

Microscopic Estimation of Arterial Vehicle Positions in a Low-Penetration-Rate Connected Vehicle Environment

Noah J. Goodall, Byungkyu “Brian” Park, and Brian L. Smith

Pre-print version. Published in the *Journal of Transportation Engineering* 140(10), October 2014.
[http://dx.doi.org/10.1061/\(ASCE\)TE.1943-5436.0000716](http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000716)

Abstract

Wireless communication among vehicles and roadside infrastructure, known as connected vehicles, is expected to provide higher-resolution real-time vehicle data, which will allow more effective traffic monitoring and control. Availability of connected vehicle technology among the vehicle fleet will likely grow gradually, and possibly remain limited, with many drivers potentially being unwilling to transmit their locations. This is problematic given that research has indicated that the effectiveness of many connected vehicle mobility applications will be dependent on the availability of location data from a minimum of 20-30% of roadway vehicles. In an effort to improve the performance of connected vehicle applications at low connected vehicle technology penetration rates, we propose a novel technique to estimate the positions of non-communicating (unequipped) vehicles based on the behaviors of communicating (equipped) vehicles along a signalized arterial. Unequipped vehicle positions are estimated based on observed gaps in a stopped queue, and the forward movement of these estimated vehicles are simulated microscopically using a commercial traffic simulation software package. In simulations, the algorithm made more correct than incorrect estimates of unequipped vehicle positions in the same lane and within 7 meters longitudinally. When applied to a previously-developed connected vehicle traffic signal control strategy in simulation, the location estimation algorithm produced small improvements in delays, speeds, and stopped delay when compared to an equipped vehicle-only scenario at penetration rates of 25% or less. The location estimation algorithm is generic, and could be applied to other connected vehicle applications to improve performance at low penetration rates.

INTRODUCTION

The concept of connected vehicles – previously known as IntelliDriveSM or Vehicle Infrastructure Integration – uses advanced wireless communications, global positioning systems (GPS), vehicle sensors, and smart infrastructure to allow vehicles and the infrastructure to communicate wirelessly. In a connected vehicle environment, equipped vehicles can share their locations, speeds, headings, and many other data in real-time with nearby vehicles and the surrounding infrastructure via wireless communications. While the exact method of communication is not specified, a dedicated short range communication (DSRC) radio linked

Noah J. Goodall, Research Scientist, Virginia Center for Transportation Innovation and Research, 530 Edgemont Road, Charlottesville, VA 22903-2454. Email: noah.goodall@vdot.virginia.gov

Byungkyu “Brian” Park, Associate Professor, Department of Civil and Environmental Engineering, University of Virginia, P.O. Box 400742, Charlottesville, VA, 22904-4742. Email: bpark@virginia.edu

Brian L. Smith, Professor, Department of Civil and Environmental Engineering, University of Virginia, P.O. Box 400742, Charlottesville, VA, 22904-4742. Email: briansmith@virginia.edu

with the vehicle's wheel speed sensor and a smart phone using its three-axis accelerometer, GPS receiver, and the cellular network would accomplish essentially the same task.

Researchers have developed several applications that utilize the data from connected vehicles to improve safety and mobility, such as traffic signal control and automated incident detection. Because of the gradual rollout of connected vehicle technologies, researchers often study the effect of different penetration rates on an application's performance. In many connected vehicle mobility applications, the performance of the application improves as the percentage of vehicles equipped with communication devices increases. For example, traffic signal control algorithm PAMSCOD reduced delay by an additional 12% when the penetration rate increased from 20% to 60% (He et al. 2012). Similarly, Priemer and Friedrich experienced a 6.5% improvement in vehicle speeds when increasing penetration rate from 10% to 50% in their traffic signal control algorithm (Priemer and Friedrich 2009).

Even though connected vehicle applications perform better at higher connected vehicle penetration rates, many are unable to outperform traditional applications at low penetration rates. The minimum required penetration rate needed varies based on the application, but is typically near 20-30% penetration rate, as seen in Table 1.

Table 1. Connected Vehicle Applications and Corresponding Minimum Required Equipped Vehicle Penetration Rates

| Application | Minimum Equipped Vehicle Penetration Rate |
|---|---|
| Traffic signal control ¹ | 20-30% |
| Freeway incident detection ² | 20% |
| Lane-level speed estimation ³ | 20% |
| Arterial performance measurement ⁴ | 10-50% |
| Queue length estimation ⁵ | 30% |

¹(Goodall et al. 2013; He et al. 2012; Priemer and Friedrich 2009; Smith et al. 2011), ²(Barria and Thajchayapong 2011), ³(Rim et al. 2011), ⁴(Li et al. 2008), ⁵(Ban et al. 2011)

The problem of low penetration rates is of particular concern because of the expected low penetration rate of connected vehicles, at least initially. The John A. Volpe National Transportation Systems Center estimates that based on vehicle fleet turnover rates, a federal mandate to install DSRC radios in all new vehicles would still only equip half of all vehicles on the road after nine years (John A. Volpe National Transportation Systems Center 2008). If the application instead relies on smart phones with GPS receivers, users would need to opt-in to share their data, with widespread adoption unlikely.

Testing in simulation has shown that when more vehicles can report their locations, the performance of connected vehicle mobility applications improves. We explore two additional questions:

1. Can the locations of some unequipped vehicles be estimated from the behavior of a few equipped vehicles?

2. Can these estimated locations then be used to improve the performance of connected vehicle mobility applications?

If both questions can be answered affirmatively, then the performance of some connected vehicle applications can be improved at low penetration rates merely by analyzing data from a few equipped vehicles and estimating the positions of unequipped vehicles. These techniques could improve mobility with minimal associated cost.

Several terms are used throughout this paper, and are defined here for consistency.

- *Connected vehicles.* The system of wireless communications among vehicles and the infrastructure.
- *Equipped vehicles (EV).* Vehicles which are equipped with the necessary hardware and software and are participating in the connected vehicle system.
- *Unequipped vehicles.* Vehicles which do not participate in a connected vehicle system, either due to lack of equipment, equipment failure, bandwidth restrictions, or a driver's unwillingness to share information.
- *Inserted vehicles.* Vehicles that exist only in a simulation, and which represent the location algorithms' estimates of the positions of unequipped vehicles.

BACKGROUND

Microscopically estimating the locations of vehicles based on wirelessly transmitted sparse data is a recent area of interest. Early work focused on freeways due to the advantages of working with a controlled access environment. Early efforts used calculated travel time information from cell tower signal triangulation (Bargera 2007; Sanwal and Walrand 1995; Westerman et al. 1996). The more accurate data from satellite navigation, such as the Global Positioning System (GPS), was integrated into later efforts (Krause et al. 2008). Nanthawichit et al. integrated GPS data with point detection data to estimate freeway travel times (Nanthawichit et al. 2003).

Herrera and Bayen used Kalman filtering techniques and Newtonian relaxation to integrate fixed-position detector data and individual vehicle positions and speeds into a high resolution traffic state estimation of a freeway (Herrera and Bayen 2010). Their algorithms were evaluated using both empirical ground truth freeway data (Federal Highway Administration 2010) and actual in-vehicle cell phones with GPS receivers (Herrera et al. 2008). With up to 25% of vehicles participating, they were able to estimate vehicle densities of multilane 36.6-meter segments, but did not estimate individual vehicle locations.

Goodall et al. used analyzed deviations in vehicle accelerations from those predicted by a car-following model, and inserted those vehicles into a simulation of a freeway network (Goodall 2013; Goodall et al. 2014). These inserted vehicles moved forward according to the car-following model within the same lane until overlapped by an equipped vehicle, and served as estimates of unequipped vehicle positions. This technique assumes that all vehicle behavior is in response to another vehicle, which precludes its implementation on an arterial where vehicles react to signals, pedestrians, upcoming turns, etc.

Recent work has estimated vehicle positions on arterials. Ban et al. calculated a vehicle's time-in-queue from the difference between its reported travel time and ideal travel time (Ban et al. 2011). Using this information, Ban et al. estimated the arrival rate at the intersection, and by assuming uniform flow rate and constant discharge rate, they were able to estimate the total

length of the queue with only 30% of vehicles reporting their locations. Their procedure required several restrictive assumptions, namely that the arrival between two vehicles in one cycle be considered the uniform arrival rate across the entire cycle, and a relatively high penetration rate so that there are at least two equipped vehicles per cycle per approach.

Using Ban et al.'s queue estimation technique, Sun and Ban attempted to reconstruct the precise trajectories of the unequipped vehicles that make up the total queue (Sun and Ban 2011). However, their algorithm assumed that the number of unequipped vehicles arriving between equipped vehicles is known, implying the use of an upstream fixed-point detector. These assumptions were relaxed in later work without a significant decrease in performance (Sun and Ban 2013). The behavior of these unequipped vehicles did not follow a car-following model, but instead followed an unrealistic linear acceleration profile. Mehran et al. used a similar approach to reconstruct vehicle trajectories using data from both fixed and probe sensors, and accounting for vehicles entering and exiting from the middle of the study section (Mehran et al. 2012).

Of the unequipped vehicle location estimation techniques discussed, all have one or more of the following four shortcomings:

1. Aggregate estimated vehicle locations spatially.
2. Use (or assume) unrealistic arterial vehicle behaviors in their models.
3. Do not estimate vehicle movements over a distance beyond a small study area of a few meters.
4. Reconstruct trajectories rather than estimate in real-time.

The location estimation techniques developed as part of this research were designed to overcome these shortcomings by continuously estimating lane-level positions of individual vehicles through a signalized corridor in real-time. Because of these fundamental differences between the location estimation algorithm and other approaches, they are not directly comparable.

DESCRIPTION OF THE ALGORITHM

The procedure to estimate the positions of vehicles consists of the following four-step process:

1. Determine from its observed behavior when an equipped vehicle is reacting to a previously unobserved unequipped vehicle based on gaps in a stopped queue.
2. Estimate the unequipped vehicle's initial position and speed, and insert it into a rolling, real-time simulation of the vehicles in the network.
3. Simulate the movements of the inserted vehicle over time.
4. Determine when the inserted vehicle estimate is no longer correct, and remove it from the simulation.

The four steps are described in greater detail in the following sections.

Step 1: Queue Gap Detection

The first step in estimating the locations of unequipped vehicles is determining when an equipped vehicle's behavior indicates the presence of a previously undetected unequipped vehicle. On an arterial, the most obvious example of this unexpected behavior occurs in a stopped queue. If an equipped vehicle leaves a large gap between itself and the leading vehicle at

any time, or between itself and the stop bar during a red phase, then an unequipped vehicle is probably in the gap.

In this step, the algorithm finds all stopped vehicles (both equipped vehicles and vehicle estimates inserted during previous time steps) within 50 meters of the stop bar on all intersection approaches. This distance of 50 meters was selected to be long enough to capture the majority of stopped queues, yet short enough that to disregard stopped vehicles at the upstream intersection. In practice, this distance can be altered to fit the geometries of the intersection.

Figure 1 demonstrates how the algorithm calculates queue length. Stopped vehicles are defined as vehicles with a speed of less than 1 meter per second. If the signal is red, then the start of the queue is the location of the stop bar. If the signal is amber or green, then the start of the queue is the location of the front of the stopped vehicle nearest the stop bar. The end of the queue is the location of the front of the stopped vehicle that is both farthest from and within 50 meters of the stop bar. The vehicles at the start and end of the queue may be in different lanes.

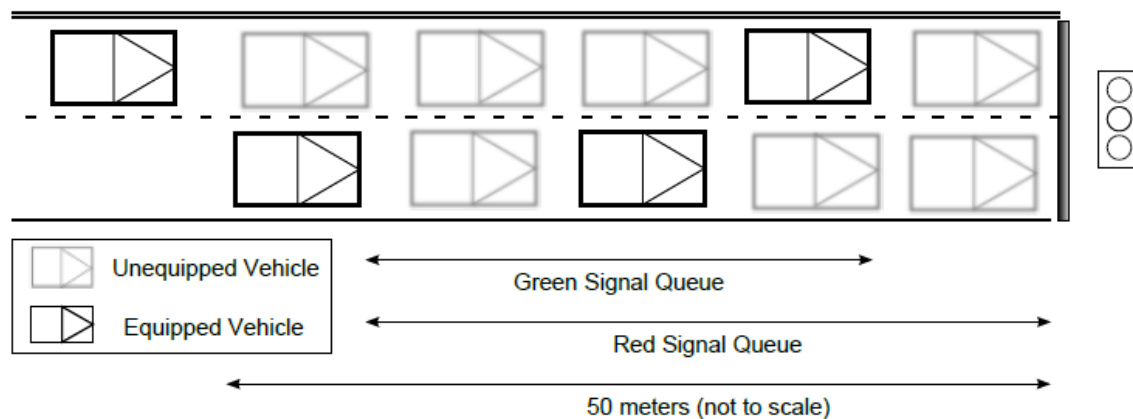


Figure 1. The queue length as calculated in the arterial location estimation algorithm, for both green and red phases.

Once the beginning and end of the queue are determined, the vehicle positions within the queue are analyzed in order to find gaps. Vehicles in this analysis are assumed to behave according to the Wiedemann car-following model, based on the empirical values collected by Wiedemann and Reiter (Wiedemann and Reiter 1992) (and in some cases extrapolated from Wiedemann and Reiter's charts by Olstam and Tapani (Olstam and Tapani 2004)). The Wiedemann model predicts a space headway of 7.25 meters between two vehicles in a stopped queue. Therefore, any headways greater than between two observed vehicles greater than 14.5 meters, i.e. double the 7.25 meter headway, suggests the presence of an unequipped vehicle and triggers an insertion. After each new vehicle is inserted, the queue is again searched for gaps until none are present.

Step 2: Inserted Vehicle Speed and Location

When a gap is found in the queue, a vehicle is inserted into the queue. The vehicle is inserted at the average stopped following distance predicted from the Wiedemann model of 7.25 meters behind the front of the leading observed vehicle. This assumes that the leading vehicle has a typical passenger car length of 4.75 meters, and the gap between vehicles is 2.5 meters. If the vehicle-to-infrastructure communication system is able to transmit individual vehicle lengths,

then the actual lengths of equipped vehicles can be used in place of assumed values. In a deployment, these values may be calibrated to reflect local conditions.

Step 3: Inserted Vehicle Movements

Once the vehicles are inserted into the gaps, their movements and interactions with other equipped and inserted vehicles can be simulated microscopically throughout the network using the traffic simulation software VISSIM (PTV 2011a). This allows the arterial location estimation algorithm to continue to estimate inserted vehicle positions, including lane changes, even as the inserted vehicle leaves the queue and travels through the network. VISSIM was used for three reasons: it allows users to access individual vehicle's information during the simulation as in a connected vehicle environment; its COM interface allows users to insert and delete individual vehicles without stopping the simulation run; and VISSIM can run multiple interacting simulations simultaneously, allowing the behavior of one model to influence the inputs to the other (PTV 2011b). The ability to run multiple interacting simulations was particularly useful when evaluating the location estimation algorithm's effect on a connected vehicle application, as a single simulation run could represent ground truth while another represented a rolling estimation of vehicle positions.

Step 4: Inserted Vehicle Deletion

Vehicle estimates that have been inserted into the simulation are checked every time step to ensure that their positions have not been overlapped by an equipped vehicle's self-reported position. Once an inserted vehicle has been overlapped, it is no longer considered a correct estimate of an actual unequipped vehicle, and is therefore removed from the simulation. Inserted vehicles that reach the end of the network are also deleted from the simulation.

This technique deviates from a previous effort for freeway vehicles in several respects (Goodall et al. 2014). The freeway algorithm inserts vehicles whenever an equipped vehicle deviates from its expected acceleration by some threshold. Because vehicles on a freeway generally react only to other vehicles, any unusual deceleration can be assumed to be a reaction to an unequipped vehicle. By contrast, vehicles on arterials may decelerate for many reasons such as turning, pedestrians, and traffic control devices, and therefore a deceleration may indicate something other than an unequipped vehicle. Because vehicles on arterials often queue, insertions in the arterial algorithm are made only during stopped queues. Another difference between the freeway and arterial algorithms is the way in which inserted vehicle movements are modeled: freeway insertions use a car-following model for simplicity, while arterial vehicles use the commercial software package VISSIM to capture complicated behaviors such as lane changing and turning movements.

EVALUATION

Testing Environment

The arterial location estimation algorithm was tested on calibrated model of US 50, a four-intersection arterial corridor in Chantilly, Virginia shown in Figure 2.



Figure 2. The arterial test network, four signalized intersections on US 50 in Chantilly, VA.

Vehicle volumes and turning movements were collected in 2003 between 3:00 PM and 4:00 PM on weekdays (Park and Schneeberger 2003). Pedestrian movements, which were very low at these intersections, were ignored in this analysis. The location estimation algorithm was replicated five times in each testing scenario, and results were found to be statistically similar for all measures-of-effectiveness within 10% at a 95% confidence level. The traffic simulation software package VISSIM was used to simulate vehicle movements (PTV 2011a). The arterial algorithm was tested at equipped vehicle penetration rates of 5, 10, 15, 25, 50, and 100%.

The arterial location estimation algorithm requires that inserted vehicles decide which direction to turn at downstream intersections. The following two methods are used to predict turning decisions:

1. **Default (dflt)** Vehicles are assigned a default 10% probability of turning left, a 10% probability of turning right, and an 80% probability of traveling straight through an intersection. This represents reasonable turning movement percentages that, in a field deployment, may have been measured manually, but are neither recent nor real-time and therefore may be slightly inaccurate.
2. **Measured (msrd)** Vehicles are assigned a probability of turning based on actual turning percentages. These percentages could be measured in real-time from connected vehicle data, although this requires the controller record aggregated vehicle movements, a potential violation of driver privacy.

Effective Penetration Rate

Measuring the performance of the location estimation algorithm is challenging. Normally the difference between observed and estimated values can be measured and averaged, which requires a one-to-one relationship between estimates and observations. The location estimation algorithm, by contrast, often has a different number of estimates (inserted vehicles) than observations (unequipped vehicles). Previous research has been forced to aggregate the results into queue length (Ban et al. 2011; Sun and Ban 2013) and vehicle density for discrete road segments (Herrera and Bayen 2010). Other efforts assume that the number of unequipped vehicles is known so that a position estimate can be produced for each unequipped vehicle (Sun and Ban 2011).

In an attempt to measure the performance of location estimation algorithm without equal numbers of estimates and observations, the effective penetration rate (PR_{eff}) metric was recently developed for similar analysis of freeway vehicles (Goodall 2013; Goodall et al. 2014), and is used here. The effective penetration rate metric is defined as the number of equipped vehicles, plus the number of “correct” inserted vehicles, minus the number of “incorrect” inserted vehicles, divided by the total number of equipped and unequipped vehicles, as shown in Equation 1.

$$PR_{\text{eff}} = \frac{\text{No. Equipped Vehicles} + \text{No. Correct Insertions} - \text{No. Incorrect Insertions}}{\text{No. Equipped Vehicles} + \text{No. Unequipped Vehicles}} \quad (1)$$

For example, consider an scenario with 30 equipped vehicles and 70 unequipped vehicles. The penetration rate is $30 / (30 + 70) = 30\%$. There are also 15 inserted vehicles. Ten of the inserted vehicles are in the same lane and within a distance ρ of a unique unequipped vehicle, and are considered correct estimates. (The value of ρ represents the measurement accuracy required by a connected vehicle application.) Five are not within a distance ρ of unique unequipped vehicles, and are considered incorrect estimates. The effective penetration rate would be calculated as $(30 + 10 - 5) / (30 + 70) = 35\%$. The process is repeated each second for each lane. See cited sources for a more detailed formulation of the effective penetration rate (Goodall 2013; Goodall et al. 2014).

RESULTS

Summary statistics of the arterial location estimation algorithm for both default and measured turning movements are provided in Table 2.

Table 2. Summary Statistics of the Location Estimation Algorithm Evaluation

| Penetration rate | 5% | 10% | 15% | 25% | 50% | 100% |
|---|-------|-------|-------|-------|-------|-------|
| Default Turning Movements | | | | | | |
| Number of insertions ¹ | 1116 | 1482 | 1963 | 2119 | 2164 | 1568 |
| Number of equipped vehicles ² | 225 | 449 | 673 | 1126 | 2260 | 4515 |
| Average inserted vehicle lifespan (s) ³ | 74.4 | 78.9 | 76.8 | 72.6 | 66.4 | 17.7 |
| Insertions per unequipped vehicle ⁴ | 0.261 | 0.367 | 0.515 | 0.627 | 0.958 | - |
| Insertions per equipped vehicle ⁵ | 4.96 | 3.30 | 2.92 | 1.89 | 0.96 | 0.35 |
| Equipped vehicle-seconds per insertion ⁶ | 33.6 | 47.3 | 52.5 | 78.8 | 155.8 | 462.3 |
| Measured Turning Movements | | | | | | |
| Number of insertions ¹ | 1078 | 1393 | 1946 | 2145 | 2190 | 1630 |
| Number of equipped vehicles ² | 225 | 447 | 670 | 1122 | 2241 | 4504 |
| Average inserted vehicle lifespan (s) ³ | 74.0 | 72.0 | 79.0 | 72.8 | 63.2 | 17.1 |
| Insertions per unequipped vehicle ⁴ | 0.252 | 0.346 | 0.513 | 0.637 | 0.977 | - |
| Insertions per equipped vehicle ⁵ | 4.78 | 3.12 | 2.90 | 1.91 | 0.98 | 0.36 |
| Equipped vehicle-seconds per insertion ⁶ | 34.4 | 46.0 | 50.3 | 76.0 | 154.9 | 438.1 |

¹ The number of unique instances of inserted vehicles, regardless of how long an inserted vehicle survives. This is roughly equivalent to the number of times an equipped vehicle triggers an insertion.

² The number of vehicles with communications capabilities. This is approximately the total number of vehicles in the scenario multiplied by the penetration rate.

³ The average length of time in seconds that an inserted vehicle moves forward before it is either overlapped by an equipped vehicle and deleted, or reaches the end of the network.

⁴ The total number of unique vehicle insertions divided by the total number of unequipped vehicles.

⁵ The total number of unique vehicle insertions divided by the total number of equipped vehicles. Each equipped vehicle triggers an average of this many insertions during its drive through the network.

⁶ This value represents the average number of vehicles inserted by an equipped vehicle as it travels through the network.

As expected, there is very little difference between the default and measured turning movements. This is because the metrics in Table 2 deal mostly with the number and initial placement of inserted vehicles rather than their later turning decisions. Inserted vehicles typically remain in the simulation for 70-80 seconds, unless the equipped vehicle penetration rate is 100%, at which point inserted vehicles remain for only 17.7 seconds on average.

Figure 3 shows a sample of vehicle trajectories of equipped, unequipped, and inserted vehicles in the network at a 15% equipped vehicle penetration rate. The unequipped and inserted vehicles never exist in the same simulation and therefore never interact with each other, but ideally each unequipped vehicle should be represented by an inserted vehicle at all times. At 1209 seconds, the equipped vehicle joins a stopped queue and inserts five vehicles, placed at even intervals. While many inserted vehicles track closely with unequipped vehicles as they

leave the intersection, some are slow to accelerate and are overlapped and deleted by equipped vehicles, such as at 1215 and 1218 in Figure 3.

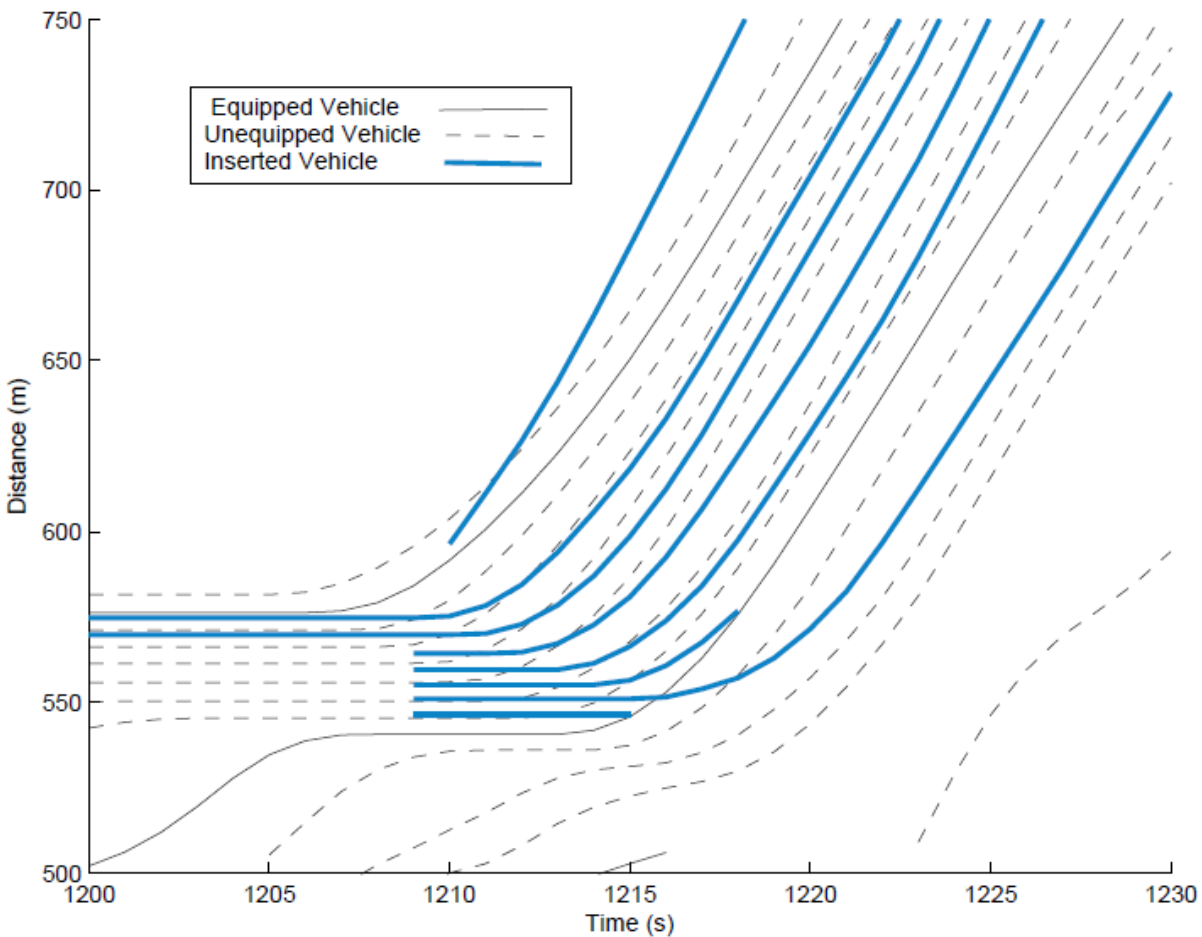


Figure 3. A sample of vehicle trajectories of equipped, unequipped, and inserted vehicles at 15% equipped vehicle penetration rate, on a single lane.

Effective Penetration Rate

As noted earlier, the effective penetration rate PR_{eff} measures the accuracy of the location estimation algorithms. It represents the percentage of all vehicles whose locations are “known”, including both equipped vehicles and unequipped vehicles with a one-to-one relationship with a inserted vehicle within the specified accuracy distance. The effective penetration rate across the entire evaluation network averaged over five runs is shown in Table 3.

Table 3. Effective Penetration Rate over the Entire Test Network Averaged over Five Simulation Runs

| Accuracy Distance ρ (m) | Penetration Rate | | | | | |
|---------------------------------|------------------|-------|-------|-------|-------|-------|
| | 5% | 10% | 15% | 25% | 50% | 100% |
| 1 | -3.3% | -2.5% | -0.9% | 9.8% | 36.9% | 96.2% |
| 2 | -0.6% | 1.4% | 4.1% | 14.3% | 40.8% | 96.2% |
| 3 | 1.3% | 4.2% | 7.6% | 17.6% | 43.4% | 96.2% |
| 4 | 2.7% | 6.2% | 10.1% | 20.0% | 45.3% | 96.2% |
| 5 | 3.8% | 8.0% | 12.4% | 22.1% | 46.9% | 96.2% |
| 6 | 4.9% | 9.6% | 14.4% | 24.0% | 48.4% | 96.2% |
| 7 | 5.8% | 11.0% | 16.2% | 25.6% | 49.7% | 96.2% |
| 8 | 6.6% | 12.2% | 17.8% | 27.1% | 50.8% | 96.2% |
| 9 | 7.3% | 13.3% | 19.2% | 28.4% | 51.9% | 96.2% |
| 10 | 7.9% | 14.3% | 20.4% | 29.5% | 52.7% | 96.2% |
| 11 | 8.4% | 15.1% | 21.4% | 30.4% | 53.5% | 96.2% |
| 12 | 8.8% | 15.7% | 22.2% | 31.3% | 54.2% | 96.2% |
| 13 | 9.3% | 16.4% | 23.0% | 32.0% | 54.8% | 96.2% |
| 14 | 9.6% | 16.9% | 23.6% | 32.7% | 55.4% | 96.2% |
| 15 | 9.9% | 17.5% | 24.3% | 33.3% | 55.8% | 96.2% |

For most penetration rates tested, the algorithm produced more accurate estimates than inaccurate estimates at $\rho = 7$ meters. At 100% penetration rate, every inserted vehicle is by definition incorrect as there are no unequipped vehicles to represent, and therefore the effective penetration rate is 96.2% regardless of ρ . These insertions may be due to vehicles with large space headways in queues or unusual equipped vehicle deceleration. Although the algorithm should never be used at the 100% penetration rate as it produces only noise, in implementation the actual percentage of equipped vehicles may not be known. In some circumstances, penetration rates may approach 100%, while the location algorithm is still running, and therefore it is useful to investigate the error produced by the location estimation algorithm in these circumstances. Ideally, at high penetration rates, the location estimation algorithm would not be activated.

Because vehicles are inserted in stopped queues, one expects effective penetration rate to be higher in the vicinity of a traffic signal. Figure 4 shows the effective penetration rates for several actual penetration rates along the test corridor for eastbound traffic. The vertical dotted lines represent the stop bar immediately prior to the traffic signal. The minimum estimate accuracy ρ is 10 meters in this example.

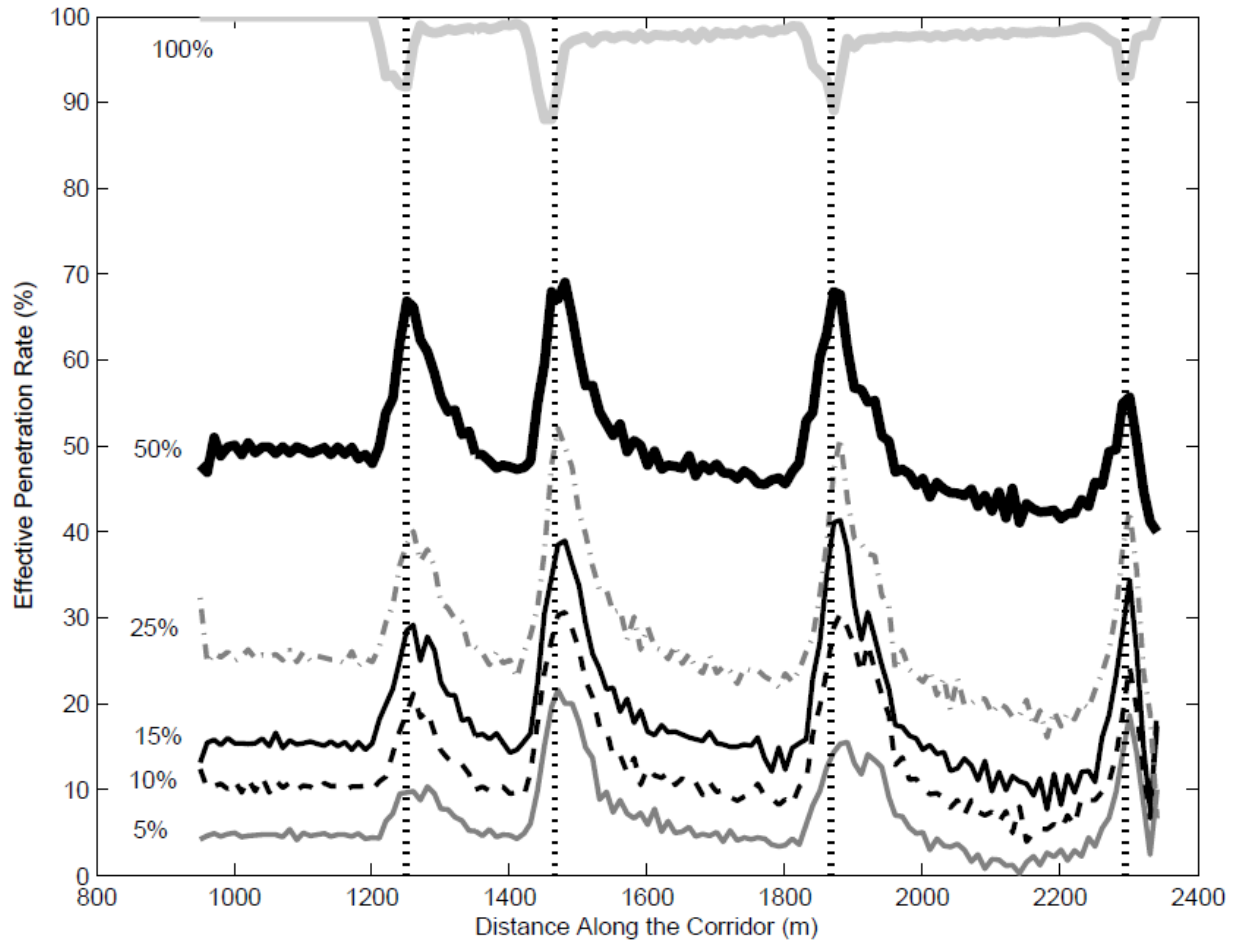


Figure 4. Effective penetration rates at different points along the corridor for eastbound traffic. Vertical dotted lines represent signalized intersections. The minimum estimate accuracy is 10 meters.

As expected, effective penetration rate is highest near the signal, as inserted vehicles are added into the queue. Effective penetration rate drops to approximately the previous penetration rate at 150 meters downstream of the intersection. One possible cause of this decrease is the platoon dispersion. When vehicles are no longer clustered in a queue, fewer vehicles remain within the minimum required estimate accuracy ρ (which is a distance rather than a headway time). Therefore the increased effective penetration rates near signals may be partially due to lower vehicle speeds at these locations and ρ representing a space headway rather than a time headway.

Traffic Signal Control Application

The ultimate objective of the arterial location estimation algorithm is to improve the performance of connected vehicle mobility applications on arterials. Because these applications perform better at higher equipped vehicle penetration rates, the research question is whether an estimate of individual vehicle locations can artificially augment the actual penetration rate and improve performance. To answer this question, the arterial location estimation algorithm is applied to the Predictive Microscopic Simulation Algorithm (PMSA) (Goodall et al. 2013).

The PMSA is a traffic signal control algorithm that observes individual vehicle locations, and predicts their behavior over a 15-second horizon using a commercial traffic simulation software package. The movements over the horizon are repeated for several possible signal phase configurations, and the scenario with the optimized objective function (in this case, minimized delay) is selected as the next phase. This process is repeated continuously. Previous testing of the PMSA in simulation has shown that higher equipped vehicle penetration rates improves the PMSA's performance, and that the PMSA begins to outperform a coordinated-actuated control system at between 10% and 25% penetration rate.

The PMSA was evaluated in simulation on a model of a four-signal arterial shown in Figure 2 using the field-collected volumes and turning movements (Park and Schneeberger 2003). Under the tested volumes, the network has an average intersection capacity utilization (ICU) of 0.75, which is a metric used by the traffic signal timing software Synchro as a surrogate for volume-capacity ratio (Husch and Albeck 2003). Equipped vehicle penetration rates of 5, 10, 15, 25, 50, and 100% were tested five times each for 30 minutes with four hundred seconds of warm up. These scenarios are referred to in the chart as equipped vehicles-only (EV-only). Additionally, both inserted vehicle turning strategies were tested: the default turning movements (LE-dft) where inserted vehicles at all intersections have a 10% probability of turning left and a 10% probability of turning right, and the field-measured turning movements (LE-msrd) where a vehicle's probability of turning is equivalent to measured turning probabilities specific to the approach.

Table 4 shows the performance of the location estimation algorithm (both using the default and measured turning movements) when compared to the PMSA using equipped vehicles only. The table shows the average delay, average speed, average stopped delay, and average total stops of all three alternatives, as well as the percentage differences between these values and the significance of these differences.

Both of the location estimation algorithms are able to significantly improve the performance of the PMSA during certain situations. Improvements were found in delay, speed, and stopped delay when penetration rate was 25% or less. Improvements in stops at these penetration rates were very small or nonexistent. The improvements in delay, speed, and stopped delay were only significant ($P < 0.05$) for two-tailed two-sample T test assuming unequal variances) at 10% and 25% penetration rate. P-values were less than 0.3 at 15% penetration rate.

The location estimation algorithm substantially worsened the performance of the PMSA at higher penetration rates of 50% and 100%, with up to 14% increases in stops at 100% penetration rate. This may be due to the poor performance of the PMSA at high volumes where low volume side streets are rarely assigned green unless they reach 120 second of red time (Goodall 2013). When this occurs, the PMSA essentially behaves as an uncoordinated pretimed plan, which, although effective at isolated intersections, fails to react to the platoon arrivals in a signalized corridor. The location estimation algorithm, by estimating more vehicles than are actually present when used at high penetration rates, may encourage this effect. Regardless of the reason, the evidence suggests that the arterial location estimation algorithm should not be used at penetration rates of 50% or higher. The algorithm produces an improvement at 25% penetration rate and lower. Although the precise transition point between improved and worsened performance was not identified, for this application the transition occurs between 25% and 50% penetration rate. Recent research suggests that the penetration rate can be estimated based on the ratio of the number of equipped vehicles and queue length (Venkatanarayana et al. 2011). In a deployment, the location estimation algorithm could be turned off if the high penetration rates are estimated.

Table 4. Performance of the PMSA When Augmented with the Data from the Location Estimation Algorithm

| PR (%) | Technique | Delay (s) | EV Diff | P-value | LE-dft Diff | P-value | PR (%) | Technique | Speed (mi/hr) | EV Diff | P-value | LE-dft Diff | P-value |
|--------|-----------|-----------|---------|---------|-------------|---------|--------|-----------|---------------|---------|---------|-------------|---------|
| 5 | EV-only | 62.6 | | | | | 5 | EV-only | 26.6 | | | | |
| | LE-dft | 61.7 | -1.37% | 0.713 | | | | LE-dft | 26.8 | 0.82% | 0.588 | | |
| | LE-msrd | 61.5 | -1.76% | 0.327 | -0.40% | 0.918 | | LE-msrd | 26.8 | 0.99% | 0.252 | 0.16% | 0.918 |
| 10 | EV-only | 54.1 | | | | | 10 | EV-only | 28.0 | | | | |
| | LE-dft | 52.7 | -2.64% | 0.164 | | | | LE-dft | 28.3 | 1.19% | 0.116 | | |
| | LE-msrd | 51.6 | -4.71% | 0.040 | -2.13% | 0.343 | | LE-msrd | 28.5 | 1.88% | 0.037 | 0.68% | 0.424 |
| 15 | EV-only | 50.0 | | | | | 15 | EV-only | 28.8 | | | | |
| | LE-dft | 49.1 | -1.78% | 0.310 | | | | LE-dft | 29.0 | 0.90% | 0.121 | | |
| | LE-msrd | 48.3 | -3.38% | 0.280 | -1.63% | 0.598 | | LE-msrd | 29.2 | 1.40% | 0.162 | 0.50% | 0.587 |
| 25 | EV-only | 47.1 | | | | | 25 | EV-only | 29.3 | | | | |
| | LE-dft | 45.2 | -4.04% | 0.015 | | | | LE-dft | 29.8 | 1.57% | 0.031 | | |
| | LE-msrd | 45.0 | -4.41% | 0.025 | -0.38% | 0.835 | | LE-msrd | 29.8 | 1.65% | 0.003 | 0.08% | 0.898 |
| 50 | EV-only | 45.2 | | | | | 50 | EV-only | 29.7 | | | | |
| | LE-dft | 46.4 | 2.66% | 0.057 | | | | LE-dft | 29.5 | -0.62% | 0.130 | | |
| | LE-msrd | 46.6 | 3.12% | 0.015 | 0.45% | 0.633 | | LE-msrd | 29.5 | -0.63% | 0.109 | -0.01% | 0.987 |
| 100 | EV-only | 48.3 | | | | | 100 | EV-only | 29.1 | | | | |
| | LE-dft | 52.8 | 9.32% | 0.000 | | | | LE-dft | 28.2 | -2.81% | 0.000 | | |
| | LE-msrd | 52.9 | 9.39% | 0.013 | 0.06% | 0.981 | | LE-msrd | 28.2 | -2.87% | 0.007 | -0.06% | 0.937 |

| PR (%) | Technique | Stopped Delay (s) | EV Diff | P-value | LE-dft Diff | P-value | PR (%) | Technique | Stops | EV Diff | P-value | LE-dft Diff | P-value |
|--------|-----------|-------------------|---------|---------|-------------|---------|--------|-----------|-------|---------|---------|-------------|---------|
| 5 | EV-only | 37.1 | | | | | 5 | EV-only | 6223 | | | | |
| | LE-dft | 35.4 | -4.56% | 0.409 | | | | LE-dft | 6493 | 4.33% | 0.255 | | |
| | LE-msrd | 35.5 | -4.32% | 0.235 | 0.26% | 0.967 | | LE-msrd | 6437 | 3.44% | 0.232 | -0.85% | 0.824 |
| 10 | EV-only | 29.6 | | | | | 10 | EV-only | 5707 | | | | |
| | LE-dft | 27.6 | -6.80% | 0.042 | | | | LE-dft | 5727 | 0.36% | 0.883 | | |
| | LE-msrd | 27.4 | -7.64% | 0.033 | -0.91% | 0.807 | | LE-msrd | 5633 | -1.30% | 0.474 | -1.65% | 0.476 |
| 15 | EV-only | 26.2 | | | | | 15 | EV-only | 5349 | | | | |
| | LE-dft | 25.3 | -3.21% | 0.214 | | | | LE-dft | 5397 | 0.91% | 0.624 | | |
| | LE-msrd | 24.8 | -5.14% | 0.263 | -2.00% | 0.665 | | LE-msrd | 5278 | -1.32% | 0.675 | -2.21% | 0.480 |
| 25 | EV-only | 24.5 | | | | | 25 | EV-only | 4997 | | | | |
| | LE-dft | 22.6 | -8.04% | 0.000 | | | | LE-dft | 4997 | -0.02% | 0.993 | | |
| | LE-msrd | 22.5 | -8.20% | 0.008 | -0.17% | 0.945 | | LE-msrd | 4973 | -0.48% | 0.800 | -0.47% | 0.819 |
| 50 | EV-only | 23.9 | | | | | 50 | EV-only | 4680 | | | | |
| | LE-dft | 24.0 | 0.15% | 0.950 | | | | LE-dft | 4967 | 6.12% | 0.050 | | |
| | LE-msrd | 23.8 | -0.58% | 0.713 | -0.73% | 0.719 | | LE-msrd | 4961 | 6.01% | 0.002 | -0.10% | 0.964 |
| 100 | EV-only | 26.4 | | | | | 100 | EV-only | 4843 | | | | |
| | LE-dft | 28.5 | 7.67% | 0.063 | | | | LE-dft | 5521 | 13.98% | 0.000 | | |
| | LE-msrd | 28.8 | 9.11% | 0.015 | 1.33% | 0.719 | | LE-msrd | 5457 | 12.68% | 0.073 | -1.14% | 0.825 |

Performance of Default vs. Measured Turning Volumes

Two versions of the arterial location estimation algorithm were tested. The first uses default static turning percentage to assign probabilities that inserted vehicles will turn at downstream intersections. The default percentages are 10% probability of turning right, 10% probability of

turning left, and 80% probability of traveling straight through. The second version of the location estimation algorithm uses the field-measured turning percentages to assign probabilities.

From Table 4, the differences between measured and default turning percentages were statistically insignificant. It is worth noting that although the significance is small ($P > 0.34$ in all cases), the measured turning movements outperformed default turning movements in all measures-of-effectiveness when the penetration rate is 25% or smaller. It is possible that the benefits of measured turning movements were small due to the characteristics of the test network. In the model of US 50 used in the evaluation, there were few lanes with shared turns (i.e. through and left or through and right combined lanes), and the few that did had low volumes. Most vehicles inserted into a queue had only one choice of where to turn, and the turning movement decisions affected only their decisions at downstream intersections. On a network with many shared turn lanes, such as an urban network, the turning decision becomes much more important. On these networks, the measured turning movements would likely outperform the default turning movements. More research is needed in order to be certain.

CONCLUSIONS

This research represents one of the first attempts to estimate the positions of individual unequipped vehicles on arterials. Previous attempts made several admittedly restrictive assumptions in their estimates, such as uniform vehicle arrival, unrealistic acceleration rates, and known equipped vehicle penetration rates. The methods presented in this paper avoid these assumptions, and additionally estimate vehicle positions beyond a single intersection. The results showed that it is possible to estimate the locations of some unequipped vehicles on an arterial with a reasonable degree of accuracy given necessary data from connected vehicles. On arterial through movements, more estimates are correct than incorrect using a distance error of approximately 7 meters, and accuracy is substantially higher within 150 meters of an intersection.

Similarly, this research was the first to demonstrate that the estimation of unequipped vehicle positions on arterials could improve the performance of a connected vehicle mobility application in simulation, in this case using the PMSA traffic signal control strategy as an example. When using data from the location estimation algorithm, the PMSA experienced slight improvements in delay, stopped delay, and speed at penetration rates between 10% and 25% when compared to using equipped vehicle data only. No differences were found between assigning turning probabilities to inserted vehicles based on field-measured turning movements or default turning movements. The improvement of the PMSA suggests that the location estimation technique may prove beneficial to other connected vehicle algorithms such as queue length estimation for transit signal priority, ramp metering at low penetration rates, or other traffic signal control strategies such as PAMSCOD. The results were limited to simulation studies, and therefore assumed high-quality measurements of connected vehicle position, speed, and acceleration. Additional research is needed to determine if the results found in simulation will translate into real-world benefits.

Future research should investigate calibration of the algorithm's default values used such as the 50-meter queue length window and vehicle length, integration with other queue estimation techniques, the performance of the algorithm on a larger grid network, the performance of the algorithm in a real-world deployment, and the benefits of the algorithm when applied to a candidate application in a connected vehicle pilot deployment.

REFERENCES

- Ban, X. (Jeff), Hao, P., and Sun, Z. (2011). "Real Time Queue Length Estimation for Signalized Intersections Using Travel Times from Mobile Sensors." *Transportation Research Part C: Emerging Technologies*, 19(6), 1133–1156.
- Bargera, H. (2007). "Evaluation of a Cellular Phone-Based System for Measurements of Traffic Speeds and Travel Times: A Case Study from Israel." *Transportation Research Part C: Emerging Technologies*, 15(6), 380–391.
- Barria, J. A., and Thajchayapong, S. (2011). "Detection and Classification of Traffic Anomalies Using Microscopic Traffic Variables." *IEEE Transactions on Intelligent Transportation Systems*, 12(3), 695–704.
- Federal Highway Administration. (2010). "Next Generation Simulation (NGSIM)." <<http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm>> (Nov. 15, 2010).
- Goodall, N. J. (2013). "Traffic Signal Control with Connected Vehicles." Doctoral dissertation, University of Virginia, Charlottesville, VA.
- Goodall, N. J., Smith, B. L., and Park, B. (2013). "Traffic Signal Control with Connected Vehicles." *Transportation Research Record: Journal of the Transportation Research Board*, 2381, 65–72.
- Goodall, N. J., Smith, B. L., and Park, B. (2014). "Microscopic Estimation of Freeway Vehicle Positions from the Behavior of Connected Vehicles." *Journal of Intelligent Transportation Systems*, in press.
- He, Q., Head, K. L., and Ding, J. (2012). "PAMSCOD: Platoon-Based Arterial Multi-Modal Signal Control with Online Data." *Transportation Research Part C: Emerging Technologies*, 20, 164–184.
- Herrera, J. C., Andrews, S., Apte, S., Arnold, J., Ban, J., Benko, M., Bayen, A. M., Chiou, B., Claudel, C., Claudel, C., Dodson, T., Elhamshary, O., Flens-Batina, C., Gruteser, M., Amin, S., Herring, R., Hoh, B., Jacobson, Q., Kumar, M., Iwuchukwu, T., Lew, J., Litrico, X., Luddington, L., Margulici, J., Mortazavi, A., Pan, X., Rabbani, T., Racine, T., Sherlock-Thomas, E., Sutter, D., Tinka, A., Tracton, K., Tossavainen, O.-P., West, T., Wiedmer, A., Work, D. B., and Wu, Q. (2008). "Mobile Century - Using GPS Mobile Phones as Traffic Sensors: A Field Experiment." ITS America, New York, NY.
- Herrera, J. C., and Bayen, A. M. (2010). "Incorporation of Lagrangian Measurements in Freeway Traffic State Estimation." *Transportation Research Part B: Methodological*, 44(4), 460–481.
- Husch, D., and Albeck, J. (2003). "Intersection Capacity Utilization." *Trafficware*, <<http://www.trafficware.com/assets/pdfs/ICU2003.pdf>> (Mar. 5, 2013).
- John A. Volpe National Transportation Systems Center. (2008). *Vehicle-Infrastructure Integration (VII) Initiative Benefit-Cost Analysis Version 2.3 (Draft)*. Federal Highway Administration, Washington, DC.
- Krause, A., Horvitz, E., Kansal, A., and Zhao, F. (2008). "Toward Community Sensing." *2008 International Conference on Information Processing in Sensor Networks*, St. Louis, MO, USA, 481–492.
- Li, M., Zou, Z., Bu, F., and Zhang, W.B. (2008). "Application of Vehicle Infrastructure Integration Data on Real-Time Arterial Performance Measurements." *Transportation Research Board 87th Annual Meeting*, Transportation Research Board, Washington, DC.
- Mehran, B., Kuwahara, M., and Naznin, F. (2012). "Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data." *Transportation Research Part C: Emerging Technologies*, 20(1), 144–163.

- Nanthawichit, C., Nakatsuji, T., and Suzuki, H. (2003). "Application of Probe-Vehicle Data for Real-Time Traffic-State Estimation and Short-Term Travel-Time Prediction on a Freeway." *Transportation Research Record: Journal of the Transportation Research Board*, 1855, 49–59.
- Olstam, J. J., and Tapani, A. (2004). *Comparison of Car-following Models*. Swedish National Road and Transport Research Institute, Linköping, Sweden.
- Park, B., and Schneeberger, J. (2003). "Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System." *Transportation Research Record: Journal of the Transportation Research Board*, 1856, 185–192.
- Priemer, C., and Friedrich, B. (2009). "A Decentralized Adaptive Traffic Signal Control Using V2I Communication Data." *2009 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO, 1–6.
- PTV. (2011a). *VISSIM 5.40-01 User Manual*. Planung Transport Verkehr AG.
- PTV. (2011b). *VISSIM 5.40-01 COM Interface Manual*. Planung Transport Verkehr AG.
- Rim, H., Oh, C., Kang, K., and Kim, S. (2011). "Estimation of Lane-Level Travel Times in Vehicle-to-Vehicle and Vehicle-to-Infrastructure-Based Traffic Information System." *Transportation Research Record: Journal of the Transportation Research Board*, 2243, 9–16.
- Sanwal, K., and Walrand, J. (1995). *Vehicles as Probes*. Institute of Transportation Studies, University of California Berkeley, Berkeley, CA.
- Smith, B. L., Venkatanarayana, R., Park, H., Goodall, N. J., Datesh, J., and Skerrit Jr., C. (2011). *IntelliDrive Traffic Signal Control Algorithms*. Cooperative Transportation Systems Pooled Fund Study, Charlottesville, VA.
- Sun, Z., and Ban, X. (2011). "Vehicle Trajectory Reconstruction for Signalized Intersections Using Variational Formulation of Kinematic Waves." *Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington, DC.
- Sun, Z., and Ban, X. (Jeff). (2013). "Vehicle trajectory reconstruction for signalized intersections using mobile traffic sensors." *Transportation Research Part C: Emerging Technologies*, 36, 268–283.
- Venkatanarayana, R., Park, H., Smith, B. L., Skerrit Jr., C., and Ruhter, N. W. (2011). "Application of IntelliDrive to Address Oversaturated Conditions on Arterials." *Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington, DC.
- Westerman, M., Litjens, R., and Linnartz, J. P. (1996). *Integration of Probe Vehicle and Induction Loop Data-estimation of Travel Times and Automatic Incident Detection*. Institute of Transportation Studies, University of California, Berkeley, Berkeley, CA.
- Wiedemann, R., and Reiter, U. (1992). *Microscopic Traffic Simulation: The Simulation System MISSION, Background and Actual State*. CEC, Brussels.