleneck (*m-1*) at time t_3 after waiting in the queue behind bottleneck (*m-1*), using the original unaffected queue discharge rate c_{m-1}. With the reduced discharge rate $c_m < c_{m-1}$ at bottleneck (*m-1*), the actual waiting time for the probe vehicle will be longer with a departing time of $t_4 > t_3$.

DRAFT

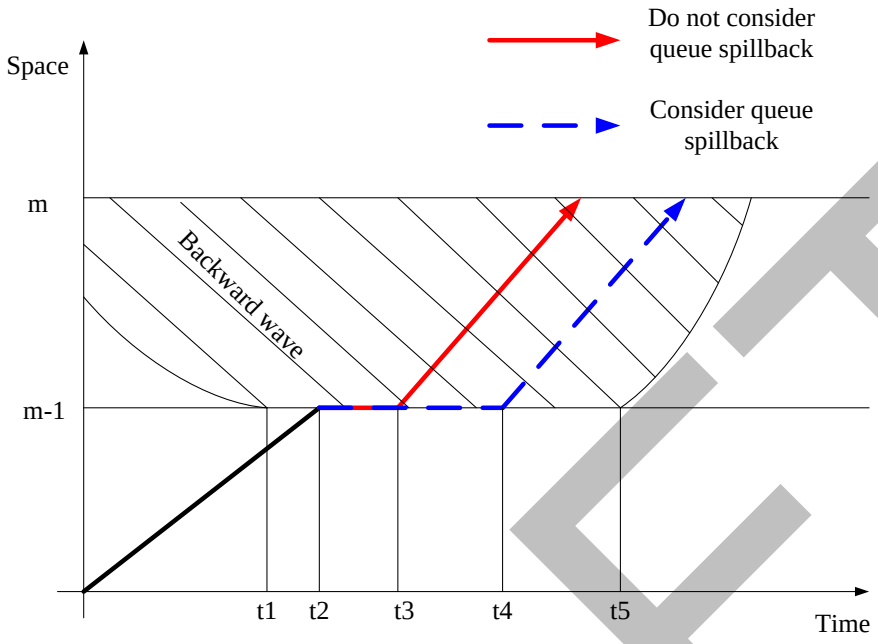


Figure 6: Queue spillback

2.5.3 Calculating the net flow rate for on-ramps

When estimating the net flow at merge or diverge bottlenecks, the flow rates in previous examples are assumed to be known and time-invariant. However, special attention must be paid to the following scenario.

If the mainline and the on-ramp are both congested, (a) the number of vehicles that can enter the bottleneck from the on-ramp and (b) the number of vehicles that can enter from the upstream segment to the bottleneck are constrained by the mainline bottleneck discharge rate. In this case, the available bottleneck discharge rate should be allocated to the upstream segment and the on-ramp proportionally, according to certain rules (Zhang and Nie, 2005). One simple rule is to split the mainline discharge rates according to the number of lanes associated with each incoming approach.

2.5.4 Considering vehicle overtaking/passing

Lastly, the FIFO property assumed on each link rules out the possibility that a vehicle can overtake and pass another vehicle. Future research will consider this condition's impact on end-to-end travel time estimation using this approach.

3. Methods for calculating end-to-end travel time distributions

3.1 Assumptions

In the previous discussion, input parameters such as the net rates f_m^{net} at merge and diverge, and the bottleneck discharge rate c_m are assumed to be deterministic. In this section, we will further consider the variations or uncertainty in the input parameters, especially in the following two applications: (1) day-to-day travel time variability estimation by considering flow

DRAFT

variations at the same time period, and (2) real-time travel time reliability prediction, where the near-future traffic flows are predicted from different sources of data with various degrees of prediction uncertainty. Emphases are placed on how to calculate the end-to-end travel time distribution based on the stochasticity of the random input parameters.

3.2 Important observations on path travel time

3.2.1 Simple corridor without merge and diverge

Consider a simple two-bottleneck corridor with no on-ramp and off-ramp, that is, f_1^{net} and f_2^{net} are equal to 0. According to Eqs. (14-16), the end-to-end travel time to bottleneck 1 for probe vehicle z entering link 1 at time t_0 is:

$$\begin{aligned} p_1^z &= t_1^z + w_1^z = t_0 + FFTT_1 + \frac{x_1(t_0) - c_1 \times t_1^z}{c_1} = t_0 + FFTT_1 + \frac{x_1(t_0) - c_1 \times (t_0 + FFTT_1)}{c_1} \\ &= t_0 + FFTT_1 + \frac{x_1(t_0)}{c_1} - (t_0 + FFTT_1) = \frac{x_1(t_0)}{c_1} \end{aligned} \tag{17}$$

And the end-to-end travel time to bottleneck 2 is:

$$\begin{aligned} p_2^z &= p_1^z + FFTT_2 + w_2^z = p_1^z + FFTT_2 + \frac{x_1(t_0) + x_2(t_0) - c_2 \times t_2^z}{c_2} \\ &= p_1^z + FFTT_2 + \frac{x_1(t_0) + x_2(t_0) - c_2 \times (p_1^z + FFTT_2)}{c_2} \\ &= p_1^z + FFTT_2 + \frac{x_1(t_0) + x_2(t_0)}{c_2} - (p_1^z + FFTT_2) = \frac{x_1(t_0) + x_2(t_0)}{c_2} \end{aligned} \tag{18}$$

By comparing $p_1^z = \frac{x_1(t_0)}{c_1}$ and $p_2^z = \frac{x_1(t_0) + x_2(t_0)}{c_2}$, we can make the following important observation. That is, the proposed formula can correctly capture the correlations between the end-to-end travel times p_1^z and p_2^z, as both values are dependent on the number of vehicles on link 1, $x_1(t_0)$. If $x_1(t_0)$ and $x_2(t_0)$ are assumed to be deterministic, the distributions of p_1^z and p_2^z are further dependent on the distribution of the bottleneck discharge rates, c_1 and c_2, respectively.

3.2.2 Simple corridor with merge and diverge

If we further consider situations where a merge and diverge occur at both bottlenecks, the path travel time formulas are expressed as follows.

$$p_1^z = t_1^z + w_1^z = \frac{x_1(t_0) + f_{net}^1 \times (t_0 + FFTT_1)}{c_1} \tag{19}$$

DRAFT

$$p_2^z = \frac{x_1(t_0) + f_1^{net} \times FFTT_1 + x_2(t_0) + f_2^{net} \times \left(p_1^z + FFTT_2\right)}{c_2} \quad (20)$$

The above equations introduce more complex dependencies for both p_1^z and p_2^z, and no additive formula or decomposed elements can be easily constructed to simplify these complicated equations. This observation reinforces many previous research results, that is, the analytical quantification of the end-to-end travel time distribution is extremely challenging.

3.3 Monte Carlo simulation

Monte Carlo simulation is widely used to simulate the behavior of various physical and mathematical systems, especially for those problems with significant uncertainty in inputs. This research aims to apply a Monte Carlo simulation method to investigate the end-to-end travel time distribution based on the proposed travel time calculation framework. In each simulation run, a realization of random input parameters will lead to a realization of random path travel time outputs, which can be regarded as estimates of the true end-to-end travel time variable. A sufficient number of simulations then provide a good representation of the travel time distributions under various traffic conditions and uncertainties.

The following procedure assumes all random variables are log-normally distributed, and calculates travel time distribution through *K* simulation runs.

Input:

The specific starting time t_0,

The distribution of the number of vehicles on each link $x_m(t_0)$, where $x_m(t_0) \sim LN\left(\mu_{x_m}, \sigma^2_{x_m}\right)$

The distribution of the net flow rate on each bottleneck f_m^{net}, where $f_m^{net} \sim LN\left(\mu_{f_m^{net}}, \sigma^2_{f_m^{net}}\right)$

The distribution of the bottleneck discharge rate c_m on each bottleneck, at time t_0 where $c_m \sim LN\left(\mu_{c_m}, \sigma^2_{c_m}\right)$

Link free-flow travel time $FFTT_m$, assumed to be constant.

Number of simulations = *K*.

For k=1 to K,

For m=1 to M

1: Based on the underlying distribution parameters (μ and σ) of the individual inputs, generate a set of random samples for the following key variables: the number of vehicles on the link, the bottleneck discharge rate and net flow rates.

2: Call the algorithm introduced in Section 2.4 to calculate the predicted end-to-end travel time for simulation *k*: $p_m^z[k]$ from this set of random samples.

DRAFT

End For

End For

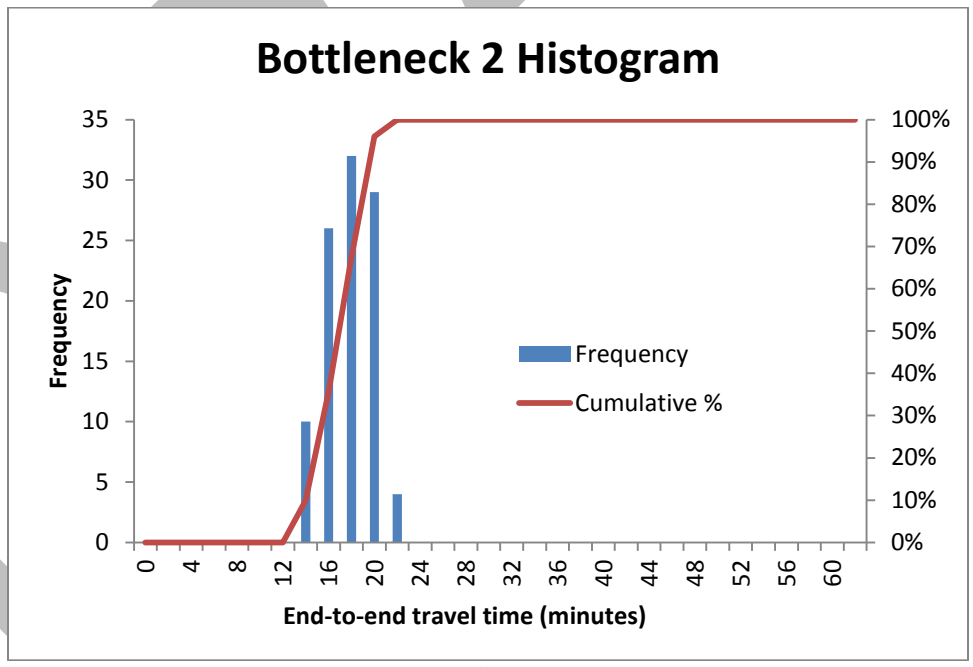
Output: Calculate the histogram, mean and variance for the end-to-end travel time from the results over *K* simulation runs.

3.4 Numerical experiments

3.4.1. Monte Carlo simulation

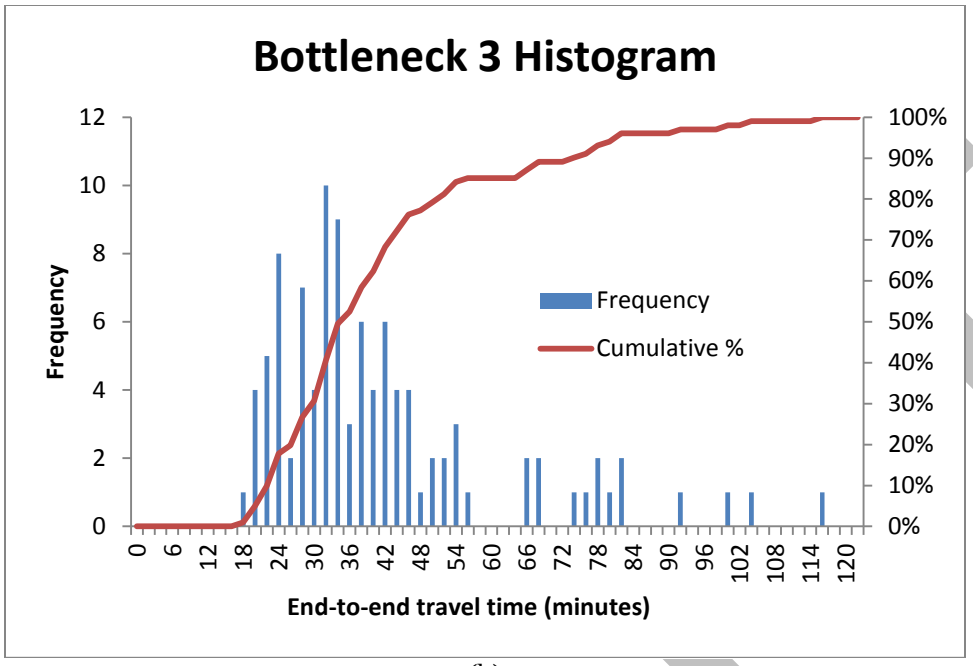
For the same example corridor (with three bottlenecks) in Section 2.3, Monte Carlo experiments were conducted to calculate the end-to-end travel time by assuming that the bottleneck discharge rates, inflow/outflow rates on ramps and existing number of vehicles on the link are all lognormal variables. K = 100 simulation runs were performed with different scenarios of stochastic input parameters.

Figure 7-(a) and (b) shows the distributions of the simulated end-to-end travel times p_2^z and p_3^z for probe vehicle z through bottleneck 2 and through bottleneck 3, respectively. Obviously, the mean travel time based on p_3^z is larger than that of p_2^z. In addition, a clear pattern of randomness propagation can be observed, as p_3^z has higher variance than p_2^z. It should be noted that, by using different input distributions for flow discharge rates and the prevailing number of vehicles on the road, the resulting travel time distributions will vary. This demonstrates the advantage of the proposed model in recognizing the impact of capacity and congestion levels on travel time reliability.



(a)

DRAFT



(b)

Figure 7: End-to-end travel time distribution

4. Model validation using NGSIM data

This section uses vehicle trajectory data available from the NGSIM (Next Generation SIMulation) project (FHWA, 2006) as ground-truth data to verify the proposed methodology and examine the sources of prediction error.

4.1. Data descriptions

The NGSIM vehicle trajectory data used in this study come from the I-80 dataset, which were collected by a video camera located at Emeryville, California. This data collection point is located adjacent to I-80, as shown in Figure 8. The site was approximately 1,650 feet in length, with an on-ramp at Powell Street (indicated in Figure 8 by the red circle). The freeway segment covered in the dataset includes six lanes, numbered incrementally from the left-most lane (HOV lane). Video data are available for three time intervals: 4:00 p.m. to 4:15 p.m., 5:00 p.m. to 5:15 p.m. and 5:15 p.m. to 5:30 p.m., on April 13, 2005. Complete, transcribed vehicle trajectories are available with a time resolution of 0.1 seconds.

DRAFT

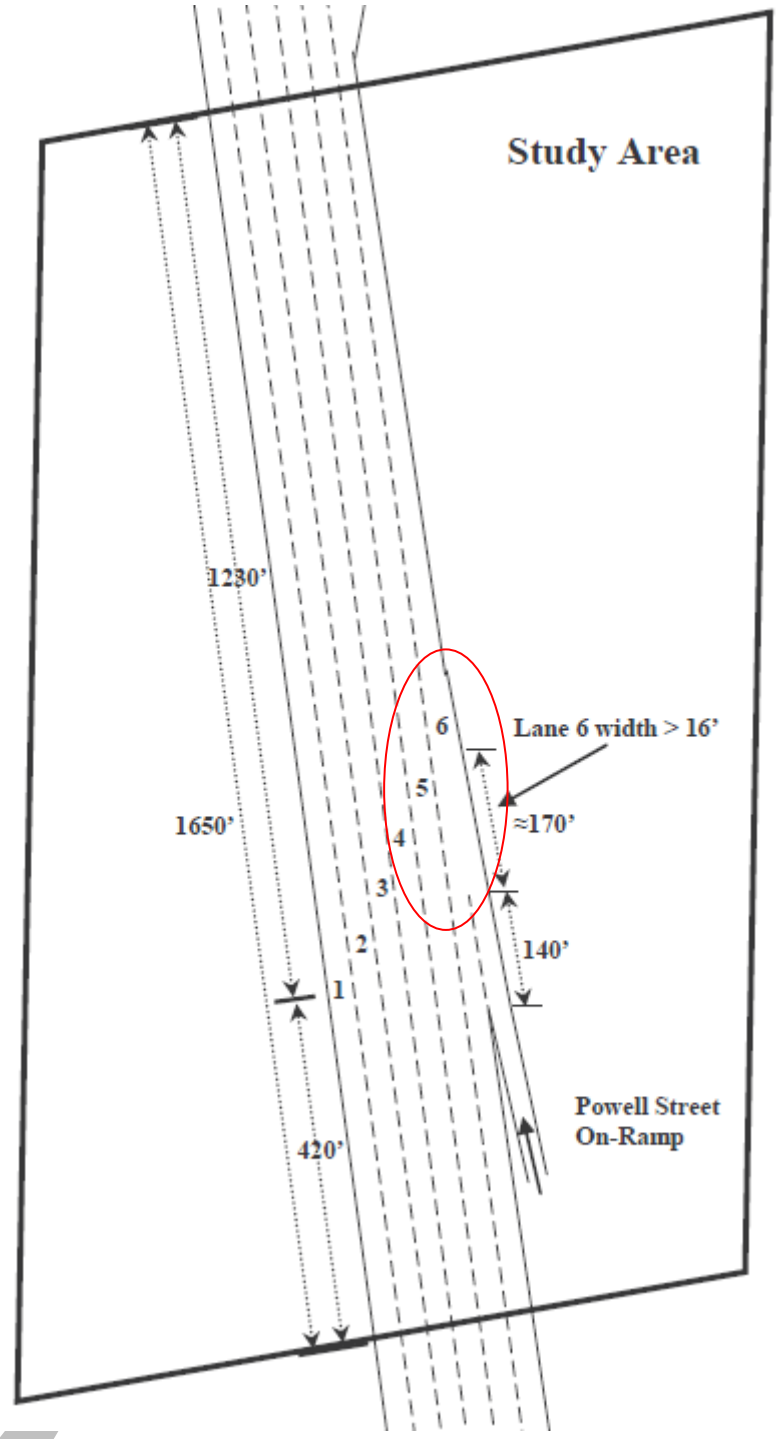


Figure 8: Schematic illustration of NGSIM study area
Source: Adapted from Figure 1, NGSIM I-80 Data Analysis Summary Report (Cambridge Systematics, 2005)

4.2. Data extraction from NGSIM dataset

1. To extract vehicle flow counts data from the NGSIM dataset, we first construct a node-link structure to represent the freeway segment in Figure 8. This stretch of

freeway is divided into two links, as shown in Figure 9, with the on-ramp connected with node 1.

2. In order to obtain the flow rate at the node/bottleneck, this study introduces a set of virtual detectors at node 1 and at node 2, respectively. Meanwhile, another virtual detector is placed on the on-ramp link so that inflow vehicles from the ramp are also counted. In addition, video cameras are assumed to be installed on both links to provide link snapshots (for probe vehicle data).

3. The vehicle trajectory data is divided into 5-minute intervals for counting vehicles. An example of one 5-minute span of vehicle trajectories on one lane is shown in Figure 10 to illustrate how the vehicle counts are collected. As mentioned before, two sets of virtual detectors A and B are placed at nodes 1 and 2 (shown as triangles in Figure 10), and video cameras C and D are also installed on both link 1 and link 2. Vehicles are counted along the vertical line drawn at the given time *t*. At time $t_0 = 0$, probe vehicle *z* = 0 enters link 1. At this time step, two vehicles are observed on link 1 and five vehicles are observed on link 2 by video cameras C and D, that is, $x_1(t_0) = 2$ and $x_2(t_0) = 5$. Similarly, probe vehicle *z* = 5 enters link 1 at time *t*. At this moment, $x_1(t) = 2$ and $x_2(t) = 4$.

   Probe vehicle *z* = 8 is worth mentioning, which enters link 1 at time *t'*. However, at time *t''*, this vehicles leaves this lane to another lane. This is shown in Figure 10 as an incomplete vehicle trajectory. During this 5-minute interval, 12 vehicles are counted by detector A, including two vehicles entering before probe vehicle z = 1, but excluding probe vehicle z = 8. Meanwhile, 13 vehicles are counted by detector B. This count includes seven vehicles before probe vehicle *z* = 1, but excludes probe vehicles 7-11, which have not yet departed from link 2.

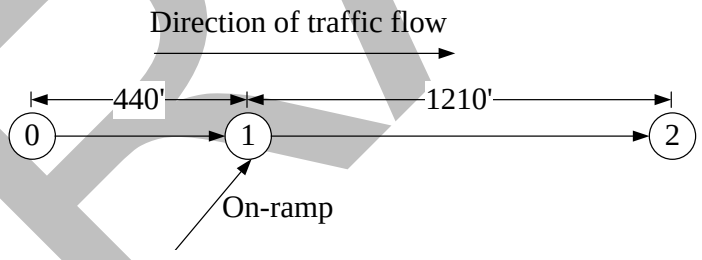


Figure 9: Node-link representation of NGSIM network

DRAFT

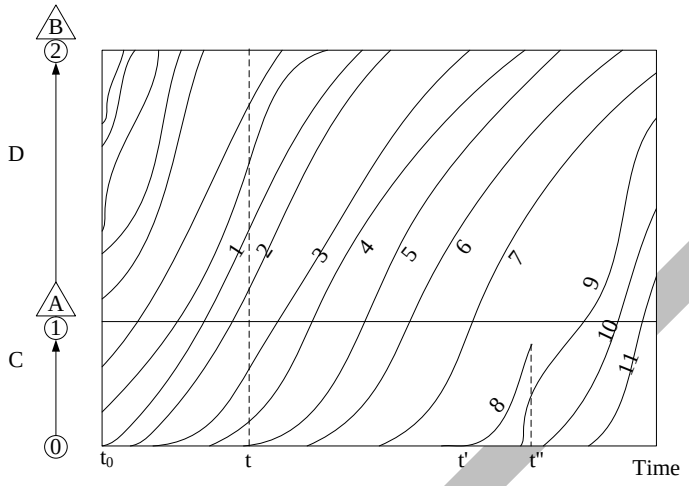


Figure 10: Vehicle trajectory on a lane

4.3. Model validation

Two end-to-end travel time prediction approaches are investigated using our model. The first approach is a lane-based approach, and the second is link-based approach.

Let us first define some notation to represent lane-specific parameters.

Z: number of probe vehicles;

n: index identifying a lane;

t_0^z : starting time for probe vehicle *z*;

n^z : starting lane number for probe vehicle *z*;

$x_m(t,n)$: number of vehicles on lane *n* at time *t*;

$\lambda_m^z(t_m^z,n)$: lane *n* specific number of vehicles behind bottleneck *m*;

$c_m(n)$: lane *n* specific discharge rate of bottleneck *m*;

$f_m^{net}(n)$: net flow rate from or to ramps by lane *n*;

$\theta(n)$: vehicle distribution rate from on-ramp to lane *n*;

$w_m^z(n)$: waiting time for probe vehicle *z* on bottleneck *m* on lane *n*;

$p_m^z(n)$: lane-based end-to-end travel time for probe vehicle *z* through lane *n*;

The following procedure is used to calculate the lane-based travel time distribution.

DRAFT

For *z* = 1 to *Z* on the link

Obtain arrival time t_0^z and starting lane number n^z for each probe vehicle *z*.

Obtain the lane-based number of vehicles $x_m(t_0, \mathrm{n})$;

Obtain the lane-specific discharge rate $c_m(n)$;

Calculate net flow rate $f_m^{net}(n)$ from the on-ramp by applying $\theta(n) \times f_m^{net}$;

Calculate the number vehicles behind bottleneck m $\lambda_m^z(t_m^z, n)$;

Calculate $w_m^z(n)$ based on $c_m(n)$ and $x_m(t_0, n)$;

Update the end-to-end lane travel time $p_m^z(n)$.

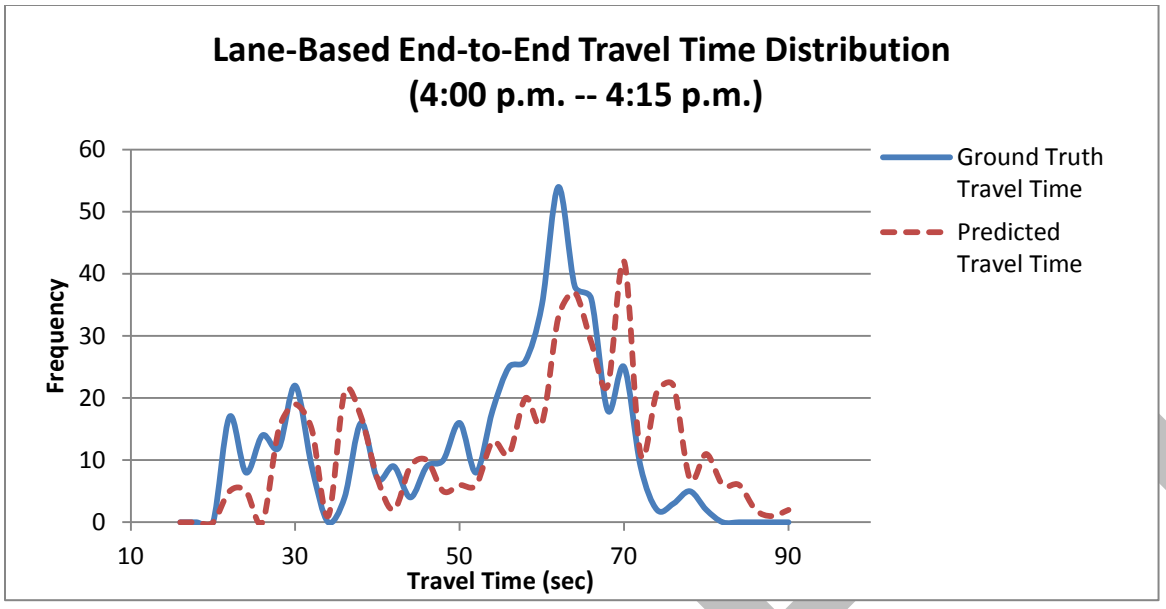
End For

Output: Lane-based path travel time distribution based on $p_m^z(n)$

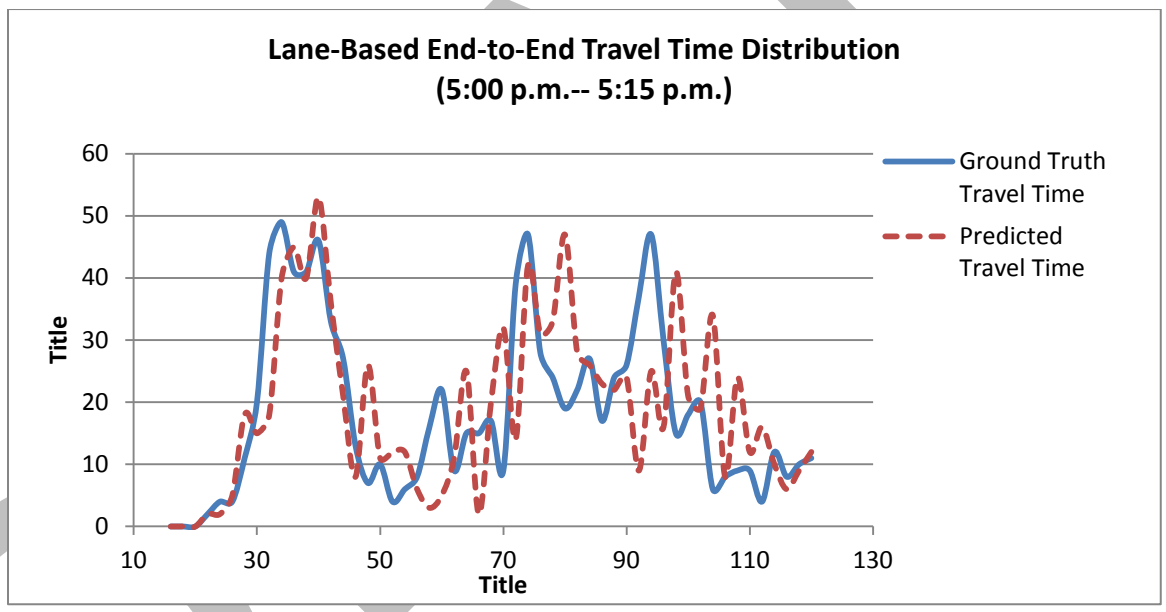
4.3.1. Lane-based end-to-end travel time prediction

The distribution of the predicted end-to-end travel times for each 5-minute interval, calculated over the 3 time periods with available data (4:00 p.m. to 4:15 p.m., 5:00 p.m. to 5:15 p.m. and 5:15 p.m. to 5:30 p.m.) are plotted in Figure 11 with the ground truth end-to-end travel time obtained directly from the NGSIM data. As it can be observed, the distribution of the predicted end-to-end travel time is very close to that of the ground truth end-to-end travel time. This demonstrates that our model is able to accurately predict the end-to-end travel time distribution.

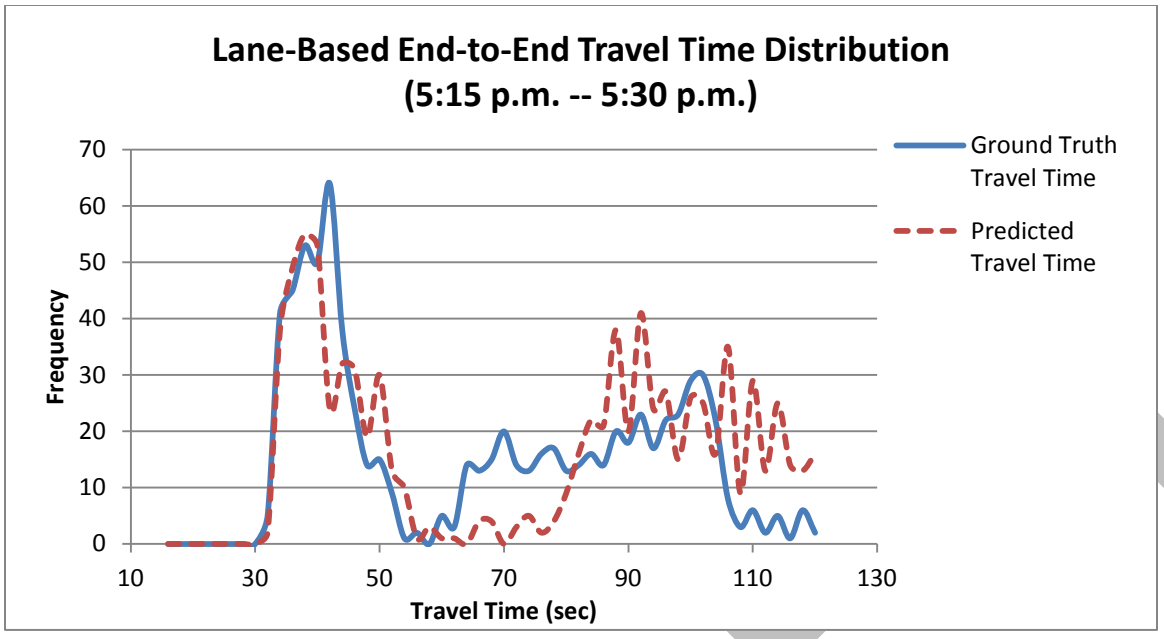
DRAFT


Lane-Based End-to-End Travel Time Distribution
(4:00 p.m. -- 4:15 p.m.)
Frequency
60
50
40
30
20
10
0
10
30
50
70
90
Travel Time (sec)
Ground Truth Travel Time
Predicted Travel Time

(a)


Lane-Based End-to-End Travel Time Distribution
(5:00 p.m.-- 5:15 p.m.)
Title
60
50
40
30
20
10
0
10
30
50
70
90
110
130
Title
Ground Truth Travel Time
Predicted Travel Time

(b)

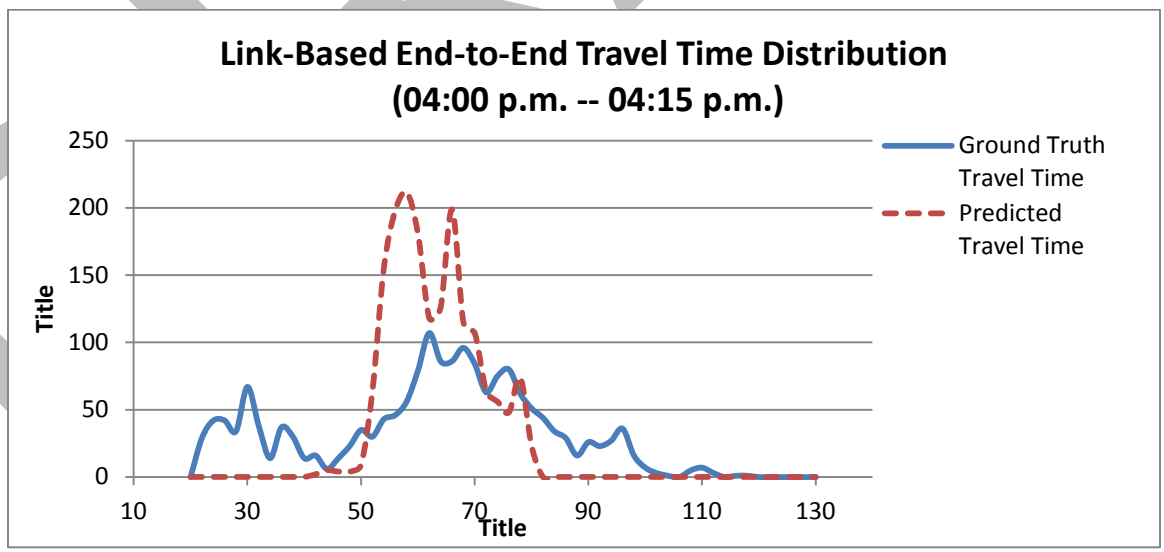


(c)

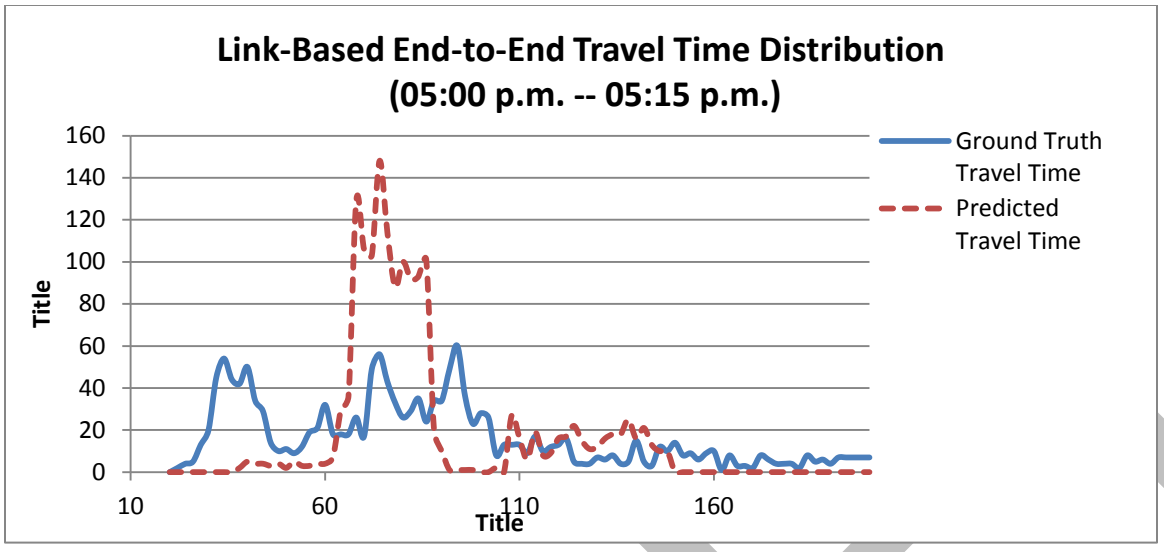
Figure 11: Lane-based end-to-end travel time distributions

4.3.2. Link-based end-to-end travel time prediction

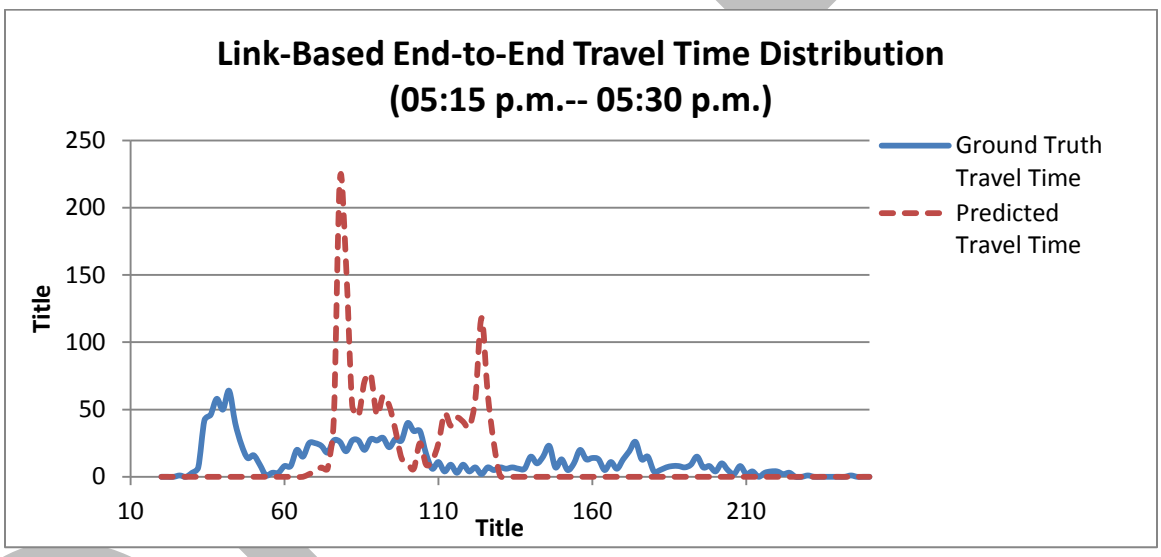
One common practice is to use link-based flow rates or density to predict travel time variability. We replace the lane-based variables in the previous approach with link-based variables $x_m(t_0)$ and c_m. That is, $x_m(t_0)$ is the existing number of vehicles on all the lanes on the link and c_m is the link discharge rate. The distribution of the predicted end-to-end travel time and true end-to-end travel time for different time intervals are shown in Figure 12.



(a)



(b)

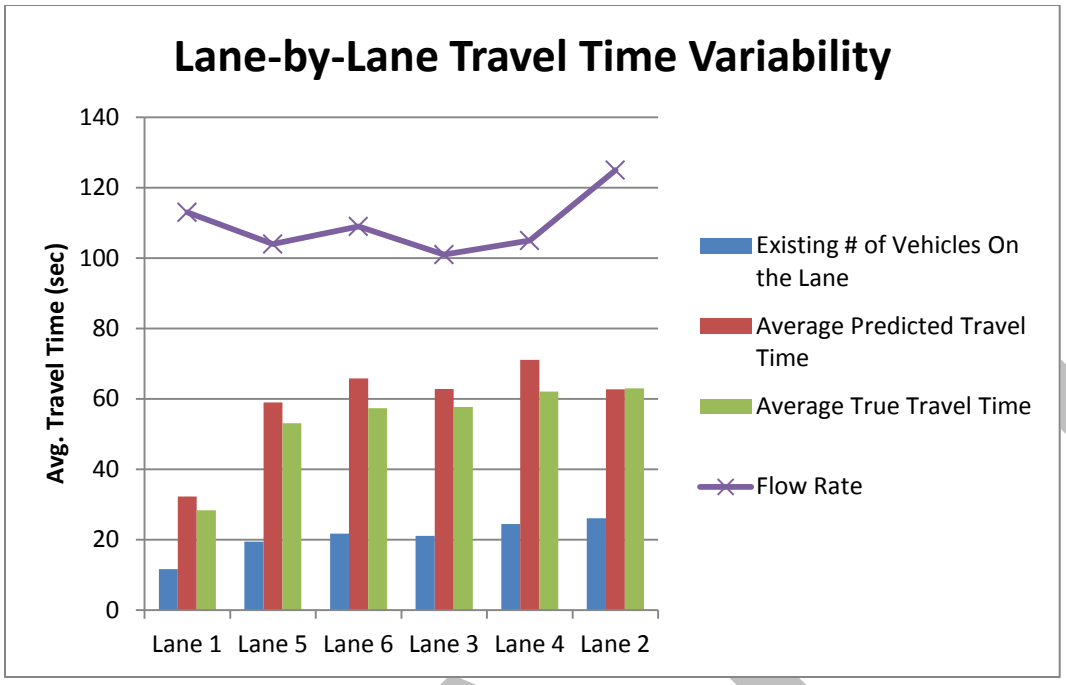


(c)

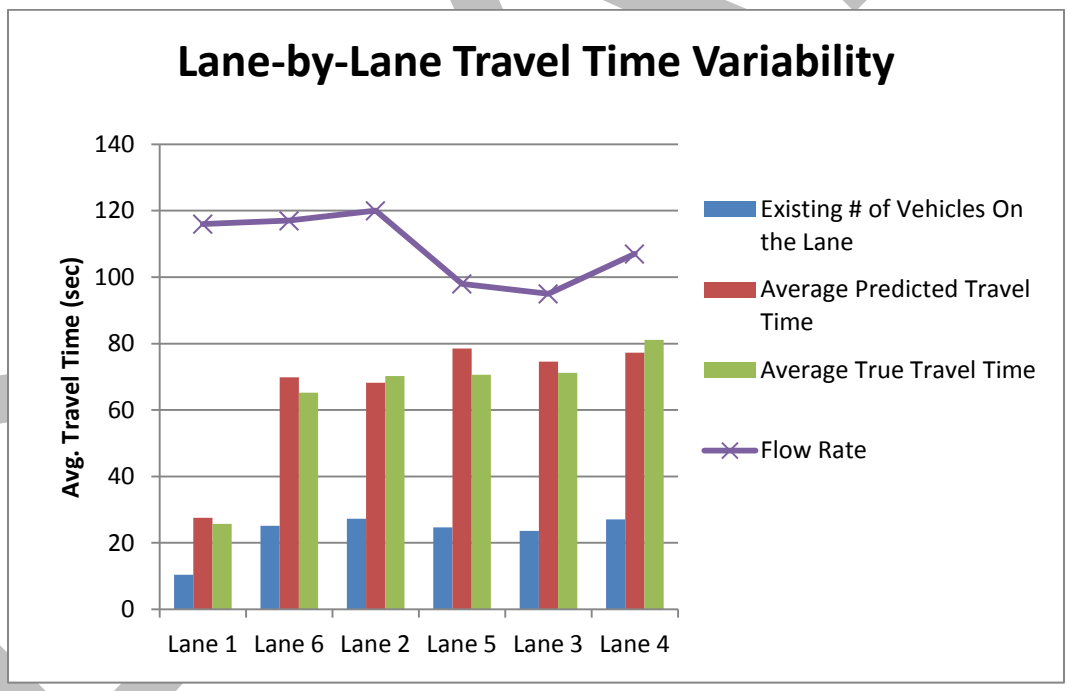
Figure 12: Link-based end-to-end travel time distributions

As it can be observed, the distribution of the predicted link-based end-to-end travel times fails to capture the wide-spread distribution in the ground truth travel times. This is explained by the fact that link-based input variables would yield the same predicted end-to-end travel times for those vehicles entering the link at the same time, regardless of their driving lanes.

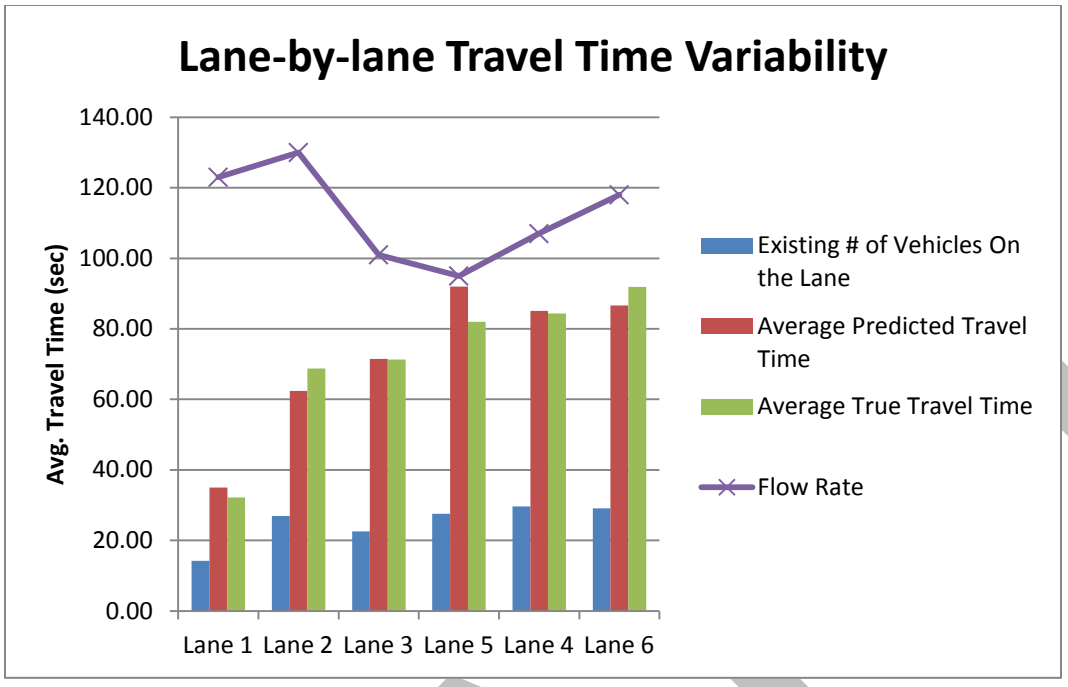
In order to understand the extent and sources of the lane-by-lane travel time variation, we use the time period between 4:00 p.m. and 4:15 p.m. as an example. Figure 13 shows the lane discharge rate, the existing numbers of vehicles on the lane, and the average true and predicted end-to-end travel times for each lane for each 5-min interval in the time period. The lane sequence is sorted by the true end-to-end travel time.



(a): 4:00 p.m. – 4:05 p.m.



(b): 4:05 p.m. – 4:10 p.m.



(c): 4:10 p.m. – 4:15 p.m.

Figure 13: Lane-by-lane travel time variability

Several observations can be obtained from Figure 13. Lane 1 (HOV lane) has the lowest existing number of vehicles on the lane and has the lowest average end-to-end travel time. The left-most lanes (lanes 1 and 2) usually have the highest discharge rates while lanes 3 and 4 usually have the lowest discharge rates. In most cases, lanes 3 and 4 have the highest average existing numbers of vehicles on the lane, as well as the highest average end-to-end travel times.

These observations imply that, due to the variation of the discharge rates and the number of vehicles on the lane, the end-to-end travel times also present strong lane-by-lane variations. As a result, we suggest using lane-based statistics to better quantify the travel time variability.

4.3.3. Prediction error sources

By comparing the predicted results with the NGSIM ground truth data, we can further uncover other possible sources of errors in the proposed travel time prediction model.

1. Aggregation errors: The link/lane discharge rates c_m and on-ramp flow rates used in the calculations are average flow rates over a certain time interval, e.g., 5-minute rates, while the existing number of vehicles on the link/lane $x_m(t_0)$ is an instantaneous value based on the entering time of a probe vehicle.
2. Measurement errors: The number of vehicles on the lane observed by the video camera at time t_0 is assumed to be error-free. In fact, there are always vehicle detection errors in NGSIM vehicle trajectory data associated with the underlying video recognition algorithm.
3. Modeling errors associated with lane changing: Since the queue model incorporates the first-in-first-out principle, lane change behavior is not considered in the calculation. This will introduce two types of errors in the model:

DRAFT

a. The model may underestimate or overestimate the number of vehicles behind the bottleneck, $\lambda_m\left(t_m^z\right)$. For example, some vehicles will enter the lane (from the other lanes) before a probe vehicle reaches the bottleneck, or some vehicles originally counted in $x_m\left(t_0\right)$ on the current lane will leave to one of the adjacent lanes, corresponding to a lower value of $x_m\left(t_0\right)$.

b. When a probe vehicle changes lane from, for example, $n^z=1$ to lane n', the discharge rate used in the calculation should be changed to the one associated with lane n'.

5. Conclusions

In this paper, a travel time estimation framework is proposed to calculate the end-to-end travel time, given the number of vehicles along the corridor, the bottleneck discharge rates and the on-ramp/off-ramp flow rates. Monte Carlo methods and mathematical approximation methods are presented to calculate the travel time distribution if the distributions of the variables are also given. Ground-truth data from the NGSIM project are applied for model validations. Several critical observations are obtained from the model validations:

(1) Our model is able to provide a analytical to predict travel time and its distribution;
(2) Using more detailed inputs, e.g., lane-level rather than link-level, could yield more accurate end-to-end travel time distribution predictions; and
(3) The variation of lane-specific parameters (the number of vehicles on the lanes and lane discharge rates) results in lane-by-lane travel time variation.

Nevertheless, the impacts of downstream queue spillback on the upstream travel time have not yet been considered in our model. In queue spillback, the discharge rate of the upstream bottleneck is constrained by the downstream bottleneck discharge rate. Moreover, it is also desirable to consider the influence of lane changes on the predicted number of vehicles waiting in the queue. These important features will be investigated more thoroughly in future research.

DRAFT

References

1. Cambridge Systematics, Inc. NGSIM BHL Data Analysis. Technical report, September 2004. Summary Report, Prepared for Federal Highway Administration.
2. Cambridge Systematics, Inc. NGSIM I-80 Data Analysis (4:00 p.m. to 4:15 p.m.). Technical report, September 2005. Summary Report, Prepared for Federal Highway Administration.
3. Cambridge Systematics, Inc. NGSIM I-80 Data Analysis (5:00 p.m. to 5:15 p.m.). Technical report, September 2005. Summary Report, Prepared for Federal Highway Administration.
4. Cambridge Systematics, Inc. NGSIM I-80 Data Analysis (5:15 p.m. to 5:30 p.m.). Technical report, September 2005. Summary Report, Prepared for Federal Highway Administration.
5. Chen, C., Skabardonis, A., and Varaiya, P. (2003) Travel-time Reliability as a Measure of Service. *Transportation Research Record*, 1855, 74-79.
6. Clark, S. and Watling, D. (2005) Modeling Network Travel Time Reliability under Stochastic Demand. *Transportation Research Part B*, 39 2 (2005), pp. 119–140.
7. Cowan, R. J. (1975) Useful Headway Models. *Transportation Research*, Vol. 9, No. 6, 1975, pp. 371–375.
8. Daganzo, C. F. 1995. Properties of link travel times under dynamic load. *Transportation Research*, Vol. 29B: 95-98.
9. Federal Highway Administration. Next Generation SIMulation Fact Sheet. Technical report, December 2006. FHWA-HRT-06-135.
10. Federal Highway Administration. Recurring Traffic Bottlenecks: A Primer Focus on Low-Cost Operational Improvements. Technical report, June 2009. FHWA-HOP-09-037.
11. Giuliano, G. (1989) Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway. *Transportation Research A* 23:387–396.
12. Golob, T. F., Recker, W.W. and Leonard, J. D. (1987) An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents. *Accident Analysis and Prevention*, 19 (5), pp. 375–395.
13. Greenberg, I. (1966) The Log-normal Distribution of Headways. *Australian Road Research*, Vol. 2, No. 7, pp. 14-18.
14. Guo, F., Rakha, H., and Park, S. (2010) Multistate Model for Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2188, pp. 46-54.
15. Hazelton, M. L. (2001) Inference for Origin–Destination Matrices: Estimation, Prediction and Reconstruction. *Transportation Research*, Vol. 35B, No. 7, 2001, pp. 667–676.
16. Kripalani, A. and Scherer, W. (2007). Estimating Incident Related Congestion on Freeways Based on Incident Severity. July, 2007. Research Report No. UVACTS-15-0-113
17. Kwon, J., Barkley, T., Hranac, R., Petty, K., and Compin, N. (2011) Decomposition of Travel Time Reliability into Various Sources. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2229, 2011, pp. 28-33.
18. Lam, W.H.K., Shao, H. and Sumalee, A. (2008) Modeling Impacts of Adverse Weather Conditions on a Road Network with Uncertainties in Demand and Supply. *Transportation Research Part B*, 42 10, pp. 890–910.
19. Lo, H.K. and Tung, Y.K. (2003) Network with degradable links: capacity analysis and design. *Transportation Research Part B*, 37 4, pp. 345–363.
20. Merchant, D.K. and Nemhauser, G.L. (1978). A Model and an Algorithm for the Dynamic Traffic Assignment Problem. *Transportation Science*, Vol. 12, pp. 183-199

DRAFT

21. Merchant, D.K. and Nemhauser, G.L. (1978). Optimality Conditions for a Dynamic Traffic Assignment Model. *Transportation Science*, Vol. 12, pp. 200-207
22. Ng, M.W., Waller, S.T. (2010). A Computationally Efficient Methodology to Characterize Travel Time Reliability using the Fast Fourier Transform, *Transportation Research Part B* 44 10, 1202-1219.
23. Oh, Jun-Seok and Chung, Younshik (2006). Calculation of Freeway Travel Time Variability. *Transportation Research Record: Journal of Transportation Research*, No. 1945, pp. 12 – 23
24. TRB (2000). *Highway Capacity Manual*. Transportation Research Board, National Research Council, Washington, D.C., U.S.A. ("HCM 2000").
25. Richardson, A.J. (2003). Travel Time Variability on an Urban Freeway. TUTI Report 25-2003, presented at the *25th Conference of Australian Institutes of Transport Research (CAITR)*, University of Adelaide, December 2003
26. Sundaram, S., Koutsopoulos, N.H., Ben-Akiva, M., Antoniou C.,Balakrishna, R. (2011). Simulation-based dynamic traffic assignment for short-term planning applications. *Simulation Modelling Practice and Theory*, 19, pp. 450–462
27. Waller, S.T. and. Ziliaskopoulos, A.K. (2001). Stochastic dynamic network design problem. *Transportation Research Record*, 1771, pp. 106–113.
28. Unnikrishnan, A., Ukkusuri, S., and Waller, S. T. (2005) Sampling methods for evaluating the traffic equilibrium problem under uncertain demand. *Transportation Research Board 84th Annual Meeting Compendium of Papers*, Washington, DC.
29. Zhou, X., Rouphail, N. M., and Li, M. (2011) Analytical Models for Quantifying Travel Time Variability Based on Stochastic Capacity and Demand Distributions. Presented at *Transportation Research Board 90th Annual Meeting*, Washington, D.C., 2011

DRAFT