Improved Travel Time Estimation for Reliable Performance Measure Development for Closed Highways

Hong Yang, Ph.D.
(Corresponding author)
Post-doctoral Associate, Department of Civil & Urban Engineering,
Polytechnic School of Engineering;
Center for Urban Science + Progress (CUSP);
New York University (NYU)
One MetroTech Center, 19th Floor,
Brooklyn, NY 11201, USA
Phone: (646) 997-0548, Fax: (646) 997-0560
E-mail: hong.yang@nyu.edu

Kaan Ozbay, Ph.D.
Professor, Department of Civil & Urban Engineering,
Polytechnic School of Engineering;
Center for Urban Science + Progress (CUSP);
New York University (NYU)
One MetroTech Center, 19th Floor,
Brooklyn, NY 11201, USA
Phone: (646) 997-0552, Fax: (646) 997-0560
E-mail: kaan.ozbay@nyu.edu

Kun Xie, M.Sc.
Ph.D. Candidate, Department of Civil & Urban Engineering,
Polytechnic School of Engineering;
Center for Urban Science + Progress (CUSP);
New York University (NYU)
One MetroTech Center, 19th Floor,
Brooklyn, NY 11201, USA
Phone: (646) 997-0547, Fax: (646) 997-0560
E-mail: kun.xie@nyu.edu

Word count: 5747 Texts + 4 Table + 3 Figures = 7497
Abstract: 207
Re-submission Date: November 15, 2014

Paper submitted for Publication in the
Transportation Research Record, Journal of Transportation Research Board after being presented
Transportation Research Board’s 94th Annual Meeting, Washington, D.C., 2015
ABSTRACT

Accurate travel time information is not only valuable for travelers, but also critical for transportation agencies quantifying the performance of their systems. Thus, increasing interest has been devoted to develop reliable approaches for estimating travel time from various sensor data. Unlike the extensively studied estimation approaches based on point sensor measurements, experience on using probe data from closed highway systems is relatively limited. To complement existing understanding, this study aims to develop an approach that estimates travel time based on probe data from an electronic toll collection (ETC) system on closed freeways. This is different from the studies relying on automatic vehicle identification systems deployed on mainlines as well as those estimated based on point detectors. The proposed approach breaks down individual journey time into section travel time and fuses the probe data from vehicles that have used the links. The results based on real-world case studies illustrate the potential of mining the ETC data for travel time estimation under both incident-free and incident conditions. In addition, the estimated results are shown to outperform the instantaneous travel time estimates based on point sensor data in terms of capturing traffic dynamics. This in turn provides more accurate information to derive reliable performance measures to depict travel time reliability.
INTRODUCTION
The advent of Intelligent Transportation Systems (ITS) enables the acquisition of various sensor data in transportation systems. A wide variety of new technologies—from probe-based data captured by freely roaming cars and buses to electronic payment systems that can nearly capture every detail of a car, bus or train commuter’s trip—provide us new opportunities for improved monitoring and management of our complex transportation systems. Thanks in large part to the sensor data being collected on a nearly 24/7 basis, researchers and transportation practitioners alike are now able to quantify a number of realistic “performance measures” that allow transportation agencies to make timely and proactive decisions. One of the major categories of performance measures such as congestion and reliability to evaluate transportation systems relies on the travel time. Thus, the accuracy of travel time information is the key to derive reliable performance measures for appropriate decision makings.

Because of the increasing needs for accurate travel time information, a large volume of studies have been attempting to determine travel time by mining traffic data from sensors such as inductive loop detectors embedded in road surfaces, automatic vehicle identification systems, Bluetooth, Global Positioning Systems (GPS), etc. Since many existing freeways are already equipped loop detectors, one of the primary approaches was naturally based on the measurements from these detectors. These point detectors usually measure traffic volume, occupancy and/or speed, thus offer opportunities to estimate travel time based on some statistical approaches as well as traffic flow theory. A number of estimation methods, algorithms and data fusing techniques have been developed and some promising results have been achieved. Nevertheless, the accuracy of the estimates is still frequently questioned because the point measurements hardly reflect the traffic dynamics over the links. For instance, many approaches cannot capture the impact of recurrent congestions/ incidents occurred between two consecutive detectors. Despite the improvement through various techniques such as vehicle identification, queuing analysis, and advanced data processing, the practical use of the approaches is still challenge when complicated procedure/models were adopted.

Instead of relying on point sensor-based approaches, there is increasing interest in applying probe vehicle data for travel time estimation. Compared with point sensor data, the probe vehicle data often directly capture the time stamps when vehicles travel across different locations. Thus, they provide more reliable information on individual vehicles’ actual travel time. Lessons learned from a limited number of practices imply that it is highly desirable to find better estimation techniques. Issues such as sample size, locations of readers, and data fusing algorithms still need more investigations, compared to the other conventional detector-based approaches.

Thus, the primary objective of this paper is to develop a travel time estimation method based on probe data. Specifically, it focused on the closed freeway systems with electronic toll collection facilities on ramps. A detailed estimation approach has been presented and its performance as well as applications have been illustrated through case studies.

LITERATURE REVIEW
Numerous research projects have been devoted to estimate highway travel times using various sensor data. One of the major categories of travel time estimation efforts was dedicated to mine the spot measurements from inductive (single and dual) loop detectors. For instance, earlier studies (1-4) focused on estimating speed based on the single-trap volume and occupancy measurements, and hence travel time. The estimated travel time represents traffic conditions over the links in a fixed time period, it hardly reflects the traffic variations, and thus the actual travel time usually differs significantly from the quotient of spot speed and the link length (5). Another category of research utilizes the measurements from dual loop detectors to re-identification vehicles and construct their pseudo vehicle trajectories for travel time estimation (6-9). It derives imaginary vehicle trajectories based on observed spot speeds or estimated section level travel times. Coifman (6) conclude that such trajectory based method work better in non-congested periods. To improve the performance of conventional trajectory based approach, a number of studies extended it and proposed the piecewise constant speed based method (10), piecewise linear speed based method (11), piecewise truncated quadratic speed based method (12), filtered inverse speed based (13), disturbance wave speed
based method (14), temporal-spatial queuing model (15), etc. Empirical studies have shown that these extended approaches can improve the estimation in terms of reduces bias and residual error (13), (partially) accounting for congestion impact (15), etc. Instead of constructing vehicle trajectory, Lucas et al. (16) developed an estimation by tracking the platoons based on aggregated measures from upstream and downstream detectors, which was shown to better capture the prevailing congestion.

Despite the popularity, a number studies still concerned about the performance of travel time estimation based on spot measurements. For example, Li et al. (17) assessed four spot speed-based travel time estimation models and concluded that all of them underestimate the actual travel time. These models did not account for the potential relationship between section length and estimation errors. Likewise, Soriguera and Robuste (18) also suggested that those estimation approaches relying on speed interpolations failed to provide better estimation because they did not account for traffic dynamics and queue evolution. Regarding the potential errors of estimation, researchers have implemented a number of countermeasures to tackle the issues. For instance, Blue et al. (19) discussed the use of neural networks to improve estimation on bottleneck sections with recurrent congestions. Similarly, Wen et al. (20) used the gray-based recurrent neural network to treat missing data and fusing data to estimate travel time. Chu et al. (21) applied adaptive Kalman filter in estimation by fusing both erratic sensor data and probe vehicle data. Yeon et al. (22) used the Discrete Time Markov Chains to address the underestimation of travel time due to inaccurately captured congestions states. Other than improving estimation approaches, a few studies focused on the issues of data collection. Specifically, a majority of them (23-26) concentrated on the optimization of sensor placement for better travel time estimation. These studies contributed a number of useful suggestions to advise the configuration of sensors.

Instead of using indirect measurements to estimate travel time, there are also a few studies that contributed to the use of probe data for travel time estimation. These probe data generally come from Automatic Vehicle Identification (AVI) systems such as tolling data, license plate matching, and media access control (MAC) address matching. For instance, Tam and Lam (27, 28) used the AVI data to estimate travel time in Hong Kong. Ozbay and Yildirimoglu (29, 30) and Soriguera et al. (31) used vehicular toll transaction records as probes to estimate freeway travel time. Wasson et al. (32) examined the potential of using media access control (MAC) address from Wifi and Bluetooth-enabled electronic devices. These approaches essentially measure the spatial travel time based the fixed location equipment to recognize and track (a portion of) vehicles in the traffic flow. Other than these approaches, some studies also explored more detailed probe data from cellular phones (33) and global positioning systems (GPS) (34) from vehicles traveling along the highways. Although the estimation based probe data provide more reliable information to capture the actual travel times for sample vehicles, it is still have some challenge issues to be resolved. For instance, what is the reasonable number of probes required for reliable estimation of travel time? Researchers such as Chen and Chien (35), Cetin et al. (36) and Dion and Rakha (37) have provided some insight into such question but more thoughts are expected. Another question as discussed by Sherali et al. (38) is where to locate the AVI readers as the data collection resources are usually limited and expensive. Moreover, it is also necessary to consider the optimal data aggregation interval size for different levels (link, corridor, etc.) of estimation given the individual samples of probes (39).

Under the context of AVI-based estimation, an increasing interest is to mine the data from electronic toll collection (ETC) systems in a closed tollway systems. The toll plazas of such system are usually located on the on- and off-ramp rather than the mainline (29). For vehicles without toll tag, they usually receive a ticket at the entrance of a toll plaza and return it with tolls at another exit toll plaza. Despite these manual payment users still have the transaction time stamps, they are often excluded from consideration due to their delays in transaction. For vehicles with toll tag (i.e., E-ZPass in east coast of U.S.), they will be detected at the origin and destination toll plazas. The difference of the time stamps naturally offers the opportunities to deduce these vehicles’ travel time. Many highways in the nation have a very high penetration of ETC uses, for instance, the penetration for passenger car traffic was 78.1% and for commercial traffic was 86.5% for New Jersey Turnpike (40). Thus, the large volume of vehicles serving as probes provide a very valuable data source for travel time estimation. As such, several studies investigate those transaction data for travel time estimation (29, 30, 41-43). Unlike the extensively studied point sensor-
based methods, the use of the toll collection data from closed highway systems remains understudied. The existing studies were still exploring the data filtering and fusing procedures. There are still very limited lessons about the travel time estimation using massive ETC data and their performance against other sensor-based approaches. These issues motivate us to further tackle the questions.

**PROBLEM STATEMENT**

As mentioned earlier, travel time is a fundamental element used to derive a number of performance measures. AVI systems record the exact times at which vehicles enter and exit the network. The EZ-Pass toll collection systems on the New Jersey Turnpike, Pennsylvania Turnpike, and New York Thruway are examples of such AVI systems operating in a closed highway. So far, 15 states, including NC, VA, WV, MD, DE, NJ, IL, IN, OH, PA, NY, MA, RI, NH and ME, are using EZ-Pass system. These systems provide very rich traffic information including traffic volume, traffic classifications, time stamps, type of payment, and tolls. One important yet understudied topic is how to better utilize the large amount of transaction data generated from the ETC systems for the sake of travel time estimation.

Since the ETC systems capture enter and exit time stamps, the experienced travel time of a vehicle completing its trip between two toll plazas is directly accessible. However, there are several challenges associated with the use of the measured travel time as an estimation of the trip travel time. The most notable challenge is that the traffic demand between two toll plazas might be too low to develop a reliable estimation of the travel time. For example, the demand between two consecutive junctions can be extremely low, since travelers do not prefer using tolling roads for short-distance trips. Furthermore, an estimation of travel time for a section based on measurements from a single origin-destination (OD) pair would be biased by the experience of those vehicles. Travel time estimation should incorporate information of other through traffic as well. Therefore, a key question is how to efficiently utilize EZ-Pass data to accurately estimate link travel times.

**PROPOSED METHODOLOGIES**

This section provides a basic understanding of the estimation of travel time using EZ-Pass vehicles as probes. It develops improved models to estimate experienced travel time for closed tolling networks.

**Travel Time Estimation Methodology**

The travel time between an entry toll plaza \( i \) and an exit toll plaza \( j \) for vehicles that access \( i \) at \( k^{th} \) time interval, as shown in FIGURE 1, can be simply represented by the following equation:

\[
t_{jk} = t_{ik} + \Delta t_{jk} + t_{jk}
\]

(1)

where \( t_{ik} \) is the travel time calculated based on the entry time and exit time; \( t_{ik} \) is the entry travel time for vehicles to reach the merge area on mainline; \( t_{jk} \) is the exit travel time for vehicles to leave the exit toll plaza; \( \Delta t_{jk} \) represents the travel time for the section between the on-ramp merge area of the entry toll plaza \( i \) and the off-ramp diverge area of the exit toll plaza \( j \).

---

**FIGURE 1**

Travel time of link \( S_{ij} \) at \( k^{th} \) time interval: \( (S_{ik}, t_{ik}) \)

On-ramp time: \( t_{ik} \)

Off-ramp time: \( t_{jk} \)

Toll Plaza i

Toll Plaza j
As shown in FIGURE 1, the link travel time $t_{Sijk}$ of the mainline section between entry toll plaza $i$ and exit toll plaza $j$ for vehicles entering at $k^{th}$ time interval is consisted of two part: (a) the travel time $\Delta t_{Oik}$ for the short segment between the off-ramp diverge area and the on-ramp merge area of toll plaza $i$ and (b) the travel time $\Delta t_{ijk}$ for the long segment between the on-ramp merge area of the entry toll plaza $i$ and the off-ramp diverge area of the exit toll plaza $j$. $t_{Sijk}$ can be represented by the following equation:

$$t_{Sijk} = \Delta t_{Oik} + \Delta t_{ijk} + \left(t_{ijk} - t_{Oik} - t_{Djk}\right)$$

(2)

Since there is no direct measurement for the travel time $\Delta t_{Oik}$ of the short segment within an interchange, a linear expansion factor is assumed based on the length of the segments. The estimation of $\Delta t_{Oik}$ is described as follows:

$$\Delta t_{Oik} = \Delta t_{ijk} \times \frac{S_{Oi}}{S_{Oi} + S_{Bij}} = \left(t_{ijk} - t_{Oik} - t_{Djk}\right) \times \frac{S_{Oi}}{S_{Oi} + S_{Bij}}$$

(3)

where $S_{Oi}$ is the length of the segment between the off-ramp diverge area and the on-ramp merge area of toll plaza $i$; $S_{Bij}$ is the length of the segment between the on-ramp merge area of the entry toll plaza $i$ and the off-ramp diverge area of the exit toll plaza $j$.

Based on equation (3), the link travel time $t_{Sijk}$ of the mainline section between an entry toll plaza $i$ and exit toll plaza $j$ at $k^{th}$ time interval is shown and rewritten in equation (2) as

$$t_{Sijk} = \left(t_{ijk} - t_{Oik} - t_{Djk}\right) \times \left(1 + \frac{S_{Oi}}{S_{Oi} + S_{Bij}}\right)$$

(4)

As the length of each segment is a constant value, only information about the travel time between two toll plazas and the corresponding on-ramp travel time and off-ramp travel time is needed to estimate the link travel time $t_{Sijk}$ in equation (4).

The travel time between a toll plaza $i$ and toll $j$ can be simply estimated by the above equation (4). However, the results might not be representative for the target link between $i$ and $j$. This is because the demand between toll plaza $i$ and toll plaza $j$ might be very low for certain periods, particularly at night time. In addition, samples can be very low if $i$ and $j$ are very close and few vehicles prefer to pay to use the link for such a short trip. As mentioned before, it is desirable to combine travel time information from vehicles associated with other OD pairs that cross the target link. Therefore, a data fusion procedure is developed for mining the probe data. The following FIGURE 2 is used to illustrate the proposed data fusion procedure.

FIGURE 2 Estimate target link travel time between $i$ and $j$ based on other adjacent OD pairs

Assuming that the objective is to estimate travel time $t_{Sijk}$ for link $ij$ at the $k^{th}$ time step. Using the aforementioned equations we can easily obtain a rough estimation of travel time for any OD pair that crosses
the target link and have \(j\) as the destination. For instance, assuming \(N\) toll plazas at the upstream toll plaza \(i\) are considered. The travel times between an upstream origin toll plaza \(i\) \(\in\{1,2,\ldots,N\}\) and toll plaza \(i\) and toll plaza \(j\) measured based on equation (4) are denoted as \(t_{i(n)jk}\) and \(t_{i(i-x)jk}\), respectively. Thus the travel time of the target link \(ij\) at the \(k^{th}\) time step can be represented by equation (5).

\[
\hat{t}_{ij} = t_{i(1)jk_1} - t_{i(1)jk_2} = t_{i(2)jk_2} - t_{i(2)jk_2} = \ldots = t_{i(n)jk_n} - t_{i(n)jk_n} \quad (5)
\]

Since there will be information delays due to vehicles traverses from one link to another, equation (6) is used to synchronize the vehicles that enter the target link at the same time interval.

\[
k = k_1 + t_{i(1)jk_1} = k_2 + t_{i(2)jk_2} = \ldots = k_n + t_{i(n)jk_n} \quad (6)
\]

Once the travel time for the target link \(ij\) is derived based on equation (5) and equation (6), the final estimated travel time of \(\hat{t}_{ijk}\) can be obtained with equation (7).

\[
\hat{t}_{ijk} = \frac{1}{N} \sum_{n=1}^{N} (t_{i(n)jk_n} - t_{i(n)jk_n}) \quad (7)
\]

Note that this equation assumes the mean value of the derived travel time \(t_{ijk}\) from different OD pairs as the final estimation. Instead of using the mean value, the median value can be another option as the final estimation of \(\hat{t}_{ijk}\). The estimated travel time \(\hat{t}_{ijk}\) can be compared with the one that is directly derived based on equation (4) using the ETC data between toll plaza \(i\) and toll plaza \(j\).

Data Fusion Algorithm and Discussion
When implementing the proposed approach for travel time estimation, there are still two important questions that need to be answered. With equation (4) in mind, the first question is how to obtain the entry travel time \(t_{iak}\) and exit travel time \(t_{ja}\). The second question is how to filter and aggregate the ETC measurements of multiple vehicles to obtain a representative travel time \(t_{ja}\) for each time interval.

One way to obtain the entry and exit travel times is to assume that there is a constant travel time given the fixed distance (between toll plaza and mainline) and the corresponding speed limits. However, despite the relatively short distances between entry and exit points, the entry travel time \(t_{iak}\) from toll booths to mainline is closely related to the traffic conditions at the entry ramp as well as the mainline. Similarly, the exit travel time \(t_{ja}\) is more related to the traffic conditions at the exit ramp and the queues in the approach zone of the exit toll booths. Depending on the traffic conditions, these two travel time values can vary. For instance, an assumption of constant entrance/exit times would not work when there is congestion along the mainline/exit toll plaza. For this type of highway configurations, congestion along the mainline would block a vehicle’s access to the mainline whereas the congested exit toll plaza would delay the vehicle from leaving the tollbooth. Therefore, the constant travel time values calculated from the entry or exit distances and the posted speed limits cannot always accurately represent the actual traffic conditions, particularly during peak periods. In instances where the historical travel time distributions for the entry section and the exit section are available, random samples of the corresponding time periods should be used instead of the assumed constant travel time values.

To obtain the representative travel time for each time interval based on observed ETC data, unrealistic observations (outliers) have to be eliminated. These individual outlier observations might be due to vehicles stopping in rest areas or various other unforeseen factors. In addition, for the periods with very few or no observations, interpolation should be made based on other available information. In other words, travel times from the previous period and the free flow travel time can be considered when performing an interpolation of travel time for those periods with missing/limited data. TABLE 1 shows an example of the algorithm used to filter and aggregate the observed travel time for obtaining the representative travel time. Some of the key parameters such as \(\alpha_i\) and \(\alpha_z\) need to be determined based on the users preference. For
instance, $\alpha_i$ can be large if the information at previous time step are more important. $\alpha_i$ defines the range
that the current observed travel compared to the travel time at previous time step.

**TABLE 1 Pseudo code for fusing travel time measurements**

At time interval $k$:

**Assumption:** there were $N$ ETC travel time observations $t_{nk}^{\text{obs}}$ ($n=1,2,...,N$)

**Parameters:** weight factors $\alpha_1$ and $\alpha_2$

**Objective:** filter outliers and estimate the travel time $t_{nk}^{\text{est}}$ based on $t_{nk}^{\text{obs}}$

Step 1: check the number of ETC measurements and found no observation

if ($N=0$) \( t_{nk}^{\text{est}} = \alpha_1 t_{k-1}^{\text{obs}} + (1-\alpha_1) t_f \)

Step 2: check the number of ETC measurements and found only one observation

if ($N=1$)

\{
if \left( t_{nk}^{\text{obs}} \geq \frac{1}{\alpha_2} t_{k-1}^{\text{obs}} \text{ and } t_{nk}^{\text{obs}} \leq \alpha_2 t_{k-1}^{\text{obs}} \right) \{ t_{nk}^{\text{est}} = \alpha_1 t_{k-1}^{\text{obs}} + (1-\alpha_1) t_f \}
else \{ t_{nk}^{\text{est}} = \alpha_1 t_{k-1}^{\text{obs}} + (1-\alpha_1) t_f \}
\}

Step 3: check the number of ETC measurements and found two observations

if ($N=2$)

\{
Subset $M$ observations $t_{nk}^{\text{obs}} \in \left[ \frac{1}{\alpha_2} t_{k-1}^{\text{obs}}, \alpha_2 t_{k-1}^{\text{obs}} \right]$, $m=1,2,...,M$ \& $M \leq N$

if ($M=0$) \{ $t_{nk}^{\text{est}} = \alpha_1 t_{k-1}^{\text{obs}} + (1-\alpha_1) t_f$ \}
if ($M>0$) \{ $t_{nk}^{\text{est}} = \alpha_1 t_{k-1}^{\text{obs}} + (1-\alpha_1) \text{median}(t_{nk}^{\text{obs}})$ \}
\}

Step 4: check the number of ETC measurements and found three or more observations

else

\{
Subset $M$ observations $t_{nk}^{\text{obs}} \in \left[ \frac{1}{\alpha_2} t_{k-1}^{\text{obs}}, \alpha_2 t_{k-1}^{\text{obs}} \right]$, $m=1,2,...,M$ \& $M \leq N$

if ($M=0$) \{ $t_{nk}^{\text{est}} = \alpha_1 t_{k-1}^{\text{obs}} + (1-\alpha_1) t_f$ \}
if ($M>0$) \{ $t_{nk}^{\text{est}} = \alpha_1 \text{median}(t_{nk}^{\text{obs}}) + (1-\alpha_1) t_{k-1}^{\text{obs}}$ \}
\}

Iteration $k$ : $k = k+1$

**CASE STUDIES**

**Travel Time Estimation**

The proposed travel time estimation method is applied to the New Jersey Turnpike (NJTPK), a part of the
Interstate Highway I-95. The studied highway is a closed tolling system that collects tolls through both a
traditional ticket system and an ETC collection system at the toll plazas. Despite the availability of the time
stamps for all vehicles entering and exiting the toll plazas, only EZ-Pass users were considered as probe
vehicles because other cash users may be delayed when picking up the tickets at an entry toll plaza,
returning tickets, and paying cash at an exit toll plaza. Data from 2011 were used. The case study section
is narrowed down to eight links between Interchanges 9 and 5 as shown in FIGURE 3.
When implementing the proposed travel time estimation approach, three time intervals are considered: 5-minute, 10-minute, and 15-minute. The 5-minute interval is considered to be the minimum updating interval to follow the traffic conditions that would evolve quickly (e.g., incident conditions). The proposed method is tested using data of multiple days randomly selected in 2011. The estimated link travel times based on different time intervals were compared through root mean square percentage error (RMSPE). RMSPEs are computed by comparing the observed travel time and the estimated travel time using the following equation:

$\text{RMSPE} = \sqrt{\frac{1}{M} \sum_{a=1}^{A} \sum_{b=1}^{B} \left( \frac{t_{\text{est},a} - t_{\text{obs},a,b}}{t_{\text{obs},a,b}} \right)^2}$

(8)

where $t_{\text{est},a}$ is estimated travel time at $a^{th}$ period, $a=1,2,\ldots,A$; $t_{\text{obs},a,b}$ is the $b^{th}$ observation in $a^{th}$ period, $b=1,2,\ldots,B$; $B$ that varies in each period is the total number of observations during the period; and $M$ is the total number of observations during all the analyzed periods.

FIGURE 4 and FIGURE 5 show two examples of travel time estimations for the link between interchange 08A and 009 in the northbound. FIGURE 4 shows the estimated travel time under normal conditions whereas FIGURE 5 illustrates the estimated travel time under incident conditions on the same weekday but from a different week.
FIGURE 4 Travel time estimation under incident-free conditions: a) observed travel time between two toll plazas based on ETC data, b) estimated link travel time

FIGURE 5 Travel time estimation under incident conditions: a) observed travel time between two toll plazas based on ETC data, b) estimated link travel time
The simple estimations shown in FIGURE 4 and FIGURE 5 represent are only based on the probe data collected between OD pairs associated with the target link. In contrast, the improved travel time estimations represent the results based on the additional probe data from OD pairs associated with the target link as well as the adjacent OD pairs. Particularly, it can be seen that the improved estimation captured the incident impact better while the simple one was responded with some time lag. In normal non-congested condition, both simple and improved estimations worked fine.

To further examine the performance of the improved method, RMSPEs errors were computed for travel time estimations of both different days and different conditions. TABLE 2 presents the computed RMSPEs. The results show that using the 5-minute interval achieved relatively larger RMSPEs compared to the RMSPEs found with longer time intervals. The 10-minute interval and 15-minute interval achieved similar results. To improve computational efficiency, the 15-minute interval will be used for deriving related performance measures.

### TABLE 2 Comparisons of estimations based on different updating time intervals

<table>
<thead>
<tr>
<th>Analyzed Links</th>
<th>RMSPE ( Day: 4/17/11)</th>
<th>RMSPE ( Day: 9/22/11)</th>
<th>RMSPE ( Day: 4/11/11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5-min</td>
<td>10-min</td>
<td>15-min</td>
</tr>
<tr>
<td>8 to 7A</td>
<td>0.341</td>
<td>0.141</td>
<td>0.143</td>
</tr>
<tr>
<td>7A to 7</td>
<td>0.427</td>
<td>0.325</td>
<td>0.295</td>
</tr>
<tr>
<td>7 to 6</td>
<td>1.058</td>
<td>0.183</td>
<td>0.186</td>
</tr>
<tr>
<td>6 to 5</td>
<td>0.331</td>
<td>0.250</td>
<td>0.265</td>
</tr>
<tr>
<td>7 to 7A</td>
<td>0.634</td>
<td>0.213</td>
<td>0.217</td>
</tr>
<tr>
<td>7A to 8</td>
<td>0.400</td>
<td>0.181</td>
<td>0.185</td>
</tr>
<tr>
<td>8 to 8A</td>
<td>0.274</td>
<td>0.184</td>
<td>0.162</td>
</tr>
<tr>
<td>8A to 9</td>
<td>0.231</td>
<td>0.120</td>
<td>0.121</td>
</tr>
</tbody>
</table>

**Comparative Analysis**

The estimated travel times based on ETC data are compared with the estimated travel time using measurements from remote traffic microwave sensors (RTMS) collected in 2011 along a section of the same facility. Travel time for each short link was calculated. The travel time \( t_{ij} \) between two interchanges was then calculated by summing these link travel times (i.e., travel time between interchanges 1 and 3 at time \( t \) is equal to the summation of the travel time \( t_{2,3} \) between \( t_{1,3} \) interchanges 1 and 2 and between interchanges 2 and 3: \( t_{1,3} = t_{1,2} + t_{2,3} \)). This summation represents the instantaneous travel time. The travel time estimation based upon the ETC data represents the experienced travel time, which is closer to the actual travel time experienced by a vehicle traveling from one interchange to another. Of course, when the traffic flow is close to a free flow condition, the travel times based on the RTMS and ETC data should also be close. However, if a certain link is congested, the RTMS located in the downstream of the congested link may report the free flow travel time because no vehicle passed would have reached this sensor due to the upstream congestion. In such a situation, the instantaneous travel time will not reflect the actual travel time. FIGURE 6 shows an example of the travel time obtained from different data sources. FIGURE 6 (a) and FIGURE 6 (b) show that under uncongested conditions the travel time based on the RTMS data is similar to the travel time estimated by probe data. FIGURE 6 (c) and FIGURE 6 (d) illustrate that the travel time based on the RTMS data is obviously different than the probe vehicle estimates during the congestion period. Thus, if the RTMS data were used to conduct travel time-based analysis (e.g., estimation of travel time reliability measures), the results will be biased, particularly when congestion occurs. It is advised, then, experienced travel time instead of instantaneous travel time should be used to quantify travel time related performance measures.
Example of Performance Measure Analysis based on Estimated Travel Times

In this section, the estimated travel time based on the methodology presented in the previous section has been used to develop travel time reliability measures. These travel time reliability measures for the eight links shown in FIGURE 3 are reported. The following TABLE 3 shows an example of computed measures based on one-year travel time information for each link. The free-flow travel time (FFTT) indicates the travel time based on a speed limit of 65 mph for each link. The mean travel time (MTT) represents an annual average of estimated travel time for all time periods. 95th percentile travel time (95th PTT), buffer index (BI), planning time index (PTI), and congestion frequency index (CFI) were also presented. TTI, BI, PTI, and CFI were calculated based on the equations presented below. The last four travel time indices (TTI) were considered for the purposes of showing the variation of travel time reliability during the AM peak period (7 am to 9 am) and PM period (5 pm to 7 pm) on different days.

\[ TTI = \frac{\text{Peak Period Travel Time}}{\text{Free Flow Travel Time}} \]  
\[ BI = \frac{95^{\text{th}} \text{ Percentile Travel Time} - \text{Average Travel Time}}{\text{Average Travel Time}} \]  
\[ PTI = \frac{95^{\text{th}} \text{ Percentile Travel Time}}{\text{Free Flow Travel Time}} \]  
\[ CFI = \frac{\text{Number of Periods with Travel Time Exceeds a Threshold}}{\text{Total Number of Periods}} \]
TABLE 3 Annual reliability measures for transportation performance evaluation

<table>
<thead>
<tr>
<th>Link</th>
<th>FFTT (min)</th>
<th>MTT (min)</th>
<th>95th PTT (min)</th>
<th>BI (%)</th>
<th>PTI</th>
<th>CFI (%)</th>
<th>Weekday AM TTI</th>
<th>Weekday PM TTI</th>
<th>Weekend AM TTI</th>
<th>Weekend PM TTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1S) 8-7A</td>
<td>6.55</td>
<td>7.14</td>
<td>9.92</td>
<td>38.9</td>
<td>1.515</td>
<td>5.2</td>
<td>1.015</td>
<td>1.100</td>
<td>1.001</td>
<td>1.099</td>
</tr>
<tr>
<td>(2S) 7A-7</td>
<td>6.18</td>
<td>6.49</td>
<td>8.12</td>
<td>25.1</td>
<td>1.314</td>
<td>2.8</td>
<td>0.984</td>
<td>1.058</td>
<td>0.989</td>
<td>1.072</td>
</tr>
<tr>
<td>(3S) 7-6</td>
<td>2.77</td>
<td>3.53</td>
<td>4.75</td>
<td>34.6</td>
<td>1.715</td>
<td>11.1</td>
<td>1.096</td>
<td>1.227</td>
<td>1.148</td>
<td>1.291</td>
</tr>
<tr>
<td>(4S) 6-5</td>
<td>6.00</td>
<td>5.59</td>
<td>6.59</td>
<td>17.9</td>
<td>1.098</td>
<td>0.7</td>
<td>0.948</td>
<td>0.932</td>
<td>0.948</td>
<td>0.925</td>
</tr>
<tr>
<td>(1N) 7-7A</td>
<td>6.2</td>
<td>7.34</td>
<td>12.96</td>
<td>76.6</td>
<td>2.09</td>
<td>8.1</td>
<td>1.097</td>
<td>1.232</td>
<td>1.081</td>
<td>1.390</td>
</tr>
<tr>
<td>(2N) 7A-8</td>
<td>6.47</td>
<td>7.17</td>
<td>11.28</td>
<td>57.3</td>
<td>1.743</td>
<td>7.2</td>
<td>0.996</td>
<td>1.176</td>
<td>0.992</td>
<td>1.261</td>
</tr>
<tr>
<td>(3N) 8-8A</td>
<td>6.00</td>
<td>6.38</td>
<td>8.11</td>
<td>27.1</td>
<td>1.352</td>
<td>2.7</td>
<td>1.054</td>
<td>1.077</td>
<td>1.031</td>
<td>1.103</td>
</tr>
<tr>
<td>(4N) 8A-9</td>
<td>8.77</td>
<td>8.52</td>
<td>9.86</td>
<td>15.7</td>
<td>1.124</td>
<td>1.0</td>
<td>0.934</td>
<td>0.973</td>
<td>0.926</td>
<td>0.979</td>
</tr>
</tbody>
</table>

If the BI is used to rank the links, the link between interchange 7 and interchange 7A in the northbound of NJ Turnpike with a BI of 76.6 percent is ranked as first. In other words, when crossing this link, travelers should incorporate increased time buffer into trip plans. Similarly, the PTI for this link is the highest, with a PTI value of 2.09. To ensure 95% on-time crossing, the total travel time needed is more than twice that of the link’s free-flow travel time. To calculate the CFI, the threshold value in equation (12) was assumed to be 1.5 times of free-flow travel time. As such, it was found that the link between interchange 7 and interchange 6 in the southbound lanes were more frequently subject to congestion as the CFI was 11.1 percent. This is followed by the link between interchange 7 and interchange 7A in the northbound. Overall, the TTI of AM periods during both weekdays and weekends were less than that of PM periods. This observation indicates that the travel time during the PM periods is higher than that of the morning peak period.

Since travel time reliability measures can be aggregated by different criteria, transportation agencies need to define their goals. For instance, if they are interested in understanding the change of the reliability measures by time, they should develop weekly, monthly, and seasonal measures. FIGURE 7 shows an example of monthly analyses of the weekday travel time reliability measures. It can be seen that all measures fluctuated according to the months. More importantly, the months with a higher travel time, buffer index, planning time index and congestion frequency index are easy to understand visually. Based on these characteristics of the travel time reliability measures, transportation agencies can implement necessary countermeasures to maintain the optimal performance of their facilities.
CONCLUSION
The main purpose of this study is to provide a travel time estimates using probe data collected by the ETC system instead of other sensors specifically deployed for traffic monitoring. However, the estimation can also present a significant challenge in achieving a high level of estimation accuracy when AVI readers are not located on the main roadway but only on the ramps. To break down path travel times into link travel times, an improved estimation method is developed. Results shows that the improved approach captures the actual travel time very well in both incident-free and incident conditions. According to the RMSPE, an update the interval of 15 min provides the best results for the highway of interest. Such an interval provides sufficient time to achieve a proper demand level and captures respond to rapid changes in the traffic flow (due to incidents, congestion, etc.). Our results, based on the field data, show that it is possible to use ETC data to provide accurate estimations of link travel times and to track roadway traffic conditions. The improved travel time estimates can be used to derive travel time reliability measures in a more rational manner than those instantaneous travel time estimates based on point sensors. As the travel time between the on-ramp / off-ramp and the toll plazas are not fixed, it is necessary to estimate the travel time for this portion too. A good strategy can be querying travel time from online free travel time information providers such as Google Map and Bing Map. When the travel time at different periods are collected, a reference distribution of these ramp travel time can be used in the equations for estimating mainline travel time.

For transportation agencies that have similar data sources, it would be possible to use the proposed approach for estimating travel time on their facilities and derive the performance measures. The estimated travel time based on the probe data from the ETC system could be used for validating, substituting, or

FIGURE 7 Variations of travel time reliability measures by month

(a) Monthly Weekday MTT and 95th PTT
(b) Monthly Weekday Buffer Index
(c) Monthly Weekday Planning Time Index
(d) Monthly Weekday Congestion Frequency Index
complementing other sensor data (i.e., RTMS, loop detector). Similarly, the proposed approach can be easily extended to other roadways that can identify vehicles through various technologies such as Bluetooth, Wi-Fi, and smartphones. Many private firms are interested in collecting traffic information using their own sensors / probe vehicles. Since the travel time from ETC systems capture actual point-to-point travel time, it will be a good reference point for comparison and validation purposes. Nevertheless, due to many issues discussed in this study, the probe data from ETC systems also need to be carefully filtered to estimate travel time in a reliable manner rather than using raw observations that might have considerable noise or measurement errors.

ACKNOWLEDGEMENTS
The contents of this paper only reflect views of the authors who are responsible for the facts and accuracy of the data presented herein. The funding for this project was partially provided by the New Jersey Turnpike Authority and University Transportation Research Center at City University of New York. This support is greatly appreciated. The authors are also grateful to the NJTA for providing the data used in case studies. The contents of the paper do not necessarily reflect the official views or policies of the any agencies. The authors thank for the insightful comments and suggestions from all the anonymous reviewers that helped improve the paper.

REFERENCES


