Probe vehicle lane identification for queue length estimation at intersections

by Semuel Rompis 21

Journal of Intelligent Transportation Systems
Technology, Planning, and Operations

ISSN: 1547-2450 (Print) 1547-2442 (Online) Journal homepage: http://www.tandfonline.com/loi/gits20

Probe vehicle lane identification for queue length estimation at intersections

Semuel Y. R. Rompis, Mecit Cetin & Filmon Habtemichael

JOURNAL OF INTELLIGENT TRANSPORTATION SYSTEMS
2017, VOL. 0, NO. 0, 1–16
http://dx.doi.org/10.1080/15472450.2017.1300887

Taylor & Francis
Taylor & Francis Group

Probe vehicle lane identification for queue length estimation at intersections

Semuel Y. R. Rompis[a,b], Mecit Cetin[a], and Filmon Habtemichael[a]

[a]Transportation Research Institute, Old Dominion University, Norfolk, VA, USA; [b]Department of Civil Engineering, Sam Ratulangi University, Manado, North Sulawesi, Indonesia

ABSTRACT
Vehicles instrumented with Global Positioning Systems, also known as GPS probe vehicles, have become increasingly popular for collecting traffic flow data. Previous studies have explored the probe vehicle data for estimating speeds and travel time; however, there is very limited research on predicting queue dynamics from such data. In this research, a methodology was developed for identifying the lane position of the GPS-instrumented vehicles when they are standing in the queue at signalized intersections with multiple lanes, particularly in the case of unequal queue. Various supervised and unsupervised clustering methods were tested on data generated from a microsimulation model. Among the tested methods, the Optimal Bayes Rule that utilizes probability density functions estimated using bivariate statistical mixture models was found to be effective in identifying the lanes. The methodology for lane identification was tested for queue length estimation. This research confirms that the lane identification is an important step required prior to the queue length estimation. The accuracies of the models for lane identification and queue length estimation were evaluated at varying levels of demand and probe vehicle market penetrations. In general, as the market penetration increases, the accuracy improves as expected. The result shows that 40% market penetration rate is adequate to reach about 90% accuracy.

ARTICLE HISTORY
Received 7 February 2016
Revised 23 August 2016
Accepted 17 February 2017

KEYWORDS
GPS probe vehicles; lane identification; market penetration rate; queue length estimation

Introduction

Instrumented vehicles, known as probe vehicles, have become increasingly popular due to the advancements in tracking, communications, and mobile computing technologies. Among several probe vehicle systems, the Global Positioning System (GPS) probe vehicle systems have been recognized as a very efficient way to collect real-time traffic data, especially at the network level (Turner, Eisele, Benz, & Holdener, 1998). This system has become more feasible since most portable devices such as smartphones and GPS navigation systems can be used as a platform to collect vehicle trajectory data. Reliable processing of such large amounts of real-time data from these GPS systems is a significant challenge and is critical for optimizing transportation systems, e.g., signal optimization and early incident detection.

Researchers have conducted various studies to obtain travel information using probe vehicle data, e.g., estimating queue lengths at signalized intersections. In this context, several approaches have been adopted, e.g., shockwave theory (Badillo, Rakha, Rioux, & Abrams, 2012; Cetin, 2012; Hao, Ban, & Whon Yu, 2015; Ramezani & Geroliminis, 2013), probe vehicles' travel time (Ban, Hao, & Sun, 2011), and probabilistic probe vehicle-based positioning (Comert & Cetin, 2009, 2011). Beside queue length estimation, other applications of these studies were estimating vehicles' fuel consumption and emissions (Cetin & Rakha, 2014).

Even though GPS-based probe vehicles were used for a number of applications, several issues have been addressed with regard to their utilization—the most essential being GPS positioning error. It is well known that the GPS instruments are expected to produce location errors in the range of 3–15 m (Cheng, Qin, Jin, Ran, & Anderson, 2011; Hunter, Wu, Kim, & Suh, 2012; Turner et al., 1998). Such magnitude of error does not allow lane-based analysis of traffic dynamics as it is difficult to unambiguously determine the probe vehicles' lane. An example of lane-based traffic analysis is characterizing queuing dynamics at a signalized intersection where the queue lengths of adjacent lanes are unequal. This study attempts to predict the lane of probe vehicles and also estimates the queue lengths on the adjacent lanes of an intersection approach.

The remainder of the paper is organized as follows. Following this introductory section, the research questions

CONTACT Semuel Y. R. Rompis semrompis@fulbrightmail.org; semrompis@unsrat.ac.id Department of Civil Engineering, Sam Ratulangi University, Manado 95162, North Sulawesi, Indonesia.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/gits.

that this paper addresses are discussed. This is followed by review of literature and discussion on methodological approach of the research. Finally, results and conclusions are presented with directions on future research.

Problem definition

Existing probe vehicle studies, to the best of the authors' knowledge, do not address prediction of the probe vehicles' lane. In the current literature, researchers either assume queuing on a single lane or consider equal parameters (queue length or emissions) for all lanes, which may not be ideal for real-life situations. Unequal queuing in adjacent lanes, including cases where there is no vehicle at all in either lane, is a fairly common phenomenon at many signalized intersections. In particular, unequal queuing is normal when one of the lanes serves shared movements (e.g., shared right-turn and through movement) or when the lanes serve traffic destined for different downstream points and demand for the lanes is unbalanced.

This study is an attempt to bridge the gap in the literature by investigating different modeling approaches to infer the lane of probe vehicles. Figure 1 shows trajectories of probe vehicles forming unequal queues on a two-lane intersection approach. The data for this figure were generated using Planung Transport Verkehr's (PTV) transportation simulation software *Vissim* (Vissim, n.d.). The blue trajectories represent vehicles in lane 1, while red trajectories represent vehicles in lane 2. From the figure, it is clear that the queuing dynamics are not the same in both lanes, i.e., one is long queue lane (LQL) while the other is short queue lane (SQL), having faster and slower queuing shockwaves, respectively. Therefore, the research questions that this study addresses are, given unequal queues at an intersection approach, how can the lanes containing probe vehicles be identified on a real-time basis and how can the queue lengths on the individual lanes be estimated?

This research focuses on undersaturated traffic conditions. This is due to the fact that unequal queues across adjacent lanes typically take place in such traffic conditions. In case of an oversaturated condition, normally there will be left-over/residual queue and thus the queue length in the adjacent lanes can be assumed to be similar. If the shockwave speeds of the lanes are similar or the same, it can be assumed that the queue lengths in each of the lanes are equal and there is no need for identifying the lane position of the probe vehicles. Several methods based on traffic flow theory and statistical learning techniques were utilized to meet the objectives of this research. The traffic flow theories are developed under ideal traffic situations, which make it prone to bias when applied to real-life traffic dynamics. On the other hand, statistical models are more suitable to account for uncertainties. Therefore, a combination of traffic flow theory and probabilistic learning techniques is chosen to account for the uncertainties related to traffic dynamics and the GPS probe system while keeping the principle of traffic flow theory.

This study utilized data generated from PTV's transportation simulation software *Vissim*. It is true that there are some differences between the real driving behavior data and the ones obtained from a simulation environment. However, modern traffic simulation tools, like *Vissim*, can fairly represent the real-world traffic conditions and data. Specifically, the authors chose to conduct this study in a simulation environment due to the following reasons: 1) Probe vehicle location data can be

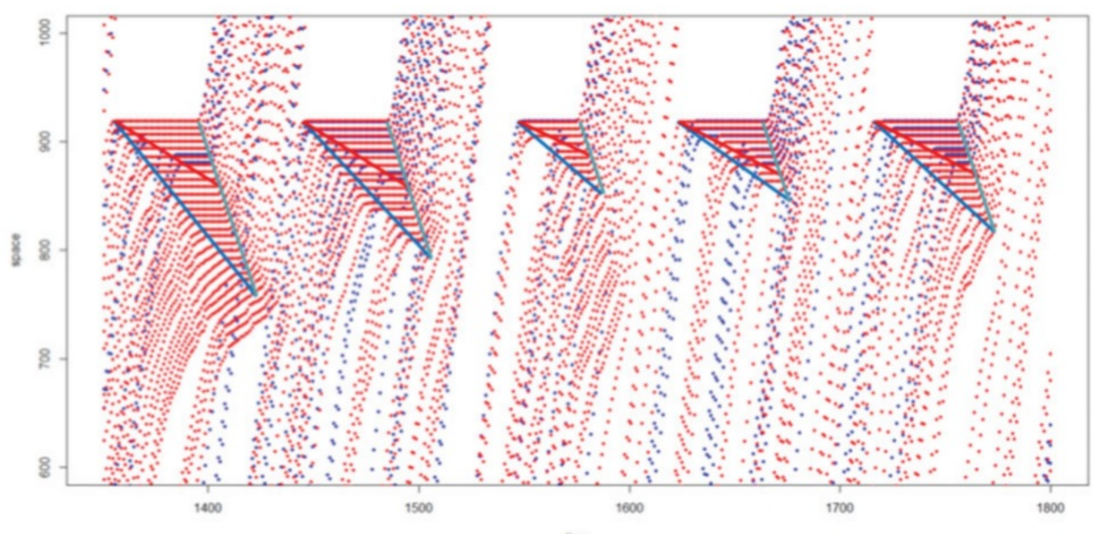

Figure 1. Vehicle trajectories show the different SW speeds of different lanes.

easily obtained from simulation as opposed from real world; 2) The proportion of probes (market penetration) and arrival rates can be easily adjusted; and 3) simulation provides a controlled environment and confounding variables (e.g. weather) can be removed from the analysis.

Related studies

Queue length is a crucial parameter in evaluating and optimizing traffic controls such as traffic signals. Several studies have been conducted to model queues at intersections or at metered on-ramps. The most common technique of modeling queues is by using LWR shockwave theory developed by Lighthill and Witham (1955) and Richard (1956), for example, the study by Liu, Wu, Ma, and Hu (2009). In their study, the intersection queue length was measured using the queue discharge process in the immediate previous cycle. By employing the LWR shockwave theory, they were able to separate the queue discharge flow state from the upstream arrival traffic state. Hao et al. (2015) developed a kinematic equation-based methodology for estimating the location of vehicles at signalized intersection by using mobile sensor data. The methodology was also able to estimate the acceleration rate of vehicles during queue discharging process. Similarly, Wang, Cai, Wu, Zheng, and Wang (2017) used shockwave theory for estimating queue length at signalized intersections by fusing data from multiple sources, such as probe vehicles and point detectors, under both over- and undersaturated conditions.

Kalman filter approach has also been applied to analyze queuing at intersections. Vigos, Papageorgiou, and Wang (2008) employed a Kalman filter to estimate the number of vehicles in signalized links based on online measurement of flow and occupancy from loop detectors. Wu, Jin, and Horowitz (2008) tested three methods for estimating queue length at metered on-ramps: Kalman filter, linear occupancy, and the HCM 2000 back-of-queue model. They concluded that the Kalman filter and linear occupancy methods are practical for real-world operations even though they have limitations, while the HCM 2000 back-of-queue model did not reliably estimate the on-ramp queue length. Others have also attempted to use optimization methods, e.g., Chang and Su, (1995). In their study, they employed the neural network models to predict the intersection queue. They claimed that starting from 3 time-steps ahead, their model was capable of providing more than 90% accuracy.

Similarly, Markov model of decision analysis has also been applied for correlation to queue length estimation at intersections. Geroliminis and Skabardonis (2005) modeled traffic between successive traffic signals as a two-step Markov decision process, while the LWR theory was used to model the traffic dynamic. Then they used this approach to estimate queue lengths and predict travel times. They claimed their model was usable for cases where the loop detector data are unknown, inaccurate, or aggregated. Another study for modeling queues at an intersection using a Markov model was conducted by Viti and van Zuylen (2004). In their study, they constructed a Markov model to calculate the dynamics of the queue. They found that the model introduced in the study was suitable for solving dynamic assignment problems.

The queue length modeling studies described above were using traffic data obtained from fixed sensor as an input. One of the first research for queue length estimation using probe vehicle data was conducted by Comert and Cetin (2009). To estimate the expected queue length and its variance, they developed an analytical formulation based on conditional probability distribution. They also discussed the effects of probe market penetration rate on the accuracy of their model. Using their work, one can estimate the queue length at a signalized intersection by knowing only the location information of the last probe vehicle in the queue. In a later study (Comert & Cetin, 2011), they developed formulations to quantify the error in queue length estimation using probe vehicle data. Similarly, Viti and Van Zuylen (2009) have developed a probabilistic approach of modeling the dynamics and uncertainty of vehicle queues at intersections controlled by fixed and actuated signals.

An attempt to characterize the shockwave profile using probe vehicle data was conducted by Ramezani and Geroliminis (2013). In their paper, they constructed the queuing profile by separating the input data into stopping and moving vehicles, and clustered the stopped vehicles into cycles. They claimed their method was helpful for identifying spillback, constructing vehicle trajectories, and estimating fuel consumption and emissions. Moreover, other studies that used probe vehicles focused on identifying the critical points in a LWR shockwave profile (Cetin, 2012; Cheng, Qin, Jin, & Ran, 2012) and estimating travel times (Ban et al., 2011). The study by Comert (2016) assumes a point queue model where space (or distance) is not explicitly accounted for. Instead, position in the queue (order of arrival) is assumed to be known.

The studies listed above do not address identifying the travel lane of the probe vehicles. As discussed before, it is difficult to unambiguously determine a probe vehicle's lane due to the GPS positioning error. This study is an effort to identify the probe vehicle's lane and thus conduct lane-based queue length estimation.

Queuing shockwave speeds from probe vehicles

The boundary between different traffic regions, e.g., transition from free flowing and queuing or vice versa, are known as shockwaves. The time and space coordinates of probe vehicles joining the queue define the queuing shockwave profile. The back of the queuing shockwave propagates backward from the stop bar at a signalized intersection with a speed that is equal to the slope of the line forming the profile. The queuing shockwave profile could be extracted from the trajectories of the probe vehicles data. One crucial factor in determining the queuing shockwave speed is the arrival rate of vehicles. Therefore, the queuing shockwaves, and thus the queue lengths, of two adjacent lanes will be significantly different when the arrival rates of the lanes are significantly different from one another as indicated in Figure 1.

In an ideal situation, where the vehicle arrival rate is assumed to be constant, or the headways between vehicles are identical, the queuing shockwave profile will be linear. However, in the real-life situation, this is not typically the case due to the random nature of vehicle arrivals.

Considering the objective of this study is to identify the lane of each probe vehicle, the queuing shockwave speed computed from individual probe vehicle in a given cycle is used as an input parameter in distinguishing the lane of the probe vehicle. Individual queuing shockwave speed is defined as the slope of a line connecting the time and space coordinates of the probe vehicle when it joins the queue to the point at the stop bar when signal turns red. The illustration of these shockwave speeds can be seen in Figure 2.

Mathematically, the individual queuing shockwave for all the probe vehicles can be calculated as shown in Eq. (1). Given the traffic dynamics of an intersection approach with unequal queues in the adjacent lanes are different, the distribution of the individual shockwave speeds corresponding to each lane is analyzed to discriminate the lane of the probe vehicles.

$$w_i^c = \frac{X_i^c - X}{t_i^c - t^c} \tag{1}$$

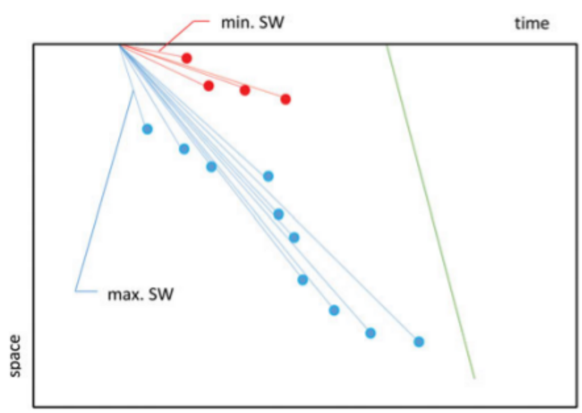

Figure 2. Individual shockwave speeds.

where $i = 1, 2, \ldots, n$; n is the number of probe vehicles in the cth cycle; w_i^c is the individual queuing shockwave speed at the cth cycle; X_i^c is position of an individual vehicle when it joins a queue at the cth cycle; X is the position of the traffic signal; t_i^c is the time of an individual vehicle when it joins the queue at the cth cycle; and t^c is the time when the signal turns red at the cth cycle. The next section describes the datasets assembled for model development and evaluation.

Traffic simulation scenarios and data

An isolated intersection with a two-lane approach road which is controlled by a traffic signal consisting of just red and green lights with durations of 45 seconds each, i.e., a cycle time of 90 seconds, was simulated in *Vissim*. The desired speed of the vehicles was set to be 60 km/h, and lane change was restricted. To have representative data, the simulation was run for 100 cycles, equivalent to 9,000 seconds. The stop coordinates of the simulated vehicles were extracted from their trajectories and further analyzed. Throughout this paper, traffic dynamics on only one arm of the intersection were considered for analysis.

Arrival rates scenarios

This study considers three scenarios of lane-based arrival rates to form unequal queues on adjacent lanes: (1) 300 and 900 vph (Scenario 1), (2) 450 and 900 vph (Scenario 2), and (3) 600 and 900 vph (Scenario 3). All scenarios represent undersaturated traffic flow conditions, as unequal queues are unlikely to be formed during saturated conditions.

Data extraction

The vehicle trajectories data used in this study are illustrated in Figure 3. The coordinates where a vehicle joined the queue are shown by the red dots. A vehicle is

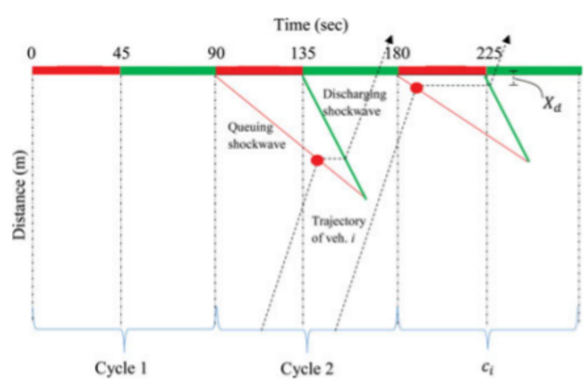

Figure 3. Trajectories of vehicles at a signalized intersection.

assumed to have joined a queue when it is at standstill, i.e., speed = 0 km/h. Similarly, a vehicle is assumed to have started discharging queue when it is no longer at standstill, i.e., speed > 0 km/h. The reason that a speed of 0 km/h is taken as a threshold is that this study mainly focuses on signalized intersections where the vehicle speeds are generally low and their spacing is small.

Market penetration scenarios

The market penetration rate of probe vehicles, i.e., the proportion of probe vehicle relative to all vehicles on the road, is an important component in probe vehicle-based studies. To address this issue, the probe market penetration rate was varied from 10% to 100% by increments of 10%. The different market penetration rates could identify the percentage of probe vehicles needed to have optimal analysis.

Unsupervised learning methods

Naïve method

As discussed before, in the case of unequal queue, the distributions of the individual shockwaves corresponding to each lane of the intersection approach are different. Therefore, the simplest way of predicting the probe vehicle's lane is by clustering the individual queuing shockwave speed corresponding to each cycle and by using an arbitrary shockwave speed boundary.

The shockwave speed boundary

Shockwave speed boundary is a threshold value that discriminates the individual shockwave speeds of each lane as a result of difference in vehicle arrival rates in a given cycle. In this study, the shockwave speed boundary is assumed to be the midpoint between the maximum and minimum individual shockwave speeds as shown in Figure 4 and is computed using Eq. (2):

$$B_c = \min(w_c) + \left(\frac{(\max(w_c) - \min(w_c))}{2} \right) \tag{2}$$

where
B_c = shockwave speed boundary at cth cycle;
w_c = set of individual queuing shockwave speed of cth cycle.

Clustering

After obtaining the boundary shockwave speed, the probe vehicles' shockwave speeds were clustered relative to the boundary. The group of vehicles that lies on either side of the boundary are assumed to be in the same lane. However, this method of clustering can provide mixed results

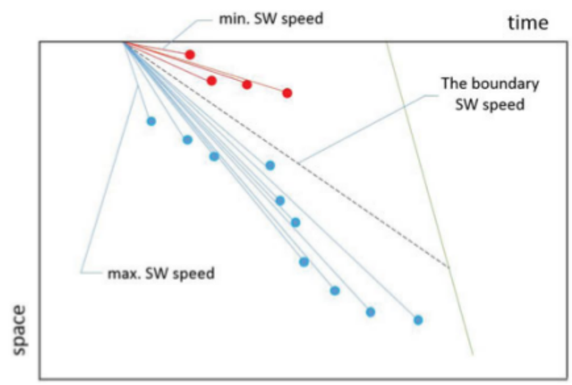

Figure 4. Illustration of individual and the boundary SW speeds.

as there could be some shockwave speed values belonging to multiple lanes. The probe vehicles' lanes were identified according to Eqs. (3) and (4):

$$\text{if } \min(w_c) \le w_i \le B_c, \quad \delta_i = 1, \ \forall\, i \in V \tag{3}$$

$$\text{if } B_c \le w_i \le \max(w_c), \quad \delta_i = 2, \ \forall\, i \in V \tag{4}$$

where
δ_i = decision variable indicating the lane: $\delta_i = 1$ if vehicle is on SQL, $\delta_i = 2$ if the probe vehicle on the LQL;
i = index to a probe vehicle;
V = the set of probe vehicles in a cycle.

Missing data cycles

To calculate the shockwave speed boundary, at least two probe vehicles are needed. However, considering the stochastic nature of traffic, there could be cycles that have less than two, or even zero, probe vehicles. The number of these occurrences becomes greater as the probe vehicle market penetration rate decreases. Therefore, discriminating the lane of probe vehicles using boundary shockwave speed is not applicable to every cycle. In this study, cycles with less than two probe vehicles are ignored as lane prediction and queue length estimation would be impossible. However, it should be noted that in queue length estimation this cycle can still use the queue length estimation from the previous cycle. A limitation of this practice is, it is possible that the arrival rates in consecutive cycles can be considerably different from each other. In such cases, using queue length from previous cycle might lead to biased estimation, but since there is no other information available this is the best that can be done in the case of real-time estimation. The available cycle was evaluated for each travel demand scenarios and market penetration.

K-means clustering method

K-means clustering is a type of hard clustering implementation where the clusters are assumed to be independent and do not overlap. In this study, the clustering was carried out for one-dimensional K-means clustering, i.e., only individual shockwave speeds are used as a variable for K-means clustering.

The inputs for this method are:

1. number of clusters (k);
2. set of points, in this case, the individual SW speeds (w_c) for the associated cycle;
3. partition the n observation into sets $p = \{p_1, p_2, \ldots, p_k\}$. In this study, there are only two partitions ($k = 2$) which is SQL (p_1) and LQL (p_2).

The objective function that the algorithm of K-means clustering minimizes is:

$$\arg\min_{p} \sum_{i=1}^{k} \sum_{w_i \in p_i} w_i - \mu_i{}^2 \tag{5}$$

where μ_i = mean of points in p_i.

Supervised learning methods

The unsupervised clustering methods discussed above force at least one probe vehicle per lane. However, this may not necessarily be the case, especially on the lane where the queue is normally short. In order to overcome this shortcoming and accommodate scenarios where all probes occupy only one of the lanes, two supervised learning methods are explored in this section. The first one is a univariate model, which utilizes shockwave speeds (w_i) only and fits lognormal probability density functions to the training data. The second is a bivariate model and utilizes both shockwave speeds (w_i) and distance of the probe vehicle from the stop bar. The probability density functions for the second model are found by fitting a mixture of normal distributions to the training data as explained in Section 7.2. Both models use the optimal Bayes rule to predict the lanes the probe vehicles are at.

Lognormal model

Two probability density functions fitted to the shockwave speeds (w_i) are needed for the Bayesian decision rule: one for the SQL and another for the LQL. To investigate the appropriate distributions, the shockwave speeds (w_i) from each lane were fitted to a variety of probability distributions using a fitting software called CumFreg program (Oosterbaan, 1994). Out of all the distributions, the lognormal distribution was found to be fitting the best to the shockwave speeds. The probability density function (pdf) of the lognormal distribution is given by Eq. (6):

$$f\ (w_i|\delta_i) = \frac{1}{w_i \sigma \sqrt{2\pi}} e^{-\frac{(\ln w_i - \mu)^2}{2\sigma^2}}, \quad w_i > 0 \tag{6}$$

where

μ = mean of the shockwaves on the log scale;

σ = standard deviation of the shockwaves on the log scale.

Once the marginal pdfs for each lane are found, the optimal Bayes rule can be applied to infer the lane of the probe vehicle from their shockwave speeds. The Bayes rule is given by Eq. (7):

$$\text{if } f(w_i|\ \delta_i = 1) > \frac{\pi(\delta = 2)}{\pi(\delta = 1)} f(w_i|\ \delta_i = 2), \quad \delta_i = 1 \text{ else } \delta_i = 2 \tag{7}$$

where

$f(w_i|\ \delta_i = 1)$ = the lognormal pdf estimated based on training data from SQL;

$f(w_i|\ \delta_i = 2)$ = the lognormal pdf estimated based on training data from LQL;

$\pi(\delta = 1)$ = the prior component of group s which is the percentage of s from the population ($\pi(\delta = 1) = \frac{n(\delta=1)}{n(\delta=1)+n(\delta=2)}$), $n(\delta = 1)$: total number of vehicles observed in SQL, $n(\delta = 2)$: total number of vehicles observed in LQL;

$\pi(\delta = 2)$ = the prior component of group s which is the percentage of l from population ($\pi(\delta = 2) = \frac{n(\delta=2)}{n(\delta=1)+n(\delta=2)}$). □

The prior components discussed above are supposed to be obtained from training data. In this study, the training data were generated from the same traffic simulation setting in *Vissim* discussed above.

Bivariate mixture model clustering

In order to improve the accuracy of the prediction, in addition to the shockwave speeds, the distance of the probe to the stop bar is also considered as an input. In that case, a bivariate pdf is needed to characterize the conditional distributions. These pdfs are found by fitting a mixture of normal distributions to the training data, which consists of two variables: w_i and d_i that are denoted by vector $x = [w_i, d_i]^{\mathrm{T}}$. Given the lane, the bivariate density is approximated by a mixture of normals as formulated in Eq. (8):

$$f\ (x|\delta = j) = \sum_{k=1}^{G_j} \tau_k f_k(x|\delta = j) \quad j \in \{1, 2\} \tag{8}$$

where

f_k = pdf of mixture component k;

j = index for the lane, either 1 for SQL or 2 for LQL;

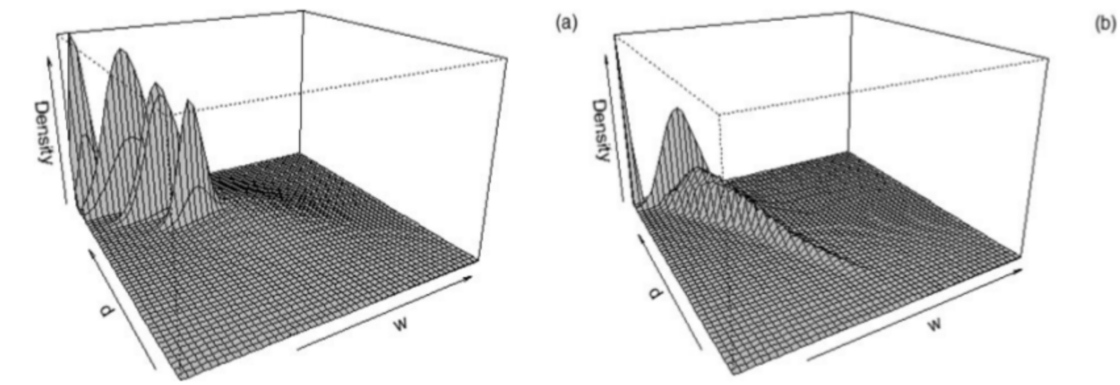

Figure 5. Bivariate mixture model pdf for (a) SQL and (b) LQL.

G_j = number of mixture model components;
τ_k = probability that an observation comes from the k-th mixture component. ($\tau_k \in (0, 1)$ and $\Sigma_k \tau_k = 1$).

Each component f_k is a bivariate normal distribution (Fraley & Raftery, 2007) and has the probability density function as in Eq. (9):

$$\phi\ (x_i; \mu_k, \Sigma_k) = \frac{\exp\left\{-\frac{1}{2}(x_i - \mu_k)^{\mathrm{T}} \Sigma_k^{-1} (x_i - \mu_k)\right\}}{\sqrt{\det\left(2\pi\, \Sigma_k\right)}} \tag{9}$$

where
μ_k = mean for component k;
Σ_k = covariance matrix.

The bivariate mixture model clustering was performed according to the following steps:

1. Using the probe vehicle individual shock-wave speeds and the distance to the traffic light while queuing, implement the mixture model to get the best pdf for each lane ($f(x|\delta = 1)$ & $f(x|\delta = 2)$).
2. Calculate the density for each probe vehicle in the data using the pdf for each lane.
3. Calculate the prior ($\pi(\delta = 1)$ & $\pi(\delta = 2)$) as discussed in the previous section.
4. Estimate the lane according to Eq. (7).
5. Calculate the probability for each lane. The probability of a probe vehicle to be in the SQL and LQL is given by Eqs. (10) and (11), respectively.

$$P(\delta i = 1) = \frac{\pi\ (\delta = 1)\ f\ (x|\delta = 1)}{\pi\ (\delta = 1)\ f\ (x|\delta = 1) + \pi\ (\delta = 2)\ f\ (x|\delta = 2)} \tag{10}$$

$$P(\delta i = 2) = \frac{\pi\ (\delta = 2)\ f\ (x|\delta = 2)}{\pi\ (\delta = 1)\ f\ (x|\delta = 1) + \pi\ (\delta = 2)\ f\ (x|\delta = 2)} \tag{11}$$

The bivariate mixture model, which identifies the best PDFs for each lane, was implemented by utilizing the package *mclust* (Fraley et al., 2015) in *R* statistical software. Once the best PDFs for both SQL and LQL are obtained, the optimal Bayes rule for allocation is implemented for assigning a lane to the probe vehicles. Figure 5 exhibits examples of PDFs for SQL and LQL estimated using the bivariate mixture model clustering method.

Evaluation of the methods and results

In this section, the results of all the methods discussed earlier are presented. The precisions of each method are calculated to assess the level of accuracy obtained from each model. The precision is computed by the following steps:

1. Categorize the predictions based on the true lane as shown in Table 1.
2. *Calculate the precisions*: The SQL identification precisions are calculated using the formula in Eq. (12), while the LQL identification precisions are calculated using Eq. (13).

$$Pc_s = \frac{T_1}{T_1 + F_1} \tag{12}$$

$$Pc_l = \frac{T_2}{T_2 + F_2} \tag{13}$$

One of the objectives of identifying the probe vehicles' lane is accurate estimation of queue length. In queue length estimation, special attention should be paid to the

Table 1. Prediction categories.

True lane	Predicted lane	Category
SQL	SQL	True 1 (T1)
LQL	LQL	True 2 (T2)
SQL	LQL	False 2 (F2)
LQL	SQL	False 1 (F1)

last probe vehicle. Therefore, in this study, the precision is calculated for every last probe vehicle in their predicted lane. Although the models in this study were developed for two lanes signalized intersection, they are applicable to roadways with more than two lanes as long as there is a significant difference in the shockwave speeds (and distances) among the roadway lanes.

Comparison of the lane identification methods

Figure 6 shows the prediction precision of the different methods of probe vehicle's lane identification for Scenarios 1, 2, and 3 corresponding to SQL and LQL under different market penetration rates. The overall prediction precision for SQL is observed to be less than that of LQL.

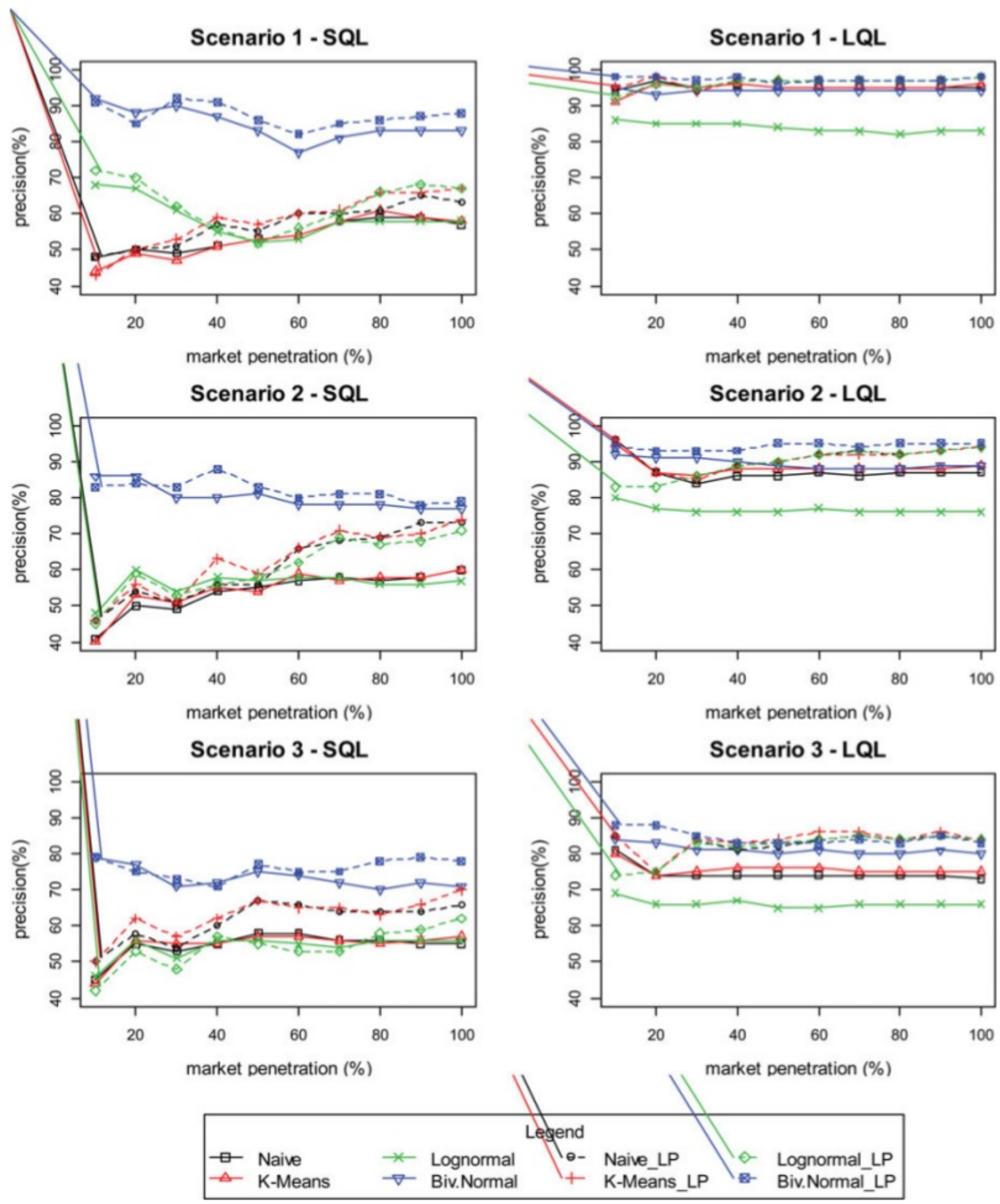

Figure 6. Lane identification prediction precision corresponding SQL and LQL using different methods under different scenarios and probe vehicle penetration rates. (The broken lines represent prediction accuray for the last probe vehicle in SQL and LQL.)

This is expected as there are more probe vehicles in the LQL relative to the SQL, which makes the predictions corresponding to LQL more accurate. For both SQL and LQL, Scenario 1 has the highest prediction precision and is followed by Scenarios 2 and 3. The reason behind this declining trend of prediction precision is that as the difference in the flow rates between the adjacent lanes decreases, predicting the lane of the probe vehicles becomes more difficult as the individual shockwave speed of the probe vehicles in the adjunct lanes tend to overlap. This can be explained by the fact that the queue lengths grow almost equally as the arrival rate of SQL gets closer to that of LQL. This is due to the fact that, as the difference in arrival rates between the LQL and SQL reduces, proper identification of the lane of the last probe vehicle becomes more difficult.

Comparing the performance of the different methods of probe vehicles' lane identification is straightforward. It is observed that the performance of bivariate mixture model clustering is superior to other methods. The prediction accuracies for probe vehicles' lane identification of the naïve, K-means clustering and Lognormal mixture models were found to be significantly lower than those of the bivariate mixture model. The Lognormal mixture model had similar prediction precision for SQL to those of naïve and K-means clustering methods. However, the performance of naïve and K-means clustering methods is slightly better than Lognormal mixture model for LQL. The prediction precision of SQL heavily depends on the market penetration rate. With increase in penetration rates, the prediction accuracies of both naïve, K-means clustering and Lognormal mixture models for SQL were observed to steadily increase. This is true as the probe vehicles' market penetration rates increase, the number of probe vehicle samples in SQL is more likely to increase, which in turn improves the prediction precision. On the other hand, the prediction precision of LQL seems to be slightly affected by increase in market penetration rate. This is best explained by the fact that, since the number of probe vehicle samples in LQL is already high, an increase in market penetration rate does not necessarily affect the prediction precision.

The application of the probe vehicles' lane identification is the estimation of lane-specific queue lengths at signalized intersections. For this purpose, lane identification of the last probe is very crucial. In Figure 6, the results of lane prediction precision for the last probes in SQL and LQL corresponding to the different methods are shown in broken lines. The prediction precision of the last probe was found to be consistently higher in all lane identification methods. This is because of the fact that the shockwave speeds corresponding to the last probes are more distinct from the ones close to the stop bar. Therefore, the lane-based queue length estimation using the lane information regarding the last probes is expected to provide better results.

Given that the naïve, K-means clustering and Lognormal mixture models use only one variable, which is individual shockwave speeds of the vehicles, as the input to identify the lanes of the probes, their performance is not expected to be accurate enough. Moreover, the naïve and K-means clustering methods suffer from the drawback that the algorithms follow hard clustering rules, which do not allow clusters to overlap, i.e., the algorithms always force probe vehicles to be assigned to both lanes even if no probe vehicle is observed in either lane, which certainly compromises the prediction accuracies of SQL. The reason that the bivariate mixture model outperformed the naïve, K-means clustering and Lognormal mixture models is that it uses two parameters, individual shockwave speeds and the distance of the probes from the stop bar, as an input to identify the lanes of the probe vehicles. Considering the prediction precision of bivariate mixture model clustering method was superior to the other methods, results from this method were further explored as discussed in the following Section.

In summary, the following can be concluded in regard to prediction precision of probe vehicles' lane identification: (1) The prediction precision of LQL is higher than that of SQL and thus SQL requires particular attention, (2) the lane identification in LQL results in satisfactory precision level, and (3) the prediction precision gets higher when only the last probe is accounted for.

Figure 6 shows that the prediction accuracies of lane position in general are not so sensitive to the penetration rate of probe vehicles. The reason behind this fact is mainly because the probe lane identifications were accomplished using the individual probe vehicles' shockwave speeds, instead of the aggregated ones, as the main variable. Because of that, the lane identification precision was not affected by the market penetration rates; rather it was affected by the number of shockwave speeds overlapping from the two lanes. The more cases of overlapping shockwave speeds appeared, the less precise the lane identification was. Thus, for the bivariate mixture model, the lane identifications were carried out for all the probe vehicles that have lesser chance of overlapping.

Figures 7 and 8 show the results for the last probe vehicles' lane prediction precision and available cycles for SQL and LQL using Bivariate Mixture Model Clustering Method in more detail. In the figures, the horizontal axis, i.e., probability of a probe vehicle being on SQL or LQL, was calculated using the prediction probability shown in Eqs. (10) and (11). In general, both figures show that for

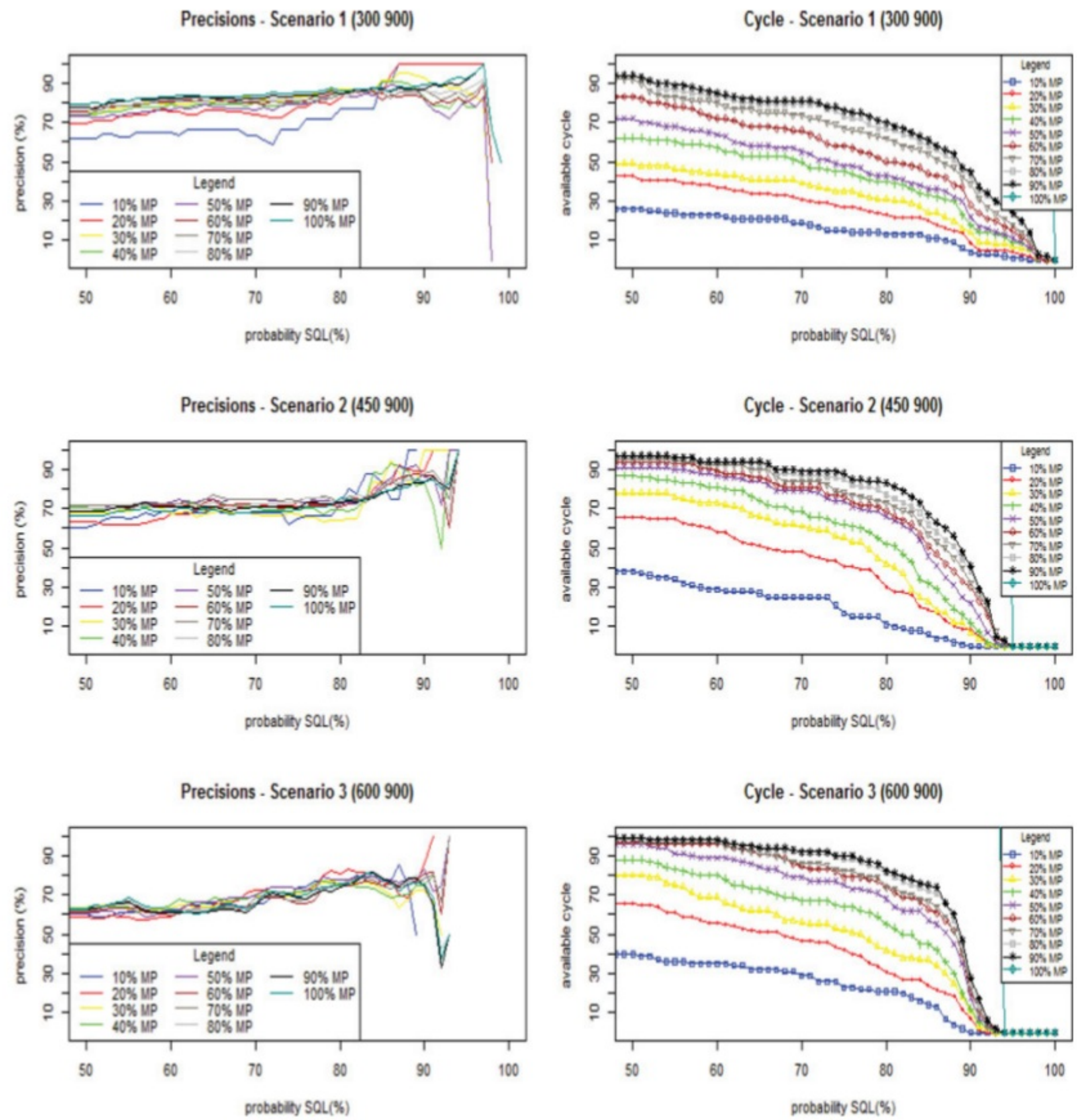

Figure 7. The last probe vehicles' lane prediction precision and available cycle for SQL using bivariate mixture model clustering method.

all scenarios, higher probability results in better precision but the number of available cycle decreases.

In Figure 7, it can be seen that there is a trade-off between precision and available cycles. The fluctuations in the precision, when the probability of SQL is very high, were mostly caused by the significant decrease in the number of available cycles. For example in Scenario 3, for the probability above 90%, the available cycle for all market penetration was less than ten.

Predictability

The probe vehicles' lane identification heavily depends on the individual shockwave speed difference between the two lanes and thus when the differences get smaller the lane prediction precision declines as in Scenario 3. Figure 9 shows examples of the low prediction probability cycles. The y-axis is the distance in meters from the stop bar, and the x-axis is the time in seconds the probe vehicles joined the back of the queue measured from the time the green traffic light ended. The red and black circles represent the stop points of the probe vehicles in the SQL and LQL, respectively, and are mixed together or overlapping. This makes it very difficult to differentiate the probe vehicles' lane position even by taking only the last probe vehicles. The individual shockwaves in these cycles normally have low prediction probability, which are usually less than 50%. In other

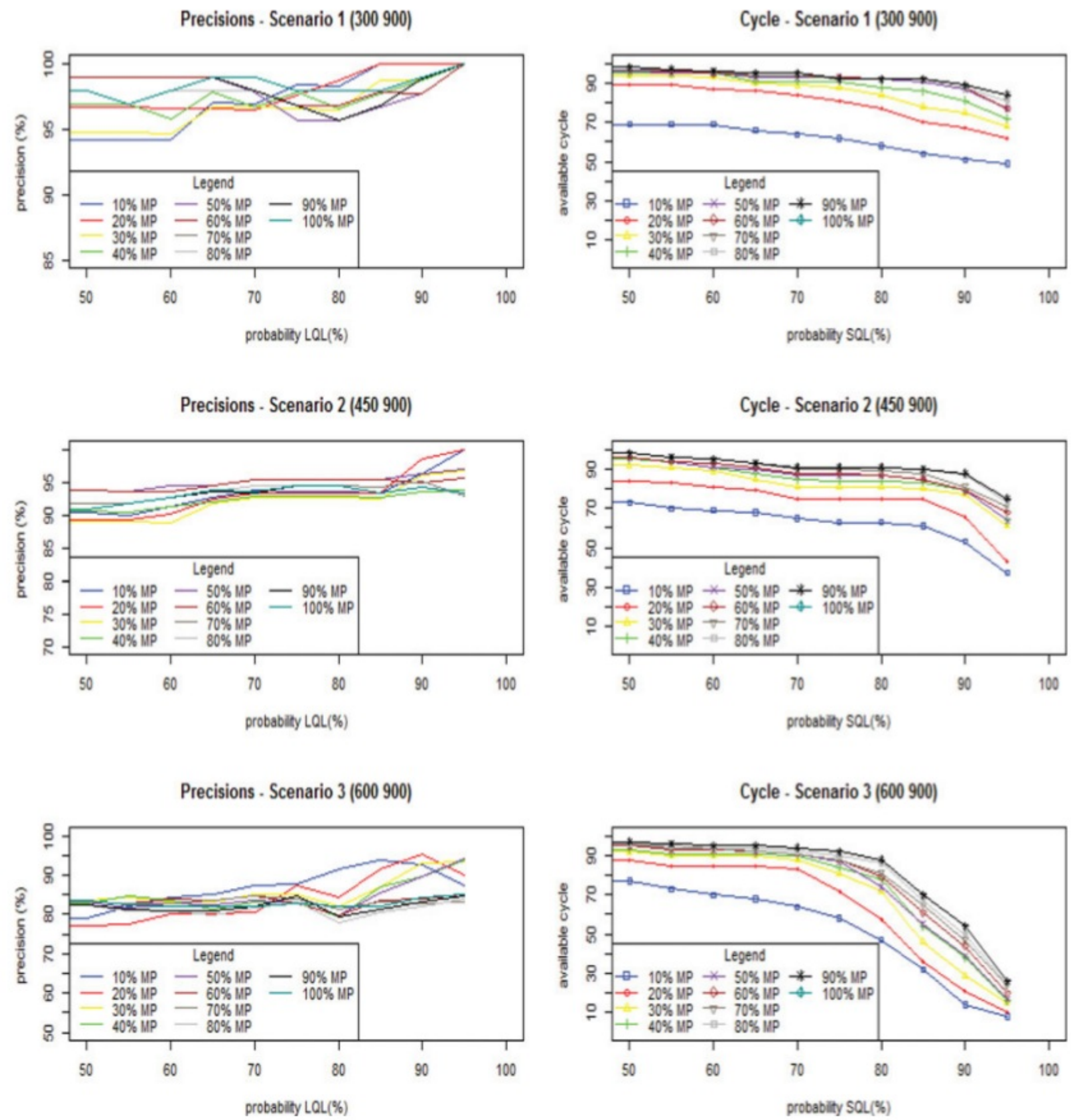

Figure 8. The last probe vehicles' lane prediction precision and available cycle for LQL using bivariate mixture model clustering method.

words, using the prediction probability, at a certain threshold which is apparently more than 50%, the identification can be carried out with high confidence.

Probability threshold

Considering there are three scenarios and ten different market penetration rates, determining the probability threshold is a bit problematic. Apparently, the desired probability threshold is the probability where the precisions are acceptable but also provide enough number of available cycles for all market penetration rates. To determine the probability threshold, the mean precision and also the average number of available cycles for all market penetration rates were calculated as shown in Table 2.

Table 2 shows that, for SQL at probability threshold of 85%, the precision for Scenarios 2 and 3 is the best and still acceptable for Scenario 1 (>85%). Except for Scenario 2, the available cycles for this threshold are greater than 40. For LQL, at the probability threshold of 75%, the combination of the prediction precision and the available number of cycles is pretty good (in all Scenarios except 3, the precisions are greater than 85%). Using these considerations, the probability thresholds for SQL and LQL were set to be 85% and 75%, respectively.

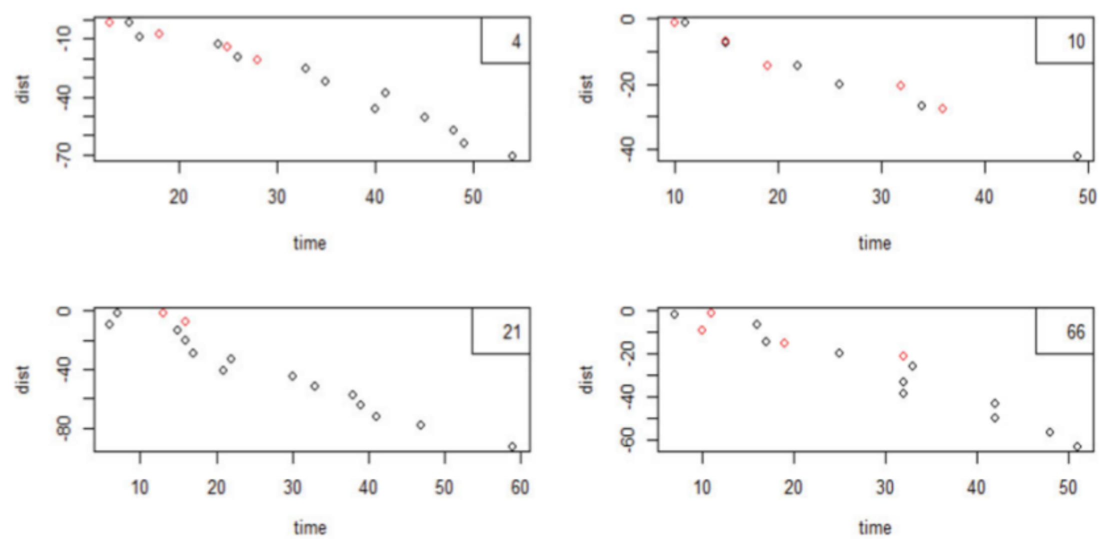

Figure 9. Example of low prediction probability cycles.

By applying the probability thresholds, the prediction accuracies and available number of cycles for both lanes in the different scenarios and market penetration rates considered are provided in Table 3.

On average, the prediction precision of lane identification for both lanes are good (≥ 85%), except for Scenario 3 where the individual shockwaves overlap making them undistinguishable. Unexpectedly, the available number of

Table 2. The mean of the precision and available cycles of all market penetration rates in each scenario.

	SQL						LQL					
	Precision (%)			No. of cycles			Precision (%)			No. of cycles		
Probability threshold (%)	Sc. 1	Sc. 2	Sc. 3	Sc. 1	Sc. 2	Sc. 3	Sc. 1	Sc. 2	Sc. 3	Sc. 1	Sc. 2	Sc. 3
50	74	68	62	71	84	86	97	91	82	93	93	93
55	76	69	64	67	83	83	97	91	82	93	91	91
60	78	70	62	63	79	82	97	92	82	92	90	90
65	79	72	64	59	76	77	98	93	82	90	88	89
70	79	71	68	58	71	73	97	94	83	90	85	88
75	80	71	70	52	67	69	97	94	84	88	84	83
80	84	72	76	47	59	61	97	94	82	86	84	75
85	87	82	76	41	43	53	98	94	86	84	83	56
90	89	87	78	27	23	17	99	95	88	81	78	39
95	88	NA	NA	14	0	0	100	96	88	74	63	19
100	NA	NA	NA	0	0	0	NA	NA	NA	NA	NA	NA

Table 3. The prediction precision and number of available cycles (threshold probabilities of 85% for SQL and 75% for LQL).

	Precision (%)						Available cycles					
	SQL			LQL			SQL			LQL		
Market penetration rate (%)	Sc. 1	Sc. 2	Sc. 3	Sc. 1	Sc. 2	Sc. 3	Sc. 1	Sc. 2	Sc. 3	Sc. 1	Sc. 2	Sc. 3
10	91	83	79	98	94	88	11	6	14	62	63	58
20	85	84	75	98	93	88	20	19	24	81	75	72
30	92	83	73	97	93	85	24	23	37	88	81	81
40	91	88	71	98	93	83	33	32	45	91	84	84
50	86	83	77	96	95	83	36	46	57	93	87	87
60	82	80	75	97	95	83	44	51	63	93	88	88
70	85	81	75	97	94	84	54	57	67	92	90	88
80	86	81	78	97	95	83	58	62	72	92	91	90
90	87	78	79	97	95	85	61	67	75	92	91	92
100	88	79	78	98	95	83	64	71	76	94	92	94
Average	87	82	76	97	94	85	41	43	53	88	84	83

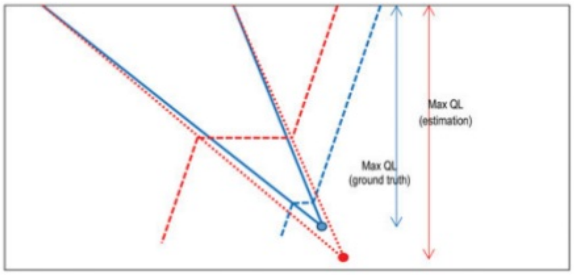

Figure 10. Maximum queue length estimation using LWR shockwave theory.

cycles in Scenario 2, for both SQL and LQL, is the lowest among the three scenarios. This means that, in this scenario, there are fewer cycles that have at least one vehicle with prediction probability within the threshold. In other words, there are fewer cycles that have a probe vehicle whose lane can be identified with high confidence. The result also shows that, except for available cycles, there is no clear relationship between the penetration rate and the lane identification precision.

Implementation of lane identification on queue length estimation

The LWR shockwave theory, which is a commonly used model queue length estimation at signalized intersections, is employed to estimate the queue length from the time and space coordinates of the probe vehicles at the moment they join queue. The reason of using this theory is because the probe vehicles provide their positions (GPS coordinates) and LWR method use both space/distance and time coordinates. However, accurate estimation of queue length requires that the lane of the probe vehicles be determined ahead of time. For this purpose, the Bivariate Mixture Model Clustering model discussed in Section 7.2 was used.

The limitation of LWR shockwave theory is that the queue lengths are estimated indirectly by the intersection between queuing shockwave and discharging shockwave as shown in Figure 10. This means that the theory can only predict the maximum queue that can take place instead of the real queue length. Considering this limitation, the errors in estimating the queue lengths are computed by

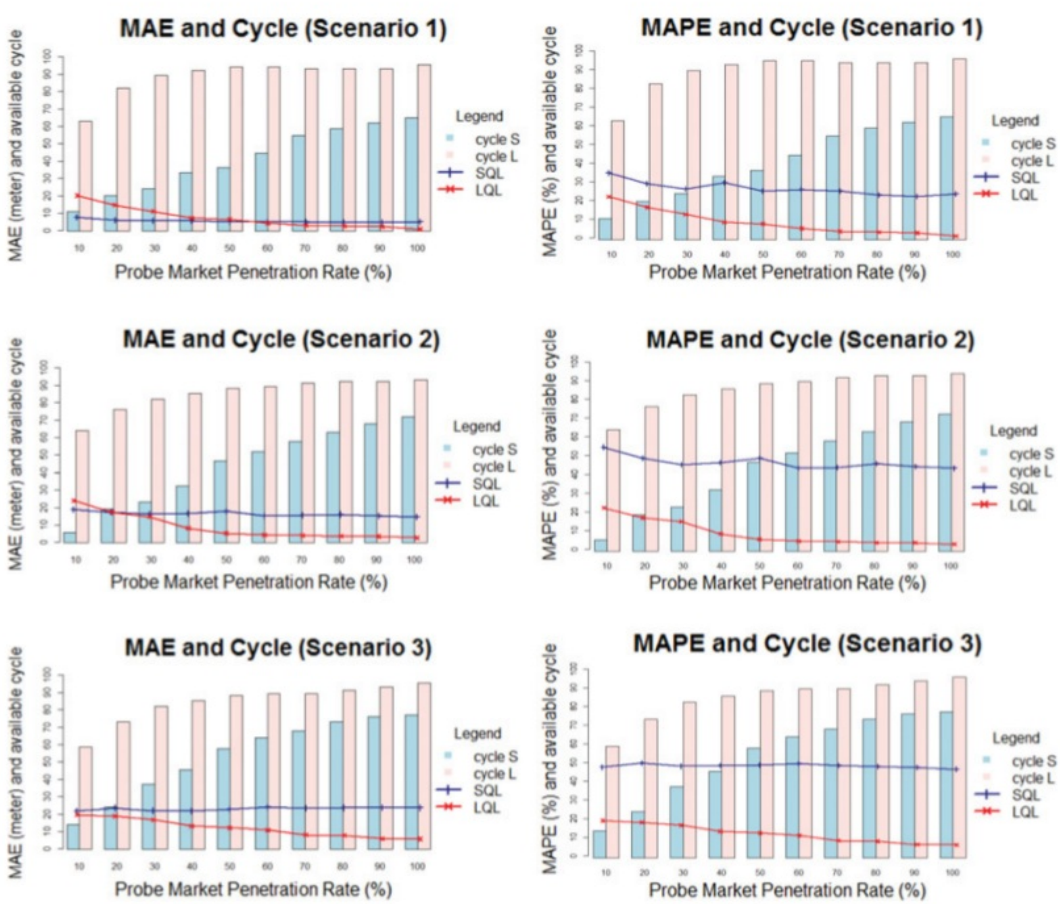

Figure 11. Queue length estimation error and available cycles.

comparing the maximum queue length calculated from the true last probe vehicle and the queue length estimated from the LWR model as illustrated in Figure 10.

Mean absolute error (MAE) is utilized to measure the bias between the estimation and the ground truth. The MAE is calculated as in Eq. (14):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{14}$$

where
n = number of samples;
e_i = error of sample i.

The mean absolute percentage error (MAPE) was utilized to understand the proportion of error. The MAPE is calculated as in Eq. (15):

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{15}$$

where
A_t = actual queue length value;
F_t = forecast queue length value.

The lane-based AE and MAPE values and also the number of available cycles of each market penetration for each scenario are presented in Figure 11. The pink and light blue bars represent the number of available cycles for SQL and LQL, while the blue and red lines represent the MAE in meters and MAPE in percentages. As expected, the figure shows that in all cycles for all market penetration rates, the number of cycles for SQL is always less than the number of cycles for LQL. The figure also exhibits that the MAE for LQL is less than that of SQL for all scenarios where the MAPEs for LQL fall below 10% for Scenarios 1 and 2 at 40% market penetration, and for Scenario 3 at 70% market penetration. The MAPEs for SQL are generally higher than LQL as the queue length is shorter, and a small deviation results in bigger percentage difference. In that case, MAE could be a better indicator of the errors on queue length estimation. Therefore, although the MAPEs of SQL are a bit large, they can actually be tolerable. On average, the MAE for Scenarios 1, 2, and 3 was found to be 5, 16.4, and 23.1 m, respectively.

One reason that the SQL results are not very good is that in the lane identification process, the vehicles that are

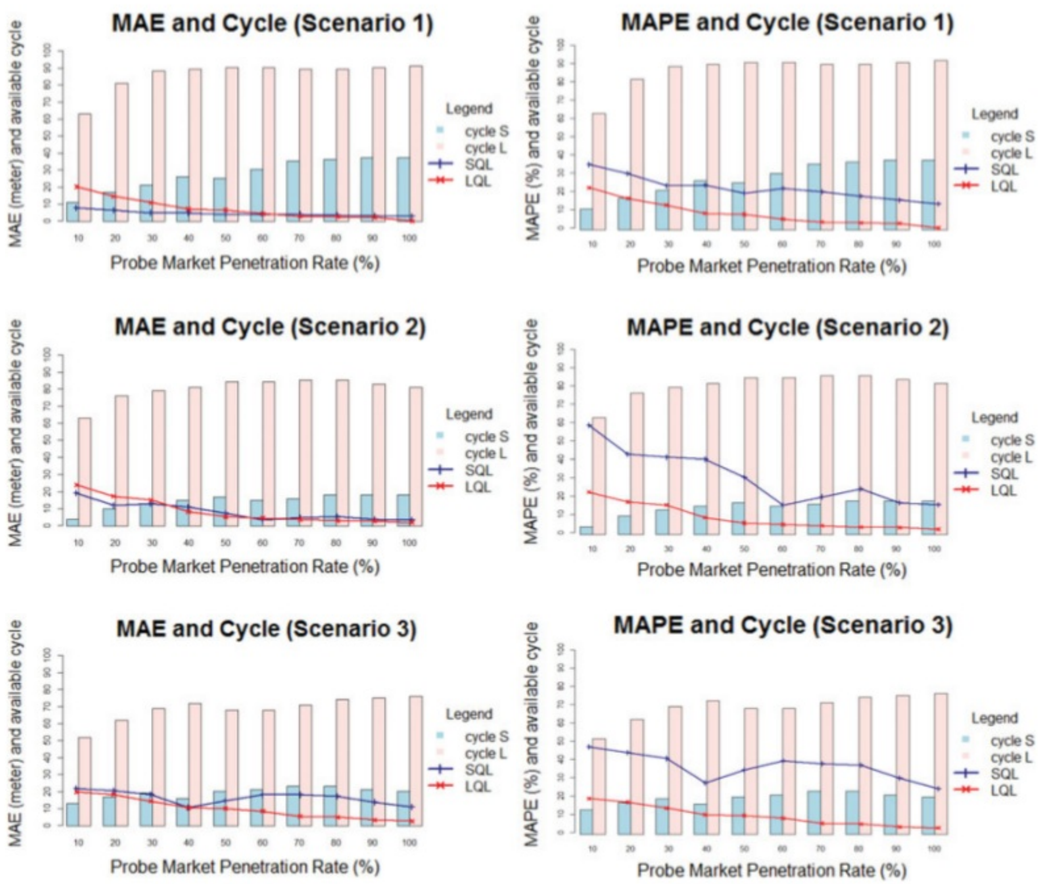

Figure 12. Queue length estimation error and available cycles by determining the last probe first.

taken to be identified properly are the vehicles that comply with a probability threshold of 85% for SQL and 75% for LQL, while all others that do not meet these thresholds were filtered out. Because of this process, the last vehicle identified in a lane might not really be the last vehicle estimated in the cycle, which produces more bias.

To deal with the aforementioned issue, the order of the process is reversed, i.e., the last probe vehicles are determined first and then the probability threshold is applied. This is done to decrease the MAE; however, it is worth noticing that there is a trade-off between the improvement of accuracy and the cycle availability as shown in Figure 12. The figure confirms the improvement in estimation accuracies, where the average biases for SQL queue length estimations are 4.5, 8.3, and 16.4 m for Scenarios 1, 2, and 3, respectively. Meanwhile, for LQL, the MAPE value is equal to or below 10% starting from 40% market penetration for all scenarios. However, the number of cycles decreases significantly for SQL for almost all market penetration rates in all scenarios.

Another important finding is that, using all approaches, the queue length estimation accuracy of LQL is always better relative to SQL. This is because the queue length estimation precision highly depends on the accuracy of the lane identification. As has been demonstrated by the result in the lane identification section, the accuracies of probe vehicles' lane identification for LQL are better than SQL for all scenarios. The results show that the more SQL demand, the more inaccurate the queue length prediction. Again, this tends to have a linear relationship with the lane identification, where the more the demand the less accurate the lane identification. The explanation of this is, the higher the demand for SQL, the more chance that the shockwave speeds corresponding to the overlap of both lanes, which decreases accuracy of lane identification.

Summary and conclusion

In this study, a methodology to identify a probe vehicles' lane in a two-lane signalized intersection was developed. The individual probe vehicles' shockwave speed and their distance from the stop bar were used in a bivariate mixture model for clustering their lanes using the Optimal Bayes Rule for allocation which utilizes probability density functions. The result demonstrates that using bivariate mixture model, the lane identification can be carried out with acceptable results when the threshold probability is greater than or equal to 85% for SQL and greater than or equal to 75% for LQL.

This study also demonstrated the use of probe vehicles for estimating queue length at an intersection. Ten probe market penetration scenarios in three demand scenarios have been tested to generate data using the traffic simulation software. The LWR shockwave theory has been utilized successfully to estimate the queue length following the probe vehicles' lane identification. The result shows that in most cases, a 40% market penetration rate is enough to reach about 90% accuracy. The results also confirm that lane identification plays an important role in estimating queue lengths of unequal demand on adjacent lanes of an intersection approach, which can be predicted well if the lane identification is conducted properly.

In general, this research has demonstrated that the probe vehicles' lane identification plays an important role in estimating queue length at signalized intersections; however, several areas are recommended to be improved, which are the limitations of the current research, i.e. 1) future research should consider the issue of heterogeneity in vehicle's type and 2) testing these methods with field data.

Acknowledgments

The authors would like to thank TranLIVE University Transportation Center for their support to this study.

References

Badillo, B. E., Rakha, H., Rioux, T. W., & Abrams, M. (2012). *Queue length estimation using conventional vehicle detector and probe vehicle data*. 15th International IEEE Conference on Intelligent Transportation Systems (ITSC) (pp. 1674–1681). IEEE. Retrieved from http://ieeexplore.ieee.org/abstract/document/6338891/

Ban, X. J., Hao, P., & 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.

Cetin, M. (2012). Estimating queue dynamics at signalized intersections from probe vehicle data: Methodology based on kinematic wave model. *Transportation Research Record: Journal of the Transportation Research Board*, (2315), 164–172.

Cetin, M., & Rakha, H. (2014). *Estimating fuel consumption and carbon footprint at signalized intersections using probe vehicle trajectories*. Transportation Research Board 93rd Annual Meeting (Report Nos. 14-2285). Retrieved from https://trid.trb.org/view.aspx?id=1288243.

Chang, G. L., & Su, C. C. (1995). Predicting intersection queue with neural network models. *Transportation Research Part C: Emerging Technologies*, *3*(3), 175–191.

Cheng, Y., Qin, X., Jin, J., & Ran, B. (2012). An exploratory shockwave approach to estimating queue length using probe trajectories. *Journal of Intelligent Transportation Systems*, *16*(1), 12–23.

Cheng, Y., Qin, X., Jin, J., Ran, B., & Anderson, J. (2011). Cycle-by-cycle queue length estimation for signalized intersections using sampled trajectory data. *Transportation Research Record: Journal of the Transportation Research Board*, (2257), 87–94.

Comert, G. (2016). Queue length estimation from probe vehicles at isolated intersections: Estimators for primary parameters. *European Journal of Operational Research*, *252*(2), 502–521.

Comert, G., & Cetin, M. (2009). Queue length estimation from probe vehicle location and the impacts of sample size. *European Journal of Operational Research*, *197*(1), 196–202.

Comert, G., & Cetin, M. (2011). Analytical evaluation of the error in queue length estimation at traffic signals from probe vehicle data. *IEEE Transactions on Intelligent Transportation Systems*, *12*(2), 563–573.

Fraley, C., & Raftery, A. E. (2007). Model-based methods of classification: using the mclust software in chemometrics. *Journal of Statistical Software*, *18*(6), 1–13.

Fraley, C., Raftery, A. E., Scrucca, L., Murphy, T. B., Fop, M., & Scrucca, M. L. (2015). *Package "mclust"*. Retrieved from ftp://193.1.193.66/pub/cran.r-project.org/web/packages/mclust/mclust.pdf

Geroliminis, N., & Skabardonis, A. (2005). Prediction of arrival profiles and queue lengths along signalized arterials by using a Markov decision process. *Transportation Research Record: Journal of the Transportation Research Board*, (1934), 116–124.

Hao, P., Ban, X., & Whon Yu, J. (2015). Kinematic equation-based vehicle queue location estimation method for signalized intersections using mobile sensor data. *Journal of Intelligent Transportation Systems*, *19*(3), 256–272.

Hunter, M. P., Wu, S. K., Kim, H. K., & Suh, W. (2012). A probe-vehicle-based evaluation of adaptive traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, *13*(2), 704–713.

Lighthill, M. J., & Whitham, G. B. (1955). *On kinematic waves. II. A theory of traffic flow on long crowded roads*. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences (Vol. 229, pp. 317–345). The Royal Society. Retrieved from http://rspa.royalsocietypublishing.org/content/royprsa/229/1178/317.full.pdf

Liu, H. X., Wu, X., Ma, W., & Hu, H. (2009). Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies*, *17*(4), 412–427.

Oosterbaan, R. J. (1994). Frequency and regression analysis. Drainage principles and applications, Chapter 6. *International Institute for Land Reclamation and Improvement, Wageningen*. Retrieved from http://www.waterlog.info/pdf/regtxt.pdf.

Ramezani, M., & Geroliminis, N. (2013). *Exploiting probe data to estimate the queue profile in urban networks*. 16th International IEEE Conference on Intelligent Transportation Systems (ITSC) (pp. 1817–1822). IEEE. Retrieved from http://ieeexplore.ieee.org/abstract/document/6728492/

Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, *4*(1), 42–51.

Turner, S. M., Eisele, W. L., Benz, R. J., & Holdener, D. J. (1998). *Travel time data collection handbook*. Retrieved from https://trid.trb.org/view.aspx?id=497690

Vigos, G., Papageorgiou, M., & Wang, Y. (2008). Real-time estimation of vehicle-count within signalized links. *Transportation Research Part C: Emerging Technologies*, *16*(1), 18–35.

Vissim, P. T. V. (n.d.). *6.0 User Manual, 2014 PTV AG*. Karlsruhe, Germany.

Viti, F., & van Zuylen, H. (2004). Modeling queues at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, (1883), 68–77.

Viti, F., & Van Zuylen, H. J. (2009). The dynamics and the uncertainty of queues at fixed and actuated controls: A probabilistic approach. *Journal of Intelligent Transportation Systems*, *13*(1), 39–51.

Wang, Z., Cai, Q., Wu, B., Zheng, L., & Wang, Y. (2017). Shockwave-based queue estimation approach for undersaturated and oversaturated signalized intersections using multi-source detection data. *Journal of Intelligent Transportation Systems, 21*(3), 167–178; http://dx.doi.org/10.1080/15472450.2016.1254046

Wu, J., Jin, X., & Horowitz, A. (2008). Methodologies for estimating vehicle queue length at metered on-ramps. *Transportation Research Record: Journal of the Transportation Research Board*, (2047), 75–82.

Probe vehicle lane identification for queue length estimation at intersections

ORIGINALITY REPORT

13%	8%	10%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	trid.trb.org Internet Source	3%
2	Chengchuan An, Yao-Jan Wu, Jingxin Xia, Wei Huang. "Real-time queue length estimation using event-based advance detector data", Journal of Intelligent Transportation Systems, 2017 Publication	1%
3	mafiadoc.com Internet Source	1%
4	www.ece.eng.wayne.edu Internet Source	1%
5	Guangchuan Yang, Zong Tian, Daobin Wang, Hao Xu. "Queue length estimation for a metered on-ramp using mesoscopic simulation", Transportation Letters, 2018 Publication	<1%
6	Anna Mary Philip, Gitakrishnan Ramadurai. "Analytical Model for Queue Length Estimation	<1%

at Signalized Intersections from Travel Time", 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019
Publication

7	digitalcommons.odu.edu Internet Source	<1%
8	Jiarong Yao, Keshuang Tang. "Cycle-based queue length estimation considering spillover conditions based on low-resolution point detector data", Transportation Research Part C: Emerging Technologies, 2019 Publication	<1%
9	export.arxiv.org Internet Source	<1%
10	theses.ucalgary.ca Internet Source	<1%
11	orbilu.uni.lu Internet Source	<1%
12	Chaopeng Tan, Lei Liu, Hao Wu, Yumin Cao, Keshuang Tang. "Fuzing license plate recognition data and vehicle trajectory data for lane-based queue length estimation at signalized intersections", Journal of Intelligent Transportation Systems, 2020 Publication	<1%
13	Xianyuan Zhan, Ruimin Li, Satish V. Ukkusuri.	

"Link-based traffic state estimation and prediction for arterial networks using license-plate recognition data", Transportation Research Part C: Emerging Technologies, 2020
Publication | <1%

14 Zsuzsanna Surányi, David B. Hitchcock, James B. Hittner, András Vargha, Róbert Urbán. "Different types of sensation seeking", International Journal of Behavioral Development, 2013
Publication | <1%

15 Eleni I. Vlahogianni, Nikolas Geroliminis, Alexander Skabardonis. "Empirical and Analytical Investigation of Traffic Flow Regimes and Transitions in Signalized Arterials", Journal of Transportation Engineering, 2008
Publication | <1%

16 tandfonline.com
Internet Source | <1%

17 Hua Wang, Changlong Gu, Washington Yotto Ochieng. "Vehicle Trajectory Reconstruction for Signalized Intersections with Low-Frequency Floating Car Data", Journal of Advanced Transportation, 2019
Publication | <1%

18 Henry X. Liu, Xinkai Wu, Wenteng Ma, Heng Hu. "Real-time queue length estimation for | <1%

congested signalized intersections", Transportation Research Part C: Emerging Technologies, 2009
Publication

---

19 "Intelligent Computing, Networking, and Informatics", Springer Science and Business Media LLC, 2014 <1%
Publication

---

20 Semuel Y.R. Rompis, Filmon G. Habtemichael, Mecit Cetin. "A methodology for calibrating microscopic simulation for modeling traffic flow under incidents", 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014 <1%
Publication

---

21 Daniel K. Kimura, David A. Somerton. "Review of Statistical Aspects of Survey Sampling for Marine Fisheries", Reviews in Fisheries Science, 2007 <1%
Publication

---

22 Gurcan Comert, Mecit Cetin. "Queue length estimation from probe vehicle location and the impacts of sample size", European Journal of Operational Research, 2009 <1%
Publication

---

23 tft.eng.usf.edu <1%
Internet Source

24	www.chkwon.net Internet Source	<1%
25	rosap.ntl.bts.gov Internet Source	<1%
26	ir.nctu.edu.tw Internet Source	<1%
27	Submitted to Queen's University of Belfast Student Paper	<1%
28	David Kennedy. "Locality and professional football club development: the demographics of football club support in late victorian Liverpool", Soccer & Society, 2007 Publication	<1%
29	Brian E. Badillo, Hesham Rakha, Thomas W. Rioux, Marc Abrams. "Queue length estimation using conventional vehicle detector and probe vehicle data", 2012 15th International IEEE Conference on Intelligent Transportation Systems, 2012 Publication	<1%
30	Jing Cao, Dawei Hu, Md., Xu Wang, Tony Z. Qiu. "Comparison of queue estimation accuracy by shockwave-based and input-output-based models", 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014 Publication	<1%

31	Seunghyeon Lee, Kun Xie, Dong Ngoduy, Mehdi Keyvan-Ekbatani. "An advanced deep learning approach to real-time estimation of lane-based queue lengths at a signalized junction", Transportation Research Part C: Emerging Technologies, 2019 Publication	<1%
32	www.tri-odu.org Internet Source	<1%
33	Craig A. Jordan, Mecit Cetin, R. Michael Robinson. "Path Clearance for Emergency Vehicles through the Use of Vehicle-to-Vehicle Communication", Transportation Research Record: Journal of the Transportation Research Board, 2013 Publication	<1%
34	link.springer.com Internet Source	<1%
35	He, Q.. "PAMSCOD: Platoon-based arterial multi-modal signal control with online data", Transportation Research Part C, 201202 Publication	<1%
36	pdxscholar.library.pdx.edu Internet Source	<1%
37	Bumjoon Bae, Yuandong Liu, Lee D. Han, Hamparsum Bozdogan. "Spatio-temporal traffic	<1%

queue detection for uninterrupted flows", Transportation Research Part B: Methodological, 2019
Publication

38	Mecit Cetin, Khairul A. Anuar. "Using probe vehicle trajectories in stop-and-go waves for inferring unobserved vehicles", 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2017 Publication	<1%
39	"Intelligent Information Processing X", Springer Science and Business Media LLC, 2020 Publication	<1%
40	Hao, Peng, Zhanbo Sun, Xuegang (Jeff) Ban, Dong Guo, and Qiang Ji. "Vehicle Index Estimation for Signalized Intersections Using Sample Travel Times", Procedia - Social and Behavioral Sciences, 2013. Publication	<1%
41	Yang Cheng, Xiao Qin, Jing Jin, Bin Ran, Jason Anderson. "Cycle-by-Cycle Queue Length Estimation for Signalized Intersections Using Sampled Trajectory Data", Transportation Research Record: Journal of the Transportation Research Board, 2018 Publication	<1%

42	escholarship.org Internet Source	<1 %
43	Majid Rostami Shahrbabaki, Ali Akbar Safavi, Markos Papageorgiou, Ioannis Papamichail. "A data fusion approach for real-time traffic state estimation in urban signalized links", Transportation Research Part C: Emerging Technologies, 2018 Publication	<1 %
44	Bing Li, Wei Cheng, Lishan Li. "Lane-Based Queue Length Estimation in Heterogeneous Traffic Flow Consisting of Cars and Buses", Journal of Transportation Engineering, Part A: Systems, 2020 Publication	<1 %
45	dot.state.mn.us Internet Source	<1 %
46	arxiv.org Internet Source	<1 %
47	Liu, H.X.. "A virtual vehicle probe model for time-dependent travel time estimation on signalized arterials", Transportation Research Part C, 200902 Publication	<1 %
48	journals.sagepub.com Internet Source	<1 %

49	Qing Cai, Zhongyu Wang, Xi Guo, Bing Wu. "New Calculating Method for HCM 2000 Queue Length Estimation Procedures with the Application of Floating Car Data", Procedia - Social and Behavioral Sciences, 2013 Publication	<1%
50	Ying Lee, Chien-Hung Wei, Kai-Chon Chao. "Evaluating the Effects of Highway Traffic Accidents in the Development of a Vehicle Accident Queue Length Estimation Model", International Journal of Intelligent Transportation Systems Research, 2016 Publication	<1%
51	Piccoli, Benedetto, Ke Han, Terry L. Friesz, Tao Yao, and Junqing Tang. "Second-order models and traffic data from mobile sensors", Transportation Research Part C Emerging Technologies, 2015. Publication	<1%
52	academic.hep.com.cn Internet Source	<1%
53	Nikolaos Geroliminis, Alexander Skabardonis. "Prediction of Arrival Profiles and Queue Lengths along Signalized Arterials by using a Markov Decision Process", Transportation Research Record: Journal of the Transportation Research Board, 2018	<1%

Publication

54	id.scribd.com Internet Source	<1%
55	etheses.whiterose.ac.uk Internet Source	<1%
56	Qiangqiang Guo, Li Li, Xuegang (Jeff) Ban. "Urban traffic signal control with connected and automated vehicles: A survey", Transportation Research Part C: Emerging Technologies, 2019 Publication	<1%
57	Yan Zhao, Jianfeng Zheng, Wai Wong, Xingmin Wang, Yuan Meng, Henry X. Liu. "Various methods for queue length and traffic volume estimation using probe vehicle trajectories", Transportation Research Part C: Emerging Technologies, 2019 Publication	<1%
58	Yiguang Xuan, Benjamin Coifman. "Identifying Lane-Change Maneuvers with Probe Vehicle Data and an Observed Asymmetry in Driver Accommodation", Journal of Transportation Engineering, 2012 Publication	<1%
59	Gurcan Comert, Mecit Cetin. "Analytical Evaluation of the Error in Queue Length Estimation at Traffic Signals From Probe	<1%

Vehicle Data", IEEE Transactions on Intelligent Transportation Systems, 2011
Publication

60 Azadeh Emami, Majid Sarvi, Saeed Asadi Bagloee. "A neural network algorithm for queue length estimation based on the concept of k-leader connected vehicles", Journal of Modern Transportation, 2019 <1%
Publication

61 G.H. Chapman, S. Djaja, D.Y.H. Cheung, Y. Audet, I. Koren, Z. Koren. "A self-correcting active pixel sensor using hardware and software correction", IEEE Design and Test of Computers, 2004 <1%
Publication

62 Xiaoqin Luo, Dongfang Ma, Sheng Jin, Yue Gong, Dianhai Wang. "Queue Length Estimation for Signalized Intersections Using License Plate Recognition Data", IEEE Intelligent Transportation Systems Magazine, 2019 <1%
Publication

63 Francesco Viti, Henk J. van Zuylen. "Probabilistic models for queues at fixed control signals", Transportation Research Part B: Methodological, 2010 <1%
Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		