Section Travel Time Estimation from Point Detection Data

UCI-ITS-TS-WP-02-14

Jun-Seok Oh
R. Jayakrishnan
Will Recker

1 Institute of Transportation Studies
University of California, Irvine, jsoh@uci.edu

2 Department of Civil and Environmental Engineering and
Institute of Transportation Studies
University of California, Irvine, jayakr@uci.edu, wrecker@uci.edu

August 2002

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu
Section Travel Time Estimation from Point Detection Data

Jun-Seok Oh
Postdoctoral Researcher
Institute of Transportation Studies
University of California
Irvine, CA 92697-3600, USA
Tel: (949) 824-1672, Fax: (949) 824-8385. Email: jsoh@uci.edu

R. Jayakrishnan
Associate Professor
Department of Civil and Environmental Engineering
Institute of Transportation Studies
University of California
Irvine, CA 92697-3600, USA
Tel: (949) 824-2172, Fax: (949) 824-8385. Email: rjvaykri@uci.edu

Will Recker
Director and Professor
Department of Civil and Environmental Engineering
Institute of Transportation Studies
University of California
Irvine, CA 92697-3600, USA
Tel: (949) 824-5642, Fax: (949) 824-8385. Email: wrecker@uci.edu

Submission date: August 1, 2002
Word count: 4,257 + 7 x 230 = 6,087

Paper submitted to the 82nd Annual Meeting of the Transportation Research Board,
Section Travel Time Estimation from Point Detection Data

Jun-Seok Oh, R. Jayakrishnan, and Will Recker
Institute of Transportation Studies, University of California, Irvine

ABSTRACT

Despite recent development efforts in many other advanced traffic surveillance systems, inductance loops are still the most widely used traffic detectors. This paper points out a deficiency of point measurements obtained from such loop detectors in estimating travel times under congested traffic conditions and proposes a theoretically sound and practically applicable travel time estimation algorithm that uses the same loop detector data. The main idea of the algorithm is based on the concept of section density that can be easily obtained by observing in-and-out traffic flows between two point detection stations, with some careful corrections for detector errors. Travel times estimated from the proposed method are compared to those of other methods via both simulated and real traffic data. While the method estimating travel time based on spot speeds tends to underestimate section travel times due to failure of capturing the traffic congestion occurring between detector stations, the proposed section-density-based method provides fairly accurate travel time estimates.

Keywords: Travel time estimation; traffic surveillance systems; kinematic traffic flow theory

INTRODUCTION

Accurate traffic surveillance systems are a core element in transportation systems. With the advent of Intelligent Transportation Systems (ITS), accurate estimation of link travel times over freeway networks has become an important issue. An urgent issue in transportation surveillance systems is how to use existing surveillance infrastructure for better travel time estimation. This study develops a theoretically sound and practically applicable method for accurate estimates of freeway travel times.

For accurate travel time estimation, many advanced techniques have been applied (1): electronic distance-measuring instruments, computerized and video license plate matching, cellular phone tracking, automatic vehicle identification, automated vehicle location, video imaging, etc. However, the existing infrastructure in most urban contexts is for point-detection of traffic, mostly using pavement loop detectors. These are generally inductive loop detectors, nearly single loop detectors. Many studies have noted that the estimated travel times from single-loop data are flawed since the estimated speed is based on the assumption of common vehicle length (2, 3).

Realizing the limitations and problems in speed estimation using single loop data, several researchers have attempted to develop better algorithms for accurate speed estimation (2 - 4). Also, there have been some efforts to estimate travel times using single-loop detectors (5 - 8). The basis of these works was stochastic models of traffic flow and estimated travel times by investigating traffic flows. Most of these studies have focused on overcoming the problems in speed estimation from single-loop data, assuming that the travel times estimated from double loop data are accurate. However, the travel times estimated from double-loop speeds may also be flawed under congested traffic conditions, which is perhaps when accurate travel time estimation is most important.

Previous delay estimation studies under congestion (or incidents) have usually employed either the shockwave theory proposed by Lighthill and Whitham (9) or deterministic queuing models. Chow (10) and Wijesinghe (12) used the shock-wave theory in estimating delays and showed that the total delay from the shock-wave is identical to that obtained from queueing models. Morales (12) showed analytical procedures estimating delay for various cases based on deterministic queueing models. The cumulative vehicle arrivals and departures were used for delay estimation. Al-Deek et al. (13) proposed a method to determine temporal and spatial range of incident effects by analyzing shock waves from loop detector data, and a method to estimate incident delays for both cases with single and multiple incidents. However, despite its theoretical soundness, the linearized shock wave analysis is limited to only a means of rough estimation due to the nonlinearity of the shock wave.

In this paper, we show that conventional methods for estimating travel time from double loop speeds tend to underestimate travel times under congested traffic conditions, especially during incidents, and propose an
analytical and pragmatic approach to estimate travel times on a freeway section using traffic data from point detection stations. A key method is the use of point detectors’ count differences for density calculations in the section between them, with appropriate techniques suggested for correcting the detector errors. Using a scheme based on simple fluid models and with proper definitions of true travel time estimates for a section, we are able to find good results from this scheme, which may have significant practical applications.

MEASURING TRAVEL TIMES

Travel time can be defined simply as the time spent in traveling from one point to another. However, the travel time within discrete time and space requires more comprehensive considerations. Travel times are represented by discretizing the temporal and spatial dimensions. In this section we describe four different travel time measures, one a theoretical benchmark and the other three practical estimates. These four are used to for performance comparison of the new technique we propose in the following section, that is based on the use of section density estimated from point detectors.

We first describe the actual (true) section travel time based on the correct section space-mean speed, though it cannot be found from just the point detectors, as vehicle positions at given points in time would also be required for it. This is however our theoretical benchmark. We then show three other types of travel time estimates (based on travel times of re-identified vehicles and based on double and singles loop speed estimation) which are used in practice and in research. These four methods for comparison are explained in this section, and we point out the biases in the latter three techniques.

True Section Travel Time

Within discrete time and space, the representative travel time can be defined as a mean travel time within the closed area defined by the time \((t_j \text{ and } t_f)\) and space \((x_i \text{ and } x_f)\) as in Figure 1. According to the most correct definitions of space-mean speed over time-space region, the true space-mean speed \(\bar{V}\) for the vehicles within the closed area is the total travel distance divided by the total travel time of all vehicles in the closed area \(\Delta t\). Of course, point detectors at the two ends of a section are not sufficient to find such data and we are considering this here only as a theoretical benchmark to compare other estimates if he had such ground truth data (or complete simulated data). An unbiased estimate of space-mean speed is:

\[
\bar{V} = \frac{\sum_{t_j}^{t_f} \min(x_i^t, x_f^t) - \max(x_i^t, x_f^t)}{\sum_{t_j}^{t_f} \min(t_j, t_f^t) - \max(t_j, t_f^t)}
\]  

(1)

where,
- \(x_i^t\) = position of the vehicle \(n\) at time \(t\)
- \(x_f^t\) = position of the vehicle \(n\) at the downstream station \(d\)
- \(t_f^t\) = time when the vehicle \(n\) passes the upstream station \(u\)
- \(t_f^t\) = time when the vehicle \(n\) passes the downstream station \(d\)
- \(N\) = the number of vehicles traversing the section during the time interval

The "true" estimate of average section travel time can then be estimated from this space-mean speed and the distance of the section. Such an average section travel time can be considered a true mean travel time that represents the section. In this study, this section travel time is used as a benchmark travel time.

\[
\Delta t = \Delta x / \bar{V}
\]

(2)

where,
- \(\Delta x\) = average section travel time
\[ \Delta x = \text{length of the section } (x_u - x_s) \]

**Figure 1** Temporal and Spatial Illustration of Section Travel Time

**Point Detector Estimate based on Vehicle Travel Times: Arrival-based Travel Time**
Recent advances in sensing technology, such as Automatic Vehicle Identification (AVI) and vehicle re-identification from vehicle waveforms using advanced loop detectors (4/5), provide information on timestamps when vehicles pass predefined locations. In such circumstances, the travel time during a time interval can be calculated by averaging individual vehicles’ travel times from upstream to downstream. That is, the arrival-based travel time \( t_{av} \) is calculated based on the vehicles that arrived at the downstream station during a given time-step.

\[
t_{av} \approx \frac{\sum (x_u - x_s)}{N}
\]

Unlike the true travel time estimate in the previous section, this estimate would not be just based on vehicle travel during the current time step (from \( t \) to \( t + 1 \) in Figure 1). This is because many of the vehicles whose travel times over the section is averaged spent part of their travel in the section in the earlier time step (see the second through the 5th vehicle trajectories in the figure). The longer the length of the section (or shorter the time period), the greater this time lag effect.

**Point-detector Estimates based Speed Estimation:**
When individual arrival times and vehicle re-identification facilities are not available at point detectors, the travel time estimation will need to be based on average speed estimates at the detectors, either with double loops or with singles loops. We describe the obvious problem with these estimates in this section.

**Double Loop Estimation**
Double loop detectors are capable of observing speeds by using Reed traps between two closely spaced detectors. Estimated travel time based on this double-loop speed has been considered fairly accurate. Generally, the section travel time between detector stations is calculated by averaging the numbers from upstream and downstream stations. The simple double-loop travel time is defined as:

\[
t_{DL} = \frac{\Delta x}{(v_u + v_s)/2}
\]

where,
\[ v_c = \text{double loop speed at upstream detector} \]

\[ v_d = \text{double loop speed at downstream detector} \]

**Single Loop Estimation**

Unlike the double loop case, speeds at single loop detector stations are estimated based on the assumption of common vehicle length. The single-loop speed \(v_{1L}\) is calculated as

\[ v_{1L} = \frac{q}{o} \times g \]

where,

- \(q\) = traffic flow
- \(o\) = occupancy
- \(g\) = average effective vehicle length

Any bias from point detection, as in the double-loop detection case, is applicable in this case also; furthermore, additional errors are introduced by the estimated point speeds as well. Previous studies have found many problems in estimating travel times using single loop data and have attempted various methods for better estimation. The main issues in travel time estimation using single loop data is how to eventually handle the effective vehicle length for accurate speed estimation. The basis of these works was stochastic models of traffic flow and estimates of travel times by investigating traffic flows.

**Bias in Travel Time Estimates Based on Point Speed Estimation**

The two travel time estimation methods based on single or double loop speed estimates would be correct only under the assumption that the traffic condition in the section is either homogenous or a linear combination of the two points. However, this assumption is no longer valid under congested traffic conditions due to incidents or other reasons. That is, the travel time estimated from double loop speeds tends to be biased under congested traffic conditions.

![Figure 2: Failure of Point Detection in Observing Congestion](image)

This limitation of point detection systems (even with double loop speeds) in measuring delays under congested traffic conditions is perhaps known to many researchers, and is easy to see in a simple shock-wave analysis. In general, the congested region can be represented by the triangular area in Figure 2. Traffic conditions are represented by temporal spatial space corresponding to the data aggregation interval and detector spacing.
Consider the four time-space cases (a) - (d) in Figure 2, each denoting a section between a pair of detector stations (1 and 2 or 2 and 3, for a given time step. We can see that only the travel time in the area of (a) can only be correctly estimated even with double-loop speed estimates. In the case of (d), point measurements can represent travel times fairly well since the section is homogeneous and speeds at both station 1 and 2 are similar. For the case of (c) and (d) with a shock wave crossing the regions, one of the two detector stations is within the congested region, so the linear combination of two speeds may be used for the estimation. The simple average of two detector stations will result in biased estimates unless the length of the queue covers exactly half of the section (or more rigorously, when the shock wave divides the rectangular region into two equal areas). Alternatively, for case (d), neither the upstream nor the downstream detector is in the congested region. Certainly, the estimate even from the double-loop speed underestimates the travel time in the section.

Though we used an example with abrupt shock waves, it is easy to see that the same problem extends to the case of gradually (continuously) varying traffic conditions as well. During a congestion period with stop-and-go traffic phenomenon, the point detectors may fail in capturing such delay unless the detectors are closely spaced. Overall, it can be stated that point detection systems tend to underestimate travel times mainly due to their limitation in capturing traffic congestion.

The discussion in this section has focused on the general estimation schemes for travel times and has also pointed out some of the problems in the existing schemes. We do however use the above three cases of travel time estimation (based on vehicle re-identification, and based on single and double loop speed estimates) for comparison purposes against the new methodology we propose, in the next section.

**Methodology Based on Section Density Estimates from Point Detectors**

**Background**

The technique we propose here is based on the assumption that we can find the true density in a section at a given point in time purely based on point detectors, if the point detectors are perfect detectors and a cumulative count on two successive detectors is available. If so, a simple subtraction of cumulative arrivals at the two detectors would yield the number of vehicles in between the detectors and thus the density of the section between them. If such a density is known at any point in time, we can make a very good estimate of the true section travel time using simple fluid model relations for traffic, as we discuss next. The reality however is different, in that the detectors in the field are not perfect and each detector has its own tendency to undercount vehicles and thus the cumulative arrival counts at two detectors (with their own "cumulative count drift") cannot be used to find the density at any time. Moreover, the cumulative counts need to start at a time when there are no vehicles in the section, which is considered not easy to determine with point detectors (i.e., when the "reset" of the cumulative counts). These are however, not insurmountable problems and we have proposed ways to handle them. Interestingly, we have not been able to find literature discussing such a simple point count subtraction method for density estimates, as the technique is perhaps the most obvious. In the absence of counter-evidence, we assume that we are not overlooking some obvious difficulty that has been discussed in any older literature that we have not seen. We have to assume that a lack of easy solutions to the drift and reset problems we identify (and have provided solutions to) is the reason why this count-subtraction technique has not been looked into as the past. Our technique in this section effectively embeds the count-subtraction scheme in the density updates through flow measurements. Before discussing the section density detection technique, we first show how it can be used for travel time estimation.

The model we are proposing here is based on the macroscopic hydrodynamic traffic flow theory developed by Lighthill and Whitham (9) and Richards (16). The main result of their work is the derivation of kinematic waves that satisfy a first order partial differential equation. The well known basic macroscopic equation of traffic is equivalent to the fluid conservation equation that characterizes compressible flow, that is,

\[
\frac{\partial \rho}{\partial t} + \frac{\partial \rho u}{\partial x} = 0 \quad \text{(or traffic generation rate)}
\]

where,

\[\epsilon = \text{flow (vehicles/hour)}\]
The basic identity for such traffic fluids is \( q = k \cdot v \). From this equation, travel time can be derived as a function of section density and traffic flow rate. The simplest representative estimate of the section travel time would be an average of the speeds that depend on the section density at the beginning and end of the time period, as in Figure 3.

\[
\tau_{SD} = \frac{\Delta x}{\bar{v}} = \frac{\Delta x \cdot k}{q} = \frac{\Delta x \left( k(t) + k(t\Delta t) \right)}{[q(t) + q(t\Delta t)]} \tag{7}
\]

where,
- \( \tau_{SD} \) = section-density-based travel time
- \( \Delta x \) = length of the section

### Measuring Section Density and Travel Time

As described in equation (7), the section travel time can be estimated using the section density and traffic flow rates. In a discrete section of freeway (Figure 4), let us consider \( k(t) \), the traffic density in a section at time \( t \), and define \( I \) as the number of lanes in the section. In addition, let us define \( q_{in}(t) \) and \( q_{out}(t) \) as the flow passing through the upstream and downstream stations during the time interval (between \( t-I \) and \( t \)), respectively. The variables \( q_{in}(t) \) and \( q_{out}(t) \) denote flows entering and exiting the section via on and off ramps, and \( L_u \) and \( L_d \) represent the number of lanes on the ramps, respectively. Then, the section density equation over discrete time is represented by:

\[
k(t + \Delta t) = \bar{k}(t) \ast [\Delta t \frac{\Delta x}{\Delta t} [I \cdot (q(t) - q_{out}(t))] + \Delta t \cdot q_{in}(t) - L_u \cdot q_{in}(t)] \tag{8}
\]
Figure 4. A Discrete Section of Freeway

From equation (8), section density can be obtained from a point detection system by observing the flow rates. However, there are two major obstacles in measuring the section density. The first issue is how to find the initial section density (the "reset" problem mentioned above). The second issue is how to adjust the systematic errors of detectors in traffic counting (the "drift" problem).

In order to set the initial section density, we start from the derivation of section density from the occupancy measures. Since the occupancy is the portion of time occupied by vehicles during a time interval, the fraction of time can be easily converted to the section density by assuming an average vehicle length. Because the initial section density is calculated based upon the assumption of average vehicle length, the proposed method is also affected by the assumed average vehicle length. However, this does not affect the travel time estimation as much as in the single-loop travel time estimation case because the average vehicle length is used only for the calculation of the initial number of vehicles in the section; all subsequent section densities are computed from the difference between in and out counts.

\[
k = \frac{\alpha \times L}{g}
\]

(9)

where,
\[L = \text{unit length of the section density}
\]
\[g = \text{average vehicle length}
\]

In the initial section density calculation, we need to consider two conditions to ensure that the section condition is homogenous and uncongested. These conditions are determined by observing point measurements from both upstream and downstream. First, approximately same occupancies at both upstream and downstream detection stations can be regarded as an indicator of homogenous conditions in the section. The speed close to the free flow speed \(v_f\) can also signify uncongested traffic conditions. Accordingly, the traffic condition for the section density initialization can be represented by:

\[
o_1 = o_2 \text{ and } v_1 = v_2 = v_f
\]

(10)

The second issue on systematic errors of detectors is a practical issue, but critical in the operation of the proposed method. This practical problem can be resolved by using calibrated parameters for individual detectors, or by calibrating an aggregated parameter for the relative systematic error between in and out detector stations. The latter is more practical considering the current detector maintenance situation. Since the cumulative values of inflow and outflow should be identical if observed over a long period, the aggregated parameter \(\alpha\) can be calibrated as:

\[
\alpha = \frac{\sum \left[ I \cdot q_1(t) + I_{\text{off}} \cdot q^{\alpha}(t) \right]}{\sum \left[ I \cdot q_1(t) + I_{\text{off}} \cdot q^{\alpha}(t) \right]}
\]

(11)
Applying the calibrated parameter (α), finally the section density equation (8) and the estimated travel time (7) can be rewritten as:

\[ k(t + 1) = k(t) + \left( \frac{\Delta t}{1 + \Delta t} \right) \left[ (l \cdot q_s(t) + I_w \cdot q_m(t)) - \alpha \cdot (l \cdot q_s(t) + I_w \cdot q_m(t)) \right] \]  

\[ DR(t) = \frac{\Delta t \cdot (k(t + 1) + k(t))}{[I_w(t) + q_s(t)]} \]  

(13)

The proposed method can be applied to both off-line and on-line estimation. Systematic identification of loop detector error and homogeneous traffic condition is the key process in converting point measurements to section measurements. Figure 5 describes the on-line operation scheme of the proposed methodology.

![On-line Operation Scheme of the Proposed Methodology](image)

**Figure 5 On-line Operation Scheme of the Proposed Methodology**

**METHODOLOGY EVALUATION**

The proposed methodology is tested and evaluated via loop detector data from both simulated and real traffic data. Whereas results from empirical data are used for comparison between methods using loop detector data, the simulated traffic data can be used for the evaluation of different methods by comparing their estimates with the true section travel time.
We first tested the proposed methodology with a set of simulated traffic data. A microscopic traffic simulation model, Paranics, was used for the comparison of aforementioned various travel times. In addition to the travel times estimated from loop detector data, we also computed the true section travel time and the arrival-based travel time with the data of individual vehicles' trajectory that are available only within the traffic simulation. The traffic data were simulated on a section of I-805 freeway in Irvine, California based on real traffic data on May 23, 2001, and an incident was injected for 10 minutes from 6:50 to 7:00. The study section included an off-ramp, so the section density was estimated from one upstream and two downstream stations. Since there was no systematic detector error in the simulation data, a parameter of 1.0 was applied for the section density computation.

The second data set was obtained from double-loop detectors on a freeway section of I-880 in Hayward, California. The data used here was collected on February 18, 1993 and we used an aggregate interval of 10 seconds. The study section does not include any on or off-ramp, and the initial section density was calculated from the occupancy value of 0.0841 at 14:00 and a parameter of 0.9960, obtained from the difference of cumulative traffic counts during 14:00 a.m. to 20:00, was applied to the downstream detector station to fix the system detector error.

As shown in Figure 6, there were substantial differences between estimates during the period from 6:50 to 7:10, while the estimates were similar during the non-congested period. This shows that the point measurements can be used for travel time estimation under non-congested traffic conditions, but they underestimate during congestion period. As compared in Table 1, the estimates from double-loop and single-loop underestimated the travel times by 25% – 30%. Alternatively, the proposed method using the section density calculated from the point measurements was able to estimate the travel time within 5% error. In the comparison, the estimate from single-loop data showed smaller difference than that from double-loop, but it was due to the assumption of average vehicle length. In the empirical analysis during the congested period, the estimates from point measurements were about 35% lower than that from section density. The comparison with empirical data also complies with the simulation results.

![Figure 6](image-url)
Table 1 Comparison of Travel Timer during Congestion Period

<table>
<thead>
<tr>
<th></th>
<th>True Session Travel Time</th>
<th>Arrival-based Travel Time</th>
<th>Double-loop Travel Time</th>
<th>Single-loop Travel Time</th>
<th>Section Density Travel Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1:405, 6:30-7:30)</td>
<td>47.06</td>
<td>46.36</td>
<td>34.36</td>
<td>75.25</td>
<td>48.76</td>
</tr>
<tr>
<td>Travel time (sec.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative %</td>
<td>100.0</td>
<td>98.51</td>
<td>72.58</td>
<td>74.90</td>
<td>103.60</td>
</tr>
<tr>
<td>Empirical Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1:380, 14:40-15:30)</td>
<td></td>
<td></td>
<td>56.17</td>
<td>57.78</td>
<td>86.84</td>
</tr>
<tr>
<td>Travel time (sec.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative %</td>
<td></td>
<td>-</td>
<td>64.69</td>
<td>66.52</td>
<td>100.0</td>
</tr>
</tbody>
</table>

CONCLUSION

This paper identifies a deficiency of point measurements obtained from such loop detectors in estimating travel times under congested traffic conditions and proposes a new travel time estimation approach without any additional data requirements or infrastructure investment. The basis of the algorithm lies in the hydrodynamic kinematic traffic model. The proposed method is capable of estimating very accurate travel times by simply reconstructing the same data set. Even though this method can be applied only to sections where all entering and exiting traffic data are available, its practical application is promising due to its ease of application and simplicity.

REFERENCES


