Estimating Queue Dynamics at Signalized Intersections from Probe Vehicle Data

Methodology Based on Kinematic Wave Model

Mecit Cetin

As vehicle-to-vehicle and vehicle-to-infrastructure communications technologies are evolving, data from vehicles equipped with location and wireless technologies provide new opportunities to observe traffic flow dynamics more precisely. A new methodology is presented for estimating the dynamics of vehicular queues at signalized intersections on the basis of the event data generated by probe vehicles. The methodology uses shock wave theory (i.e., the Lighthill–Whitham–Richards theory) to estimate the evolution of the back of the queue over time and space from the event data generated when probe vehicles join the back of the queue. The time and space coordinates of these events are used to develop a formulation for determining the critical points needed to create time–space diagrams or shock waves to characterize the queue dynamics. The methodology is applied to sample data generated from the microscopic traffic simulation software VISSIM. It is found that the proposed methodology is effective in estimating queue dynamics at traffic signals.

Within the past decade, there has been a growing interest in equipping vehicles with wireless communications and location technologies to enable new applications for improving roadway safety and efficiency. The U.S. Department of Transportation's Connected Vehicle Research program is a significant effort to make the vision of seamless vehicle-to-vehicle and vehicle-to-infrastructure communications a reality (*1*). Within this vision vehicles are aware of their own locations in the transportation system and will exchange useful information with other vehicles as well as with the infrastructure. By tracking the positions of these probe vehicles along roadway segments, a wealth of information is generated to precisely characterize traffic flow dynamics. This knowledge in turn allows the system operators to improve system efficiency by taking relevant control actions (e.g., retiming traffic signals and responding faster to incidents).

Even though the use of probe vehicles as a traffic data source has been investigated, for example, in the context of travel time estimation (*2–5*), there has been limited effort toward investigating the use of probe data for queue length estimation (*6–9*). With the deployment of vehicle-based information collection technologies, there will be great interest in capitalizing on the probe data for traffic signal timing and other intelligent transportation system (ITS) applications. In order to support the development of these applications, research is needed to understand how probe vehicle technology could potentially improve the estimation of desired parameters.

The focus here is on the use of probe vehicle data in the context of studying queuing phenomena observed at traffic signals. More specifically, by using the time and location data generated by probe vehicles, the evolution of the back of the queue is estimated over time. The methodology makes use of shock wave theory to estimate the evolution of the back of the queue over both time and space from the event data generated when probe vehicles join the back of the queue. Another common approach to study queuing dynamics is the cumulative curve method (*10*), but that method requires a count of all vehicles at specific locations (which is not available when market penetration of probes is less than 100%) and does not always represent the spatial extent of the queue explicitly. Although some researchers demonstrate that the spatial extent of the queue can be determined from cumulative count curves (*11, 12*), the construction of the back of the queue is easier in shock wave theory when the arrival rate is time dependent (*11*).

The kinematic wave model used here to describe traffic flow dynamics is the widely known Lighthill–Whitham–Richards (LWR) model (*13–15*). This model arises from the principle of the conservation of vehicles and a fundamental diagram that relates flow to density, and is well known to reproduce the important traffic flow features (e.g., shock waves). The original LWR model has been extended in numerous studies to represent additional complexities (*16–18*). It has also been used in recent studies by Ban et al. (*6, 7*) and Liu et al. (*19*) that are particularly related to this study.

Liu et al. use high-resolution event-based data from traffic signals and fixed sensors to predict the maximum queue lengths for each cycle at intersections (*19*). Ban et al. estimate queue lengths by using the travel times and delays measured by mobile sensors (probe vehicles) between some predefined virtual points before and after the signalized intersections (*6*). By using the delay patterns, they identify the critical points when the queue is maximized, minimized, or cleared within a cycle. Both studies make use of the LWR theory to predict changes in vehicle trajectories and the shock waves created by the traffic signals. Other researchers also use the LWR theory in estimating queue lengths (*20*).

Researchers also use statistical methods to predict queues at signalized intersections. For example, Comert and Cetin developed analytical models for errors in queue length estimation based on probe vehicle data (*8, 9*). Researchers have also developed Markov chain models to predict average queues and their standard deviations (*21*). However, these models are developed on the basis of point queue models and do not represent space.

Department of Civil and Environmental Engineering, Old Dominion University, Kaufman Hall 135, Norfolk, VA 23529-0241. mcetin@odu.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2315, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 164–172.
DOI: 10.3141/2315-17

The methodology presented here differs from the previous studies, in particular work by Ban et al., in that the queue estimate is simply based on the time–space coordinates of vehicles when they join the back of the queue in each cycle (*6*). From the vehicles observed in a cycle, data only from the first and last vehicles are used in the estimate. Furthermore, the formulation does not make any assumptions about the probe percentage nor does it require probe observations in each cycle, in contrast to the minimum requirement of two probe observations per cycle as indicated by Ban et al. (*6*). In addition, the methodology does not impose any restrictions on the extent of the oversaturation or the number of stops a single vehicle makes to get through the intersection. However, the methodology presented here has limitations since undersaturated conditions have not yet been handled, which is left for future work.

PROBLEM DEFINITION AND SETTING

To develop the formulation to estimate the queue, a number of simplifying assumptions need to be made. Figure 1 shows a sample time–space diagram with vehicle trajectories in which there is a residual queue at the end of each cycle. Shock wave lines demarcating discontinuities in traffic states based on the LWR theory are added. Of particular importance are the critical points indicated by R_n and Q_n, which characterize the dynamics of the back of the queue, where n denotes the cycle number.

The specific conditions for which the formulation is developed are described in the following subsections.

Roadway Geometry and Signal Timing

A single-lane road leading to a signalized intersection is considered. The traffic signal is assumed to operate on a fixed cycle with alternating green and red phases of equal duration. This feature should not be considered as a limitation of the work since the formulation can be easily adapted to varying cycle and phase lengths.

Probe Data

Even though probe vehicles can technically collect data continuously (e.g., every second), such a large data set is not needed to estimate the queue dynamics. In this study, only the location and time data when probe vehicles join the back of the queue are used. In other words, for each probe vehicle, a single data point is needed, which contains its location on the link and the time instant when it joins the back of the queue.

Probe Vehicle Population

No assumption is made about the percentage of probe vehicles in the traffic stream.

Traffic Flow

It is assumed that vehicles arrive randomly at the intersection with an unknown rate. The methodology does not account for scenarios in which vehicle platoons form because of an upstream signal.

Shock Wave Speeds

It is assumed that the speeds of the backward-moving shock waves A_nR_n and B_nQ_n in Figure 1 are known and equal. On the basis of the LWR theory, shock wave speeds are simply found by $\Delta q/\Delta k$, where the numerator is the difference in flows of the two traffic states separated by these shock waves and the denominator is the difference in

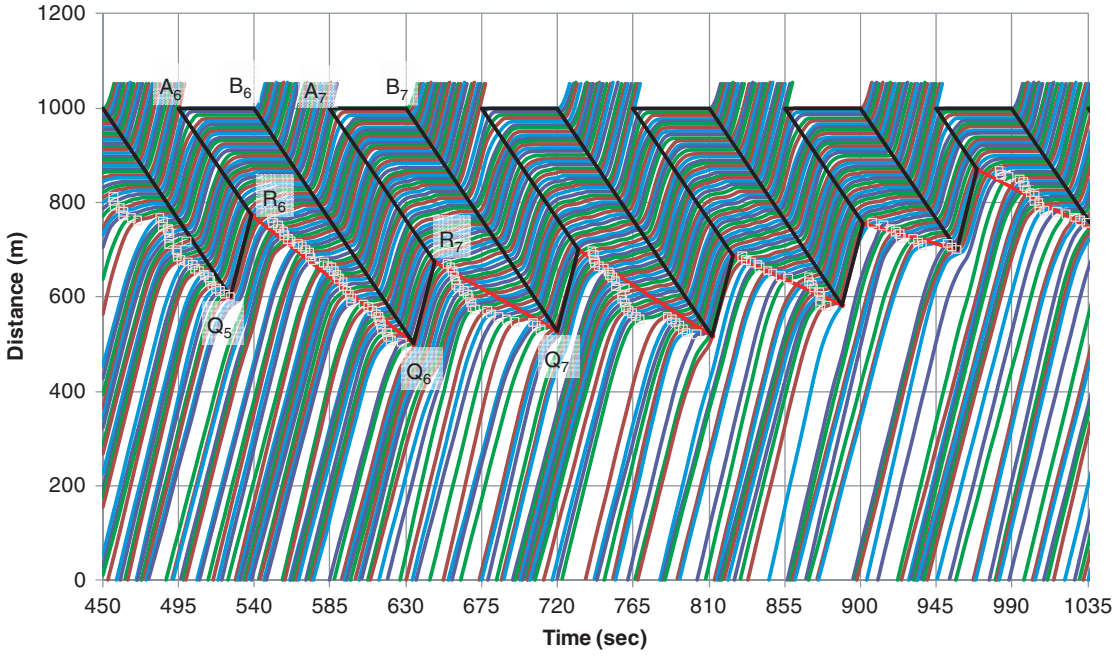

FIGURE 1 Vehicle trajectories and shock wave diagram (□ = linear regression line to probe vehicle data).

densities. Since these differences will be the same for shock waves A_nR_n and B_nQ_n in Figure 1, their speeds will also be the same.

This principle is used in other studies as well (*6, 19*). These speeds can be determined if a fundamental diagram is available or from vehicle trajectory data as shown in Figure 1. In addition, the speed of the forward shock waves Q_nR_{n+1} is known. These speeds can be assumed to be equal to the free-flow speed. A detailed discussion on how these speeds can be estimated is not within the scope of this study. The speeds of R_nQ_n shock waves for the queue growth are to be estimated from the probe data. These speeds depend on the cycle-to-cycle variations in vehicle arrivals.

Online Versus Offline Application

The methodology presented here is for offline applications; however, it can be extended for online prediction, which is left for future research.

Queue Conditions

The formulation is developed for oversaturated conditions in which there is a residual queue at the end of each cycle. The formulation will be extended to model both undersaturated and oversaturated conditions in the future.

METHODOLOGY

As mentioned earlier, the only input data from probe vehicles are their time and location information when they join the back of the queue. The goal is to estimate the critical points R_n and Q_n for each cycle n. The proposed methodology uses the LWR theory to develop a formulation to determine these unknown points. If the speeds for the shock waves representing queue growth (i.e., all R_nQ_n lines) are known for all cycles, the problem can be solved. However, these speeds (or the slopes for the straight R_nQ_n line segments) are unknown and need to be predicted from the probe vehicle data.

One might be tempted to fit a linear regression line to the probe vehicle data, shown with square symbols (□) in Figure 1, for each cycle to estimate these speeds (or slopes for the straight R_nQ_n line segments), but this method will not work since there may not be any probe observations in a cycle. Furthermore, such estimates may not be reliable since the number of data points can be very small, which may lead to unrealistic estimates. In addition, estimating these speeds independently will lead to inconsistencies since there is a fixed relationship between the coordinates of point Q_n and R_{n+1} (explained later in the discussion of Equations 3 and 4).

The approach proposed in this study is as follows. Within each cycle, the first and last probe vehicle observations (if any) are identified. Using the last probe observation in cycle n and the first observation in cycle $n + k$ ($n + k$ is the next cycle with some probe observations), a constant shock wave speed is estimated (α in Figures 2 and 3). In other words, the vehicle arrival rate between these two observation points is assumed to be constant (since shock wave speed depends on flow rate), which is realistic since there is no additional available information to be used. Once this constant speed is determined, the coordinates for all critical points (i.e., Q_n, R_{n+1}, Q_{n+1}, R_{n+1}, . . . , R_{n+k}) between the two probe observations are then estimated as explained in detail later. First, some notation and general relationships are introduced.

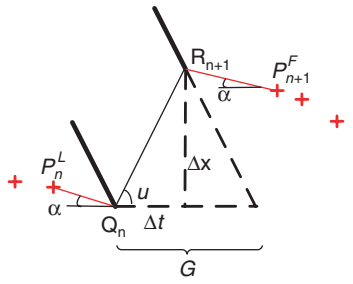

FIGURE 2 Observed time–space coordinates of some probe vehicles in two consecutive cycles (P^L_n and P^F_{n+1} denote, respectively, last probe vehicle in cycle n and first probe vehicle in cycle $n + 1$; u = shock wave speed; G = green phase).

Figure 4 shows a sample shock wave diagram and the critical points R_n and Q_n. The forward-moving shock waves (or interfaces) connecting Q_n and R_{n+1} have speed u for all n. On the basis of the assumptions made previously, the equations describing the shock wave lines starting at the stop bar and moving backward with a constant speed w can be easily written for a cycle n:

$$X_n^R = X_0 + w\left(t - \frac{2n+1}{2}C\right) \tag{1}$$

$$X_n^G = X_0 + w\left(t - nC\right) \tag{2}$$

where

X_n^R, X_n^G = shock waves starting at end of green and red phases, respectively, for cycle n;
X_0 = space coordinate of stop bar;
w = speed of shock wave; and
C = cycle length.

Since both u and the green phase length are constants, Δx and Δt shown in Figure 4 become constants for all cycles, which creates a fixed relationship between R_{n+1} and Q_n as explained later.

From the geometry in Figure 2, the relationship between the coordinates of points R_{n+1} and Q_n can be written as follows:

$$X_{n+1}^R = X_n^Q + \Delta x \tag{3}$$

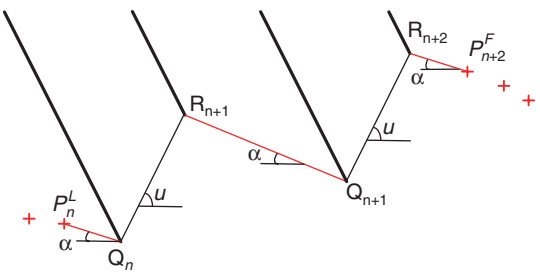

FIGURE 3 Observed time–space coordinates of some probe vehicles that join back of queue of cycle n and cycle $n + 2$.

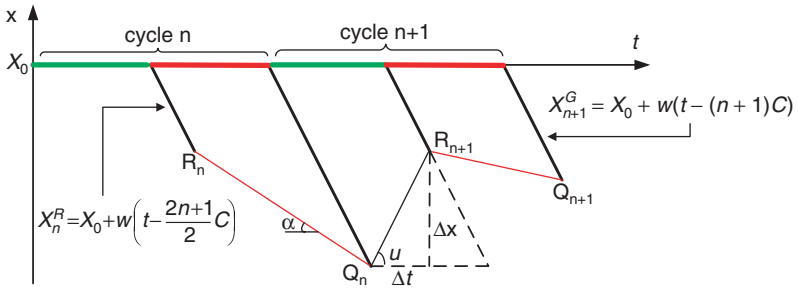

FIGURE 4 Shock waves at traffic signal under oversaturated conditions.

$$t^R_{n+1} = t^Q_n + \Delta t \tag{4}$$

where X^R_n and t^R_n denote the space and time coordinates for point R_n. Likewise, X^Q_n and t^Q_n denote the space and time coordinates for point Q_n.

With the known shock wave speeds u and w and the length of the green phase, G,

$$u = \frac{\Delta x}{\Delta t} \tag{5}$$

$$w = -\frac{\Delta x}{G - \Delta t} \tag{6}$$

Solving Equations 5 and 6 for the unknowns, Δx and Δt, yields

$$\Delta t = \frac{wG}{w - u} \tag{7}$$

$$\Delta x = u\frac{wG}{w - u} \tag{8}$$

From Equations 3 and 4, it should be clear that once point Q_n is known, point R_{n+1} is determined exactly.

Now, with the aid of Figure 2 the coordinates for points Q_n and R_{n+1} will be estimated from the probe vehicle data. For each cycle the time–space (t–s) coordinates of all probe vehicles (indicated by plus signs in Figure 2) when they join the back of the queue are known. Among these the first and last probe vehicles provide sufficient information to predict points Q_n and R_{n+1}.

Let (t^F_n, x^F_n) and (t^L_n, x^L_n) represent the coordinates of the first and last vehicles, respectively, when they join the back of the queue of cycle n. The difference in time (and space) coordinates of the first probe in cycle $n + 1$ and the last probe in cycle n can be written as follows:

$$t^F_{n+1} - t^L_n = \left(t^Q_n - t^L_n\right) + \Delta t + \left(t^F_{n+1} - t^R_{n+1}\right) \tag{9}$$

$$x^F_{n+1} - x^L_n = \left(x^Q_n - x^L_n\right) + \Delta x + \left(x^F_{n+1} - x^R_{n+1}\right) \tag{10}$$

Since it is assumed that the unknown arrival rate is constant between the two observed points,

$$\alpha = \frac{-\left(x^Q_n - x^L_n\right)}{\left(t^Q_n - t^L_n\right)} \tag{11}$$

Similarly, for the first point in the cycle $n + 1$,

$$\alpha = \frac{-\left(x^F_{n+1} - x^R_{n+1}\right)}{\left(t^F_{n+1} - t^R_{n+1}\right)} \tag{12}$$

where α is simply the tangent of the angle α shown in Figure 2 or the negative of the corresponding shock wave speed.

Equation 10 can be rewritten as follows by substituting the results from Equations 11 and 12:

$$x^F_{n+1} - x^L_n = -\alpha\left(t^Q_n - t^L_n\right) + \Delta x - \alpha\left(t^F_{n+1} - t^R_{n+1}\right) \tag{13}$$

or equivalently

$$\alpha\left[\left(t^Q_n - t^L_n\right) + \left(t^F_{n+1} - t^R_{n+1}\right)\right] = \Delta x - \left(x^F_{n+1} - x^L_n\right) \tag{14}$$

The summation of the two terms within the square brackets can be replaced by its equivalent from Equation 9:

$$\alpha\left[\left(t^F_{n+1} - t^L_n\right) - \Delta t\right] = \Delta x - \left(x^F_{n+1} - x^L_n\right) \tag{15}$$

Finally, the unknown α can be determined from the known coordinates of probe vehicles as follows:

$$\alpha = \frac{\Delta x - \left(x^F_{n+1} - x^L_n\right)}{\left(t^F_{n+1} - t^L_n\right) - \Delta t} \tag{16}$$

Now the coordinates for the critical points can be determined. The critical point Q_n is on the shock wave line X^G_n. Therefore, its coordinates should satisfy the following (according to Equation 2):

$$x^Q_n = x_0 + w\left(t^Q_n - nC\right) \tag{17}$$

Another relationship between the t–s coordinates of the critical point Q_n can also be written in terms of the known point P^L_n and α:

$$x^Q_n = x^L_n - \alpha\left(t^Q_n - t^L_n\right) \tag{18}$$

Solving Equations 17 and 18 simultaneously for the time coordinate t^Q_n results in the following:

$$t^Q_n = \frac{\left(x^L_n - x_0\right) + wnC + \alpha t^L_n}{w + \alpha} \tag{19}$$

Similarly, the *x*-coordinate can be determined as follows:

$$x_n^Q = x_0 + \alpha \frac{\left(x_n^L - x_0\right) - \alpha nC + \alpha t_n^L}{w + \alpha} \tag{20}$$

or

$$x_n^Q = x_0 + w\left(t_n^Q - nC\right) \tag{21}$$

since Q_n is on shock wave X_{n+1}^G (see Figure 4).

In summary, the *t–s* coordinates of Q_n can be found in terms of the observed probe data and known parameters by using Equations 19 and 21. Once the coordinates for the critical point Q_n are determined, the critical point R_{n+1} can be easily found by using Equations 3 and 4.

The previous formulation is for the scenario in which there are probe vehicle observations in two consecutive cycles. A more generic formulation is needed to account for other scenarios. Let n and $n + k$ be any two cycles for which probe observations are available. It can be shown that the unknown slope α for the back-of-queue shock wave is as follows:

$$\alpha = \frac{k\Delta x - \left(x_{n+k}^F - x_n^L\right)}{\left(t_{n+k}^F - t_n^L\right) - k\Delta t} \tag{22}$$

Equation 22 is the same as Equation 16 except that Δx and Δt are multiplied by k, which counts the number of green phases between cycles n and $n + k$. Equation 22 can be verified by following steps similar to those used for Equation 16.

Once α is determined from Equation 16, coordinates of Q_n can be determined by using Equations 19 and 20; subsequently R_{n+1} can be found by using Equations 3 and 4. To complete the formulation, point Q_{n+1} needs to be found as well. This step can be done by finding the intersection point of shock wave X_{n+1}^G (see Figure 4) and the linear line connecting R_{n+1} to Q_{n+1}. After some algebraic operations, the coordinates for point Q_{n+1} can be found as follows:

$$t_{n+1}^Q = \frac{\alpha t_{n+1}^R + x_{n+1}^R - x_0 + w\left(n+1\right)C}{w + \alpha} \tag{23}$$

Similarly, the *x*-coordinate can be determined as follows:

$$x_{n+1}^Q = x_0 + w\left(t_{n+1}^Q - \left(n+1\right)C\right) \tag{24}$$

The formulation for estimating all critical points R_n and Q_n from the probe vehicle data is now complete.

APPLICATION TO SIMULATION DATA

In order to show the application of the formulation and to test its performance in estimating the queue dynamics, a simple network is created in the microscopic simulation software VISSIM to generate the needed data. A single one-lane link of approximately 1 km is created. A traffic signal with 45-s green and 45-s red phases is created at the end of the link (cycle length = 90 s). All vehicles are passenger cars with the desired speed of 50 km/h and they enter the network at a constant rate but randomly. Simulation resolution is set to five time steps per second. All other parameters are kept at the default values built within VISSIM. After the simulation has been run with an arbitrary high vehicle input rate, the capacity of the signalized intersection is determined to be 1,050 vph after analysis of the output data.

The vehicle trajectories shown earlier in Figure 1 are from a scenario in which the vehicle input rate is set to 1,050 vph. The data shown with squares (□ in Figure 1) are the coordinates of the probe vehicles when they join the back of the queue. Vehicles are assumed to be stopped when their speeds drop below 5 km/h. These coordinates are found by simply determining the first time the vehicle speeds drop below this threshold.

In order to evaluate the performance of the formulation developed here, several scenarios are considered. These scenarios are created by varying the available probe vehicle data and the input flow rate. Figures 5 and 6 show the same simulation data as in Figure 1, except that the vehicle trajectories have been removed for clarity. In both plots, the signal is located at a distance of 1,000 m. Figures 5 and 6 show (*a*) the coordinates of all vehicles and a set of randomly selected probe vehicles when they first join the back of the queue, (*b*) the shock wave lines for the back of queue when all vehicle data are used (Back of Queue_all), and (*c*) the estimated shock wave lines for the back of the queue when only probe data are used (Back of Queue_Probes). In addition, the backward-moving shocks created because of signal phases are indicated with dotted lines. These dotted lines correspond to the scenario in which all vehicles are used to determine the shock waves and are only included for reference. In all scenarios, the critical points to construct the shock wave lines are found by the formulation presented in the previous section.

One could argue that rather than connect points R_n to Q_n directly (see Figure 1) to create the shock wave diagram, the critical points can be connected to the intermediate probe observations to generate a more precise description of the back of the queue. Even though this step is possible, it requires additional computations and does not seem to be important for practical applications. Therefore, the back of the queue here is simply characterized by the straight line segments connecting R_n to Q_n directly.

Figure 5 is for a scenario in which 10% of the vehicles are randomly selected to be probes, whereas in Figure 6 the probe percentage is 5%, a relatively low value compared with those from other studies (*6*). As can be observed in Figure 5, in one cycle (Cycle 10) no probe vehicles are observed. (The critical points *R* and *Q* for each cycle can be traced back to the stop line by following the dotted lines to read the clock time and determine the cycle number.) For the second scenario with 5% probes in Figure 6, no probe vehicles are observed in three consecutive cycles. In both scenarios, the estimated back-of-queue time–space "profile" from the limited probe data is reasonably close to the true profile obtained on the basis of data from all vehicles. Only a visual comparison between the estimated and true queue profiles is shown here. It is possible to calculate performance measures like total delay and travel times from the diagrams in Figures 5 and 6 to get a more quantitative comparison between the two cases. However, this exercise is left for future work. In any case, given the close correspondence between the estimated and true queue profiles, the quantitative results should be reasonably close.

To further test the methodology, additional scenarios are created by increasing the flow rate from 1,050 vph to 1,150 vph. The flow rate is kept constant throughout the simulation duration. Figures 7 and 8 provide similar information as in Figures 5 and 6, respectively, except that the flow rate to create the scenarios is different. In each graph, the shock waves change slope because of the randomness in vehicle arrivals as they are generated in VISSIM based on the speci-

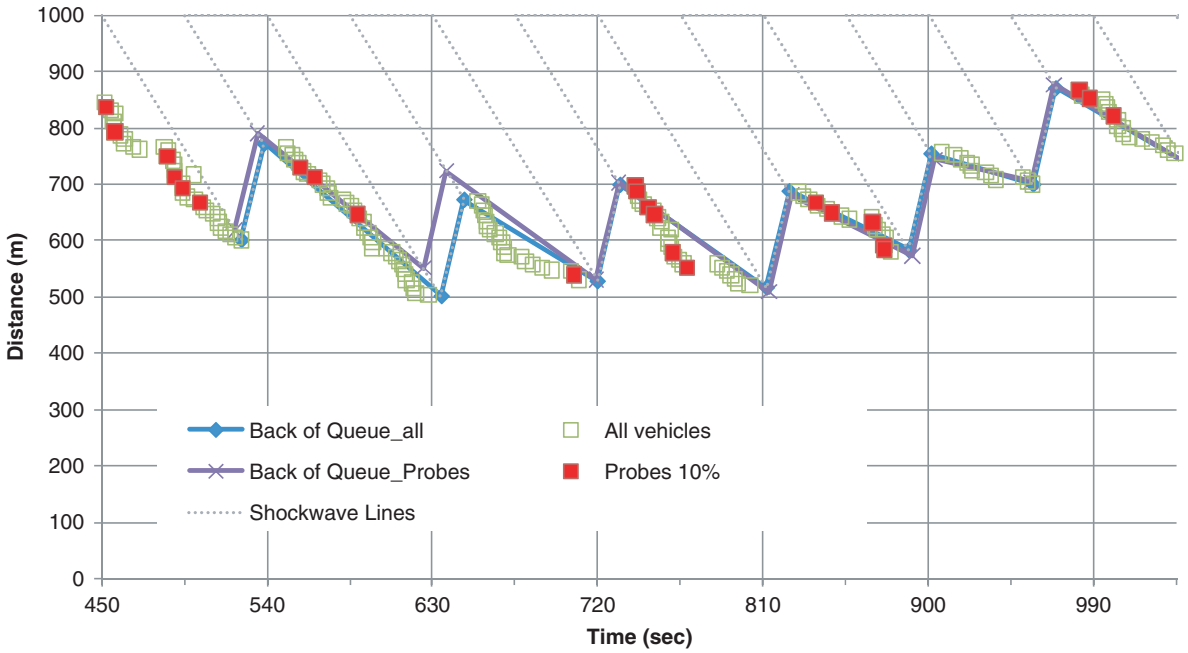

FIGURE 5 Back of queue estimated with all vehicles and 10% probes (input volume = 1,050 vph).

fied input flow rate. As can be observed, in the first several cycles shown in Figures 7 and 8 there is a more substantial growth in the queue in comparison with Figures 5 and 6. In the scenario with 10% probes shown in Figure 7, at least one probe data point is observed in each cycle, whereas in the scenario with 5% probes in Figure 8 two cycles do not have any probe vehicles. In both cases, the estimated back-of-queue profiles based on the limited probe data follow the true profile reasonably well.

The results shown so far in Figures 5 through 8 are for a single random sample of probe vehicles at a given rate. To investigate the sample variability, Figure 9 shows the back-of-queue profiles estimated from 10 replicas generated at a 5% probe rate when the vehicle input is 1,150 vph. The critical points for the true profile are indicated by solid diamond symbols. Overall, most of the estimated profiles follow the true profile reasonably closely. In several of the replicas, there is some deviation from the true answer, especially

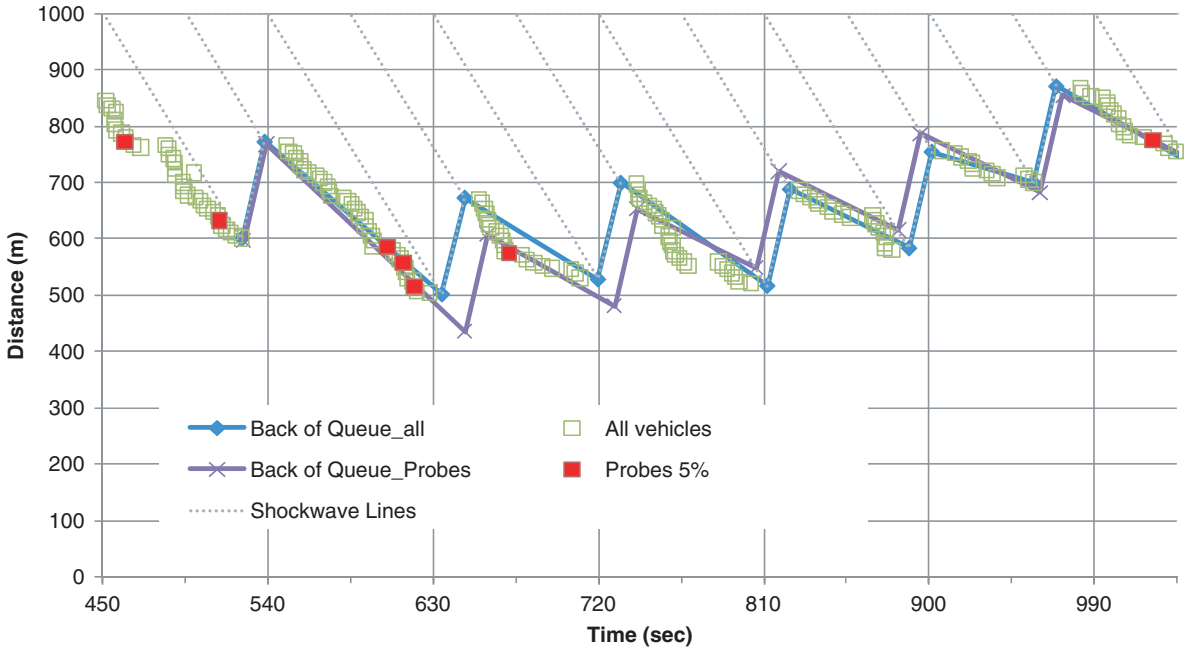

FIGURE 6 Back of queue estimated with all vehicles and 5% probes (input volume = 1,050 vph).

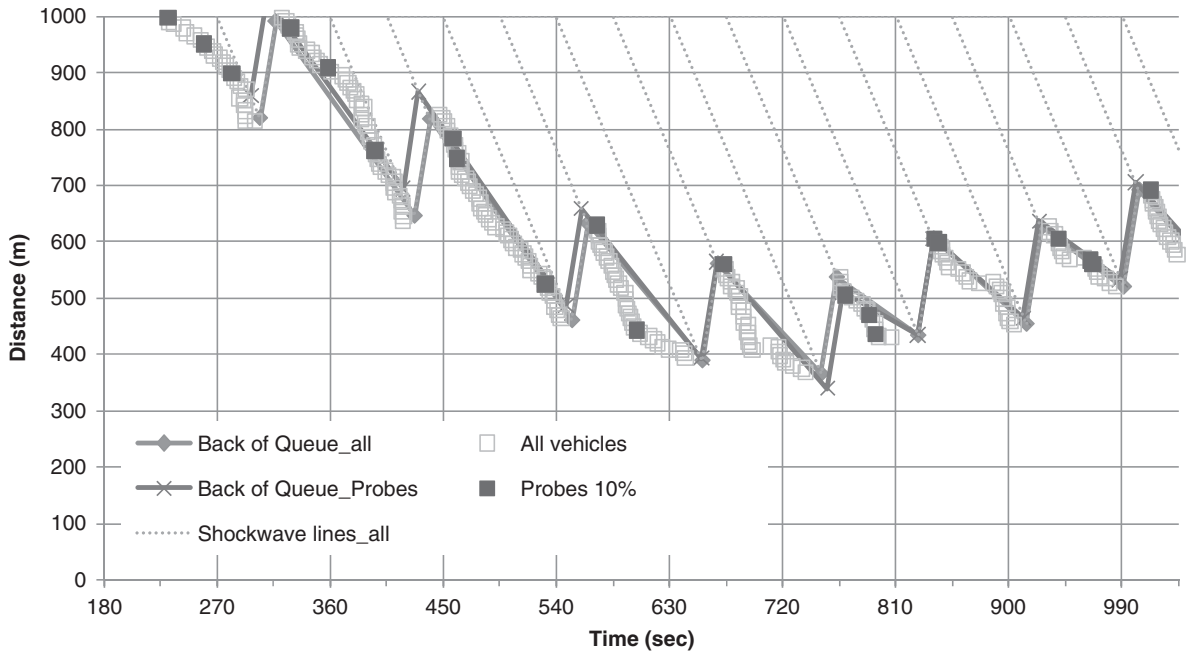

FIGURE 7 Back of queue estimated with all vehicles and 10% probes (input volume = 1,150 vph).

for Cycles 4 and 5, when there is a sudden jump in queue size. After examination of the data, it is observed that these differences tend to get larger when there are no probe data points near the critical points. For example, in Figure 8 relative to Q_3 and Q_5, probe observations are farther away from Q_4. Consequently, Q_4 is not estimated as accurately as Q_3 and Q_5. This is an important observation that may allow the modeler to assign reliability bounds to the accuracy of the predictions. More precisely, as the time separation between points P_n^L and P_{n+1}^F (the last probe in cycle n and first probe in cycle $n+1$) diminishes, the accuracy of the prediction for the critical point Q_n increases.

To gain further insights into the impacts of probe vehicles on queue prediction accuracy, summary statistics corresponding to four probe levels (i.e., 5%, 10%, 15%, and 20%) are shown in Table 1 for the scenario in which the input flow rate is set to 1,150 vph. The VISSIM model is run only once to generate vehicle trajectories. For each of the four levels, the probe vehicles are selected randomly among all vehicles to create 20 replicas. The back of the queue is

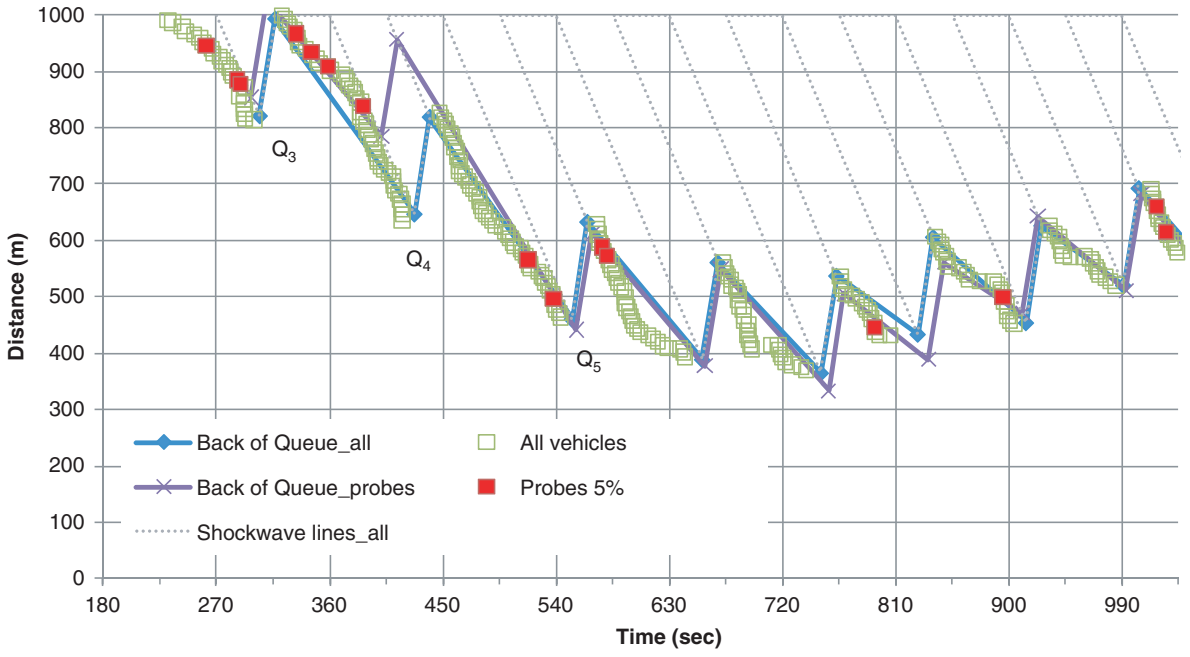

FIGURE 8 Back of queue estimated with all vehicles and 5% probes (input volume = 1,150 vph).

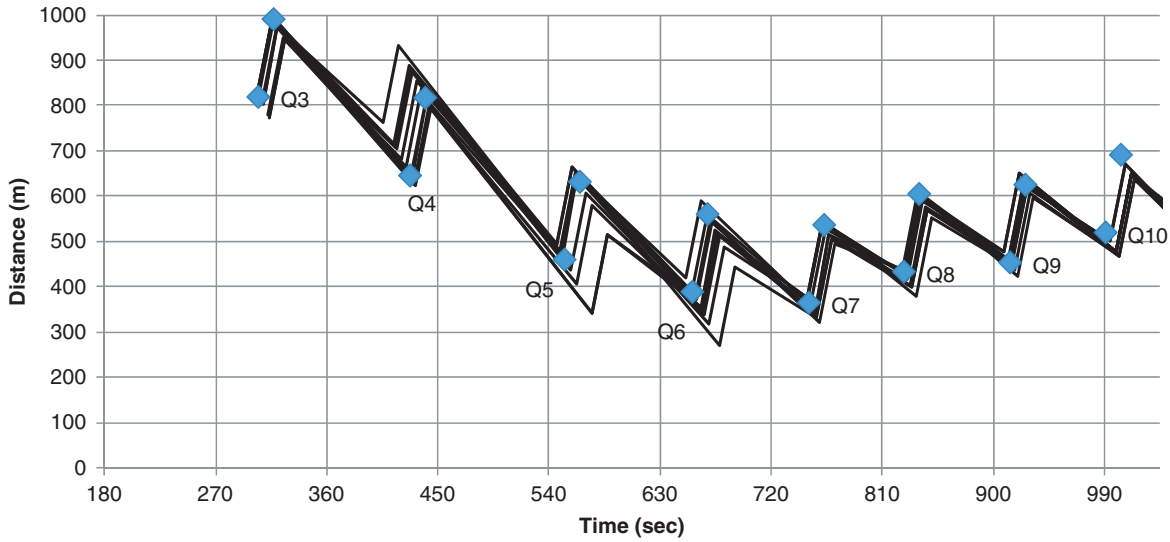

FIGURE 9 Ten back-of-queue profiles estimated with randomly selected 5% probes.

predicted for a total of 9 cycles in each replica. Therefore, there are 180 queue predictions ($20 * 9 = 180$) for each probe level. The queue is measured in meters from the stop bar (i.e., the vertical coordinate of point Q_n measured from the stop bar) and is compared with the predicted value in each cycle in each replica. The error is found by subtracting the estimate from the actual and by dividing the result by the actual queue length. The average and standard deviation of the error are shown in Table 1 along with the percentage of cycles (out of 180) that exhibit errors larger than ±10% and ±20%. Overall, it can be observed that error decreases as probe percentage increases, as expected. It is found that the predicted queue lengths are longer than the actual values since the average error is negative. This finding is perhaps an artifact of using a single VISSIM run to generate the input data. Additional runs at the same input flow level and at different flow levels will be conducted in the future to investigate the performance of the method more comprehensively.

DISCUSSION OF RESULTS

As illustrated in the previous section, the proposed method is effective in estimating the queue dynamics from limited probe vehicle observations. The methodology can be particularly useful, even with a relatively small number of probe observations, to evaluate the performance of signalized intersections experiencing significant congestion or oversaturation. As probe vehicle data become available, it is important to develop similar methodologies to make the best use of such data.

The methodology shown here provides a much richer understanding of congestion and system performance than just the travel times and delays that are typically estimated from probe data. For example, the evolution of queues over time (e.g., peak hours) can be estimated from limited probe data. Since signal timing data (e.g., phase and cycle lengths) are collected by some traffic operations centers, the formulation presented here can be employed to make the best use of the probe data.

The methodology can be extended to real-time applications to predict the back of the queue at every cycle (e.g., at the end of the red phase), which can help better optimize signal timing. For example, predicting how far the queue will grow for each cycle (i.e., more precisely, the critical points Q_n as described here) helps to determine the green time needed to clear the queue so that no residual queue is left in the next cycle. For example, by predicting Q_6 in Figure 1 at time $t = 540$, one can calculate the length of the next green time needed to clear the queue. This calculation can be done by extrapolating the Q_6 and R_7 line until it intersects the stop bar line, that is, $x = 1{,}000$.

In addition, the work presented here can help with the development of data collection policies from probe vehicles since the methodology shows what type of data is more useful than others to predict queues and consequently system performance. For example, in the context of this study, only event data when vehicles join the back of the queue for the first time are utilized. Even though vehicles may make subsequent stops before departing at the stop bar (e.g., under oversaturated conditions), the data pertaining to these other events are not needed for the methods developed here.

TABLE 1 Summary Statistics for Back-of-Queue Prediction at Different Probe Levels

Probe Level (%)	Avg (%)	SD of % Error	Percentage of Samples with Errors Greater or Less Than Given Thresholds			
			<–10%	>10%	<–20%	>20%
5	–4	10	20	6	6	2
10	–5	7	18	1	4	1
15	–2	6	8	3	0	1
20	–2	5	4	1	1	0

NOTE: Input volume = 1,150 vph. SD = standard deviation; avg = average.

CONCLUSIONS

A new methodology is presented to estimate the queue dynamics at signalized intersections based on probe vehicle data. The methodology is developed by relying on the LWR theory to determine the shock waves created by traffic signals. By analyzing the time–space

coordinates of probe vehicles in relation to these shock waves, a formulation is developed to estimate the back of the queue. Among all probe vehicles observed in each cycle, data only from the first and last probe vehicles are used as input in the estimation. The methodology does not require probe observations in every cycle. The application of the developed formulation is illustrated on the basis of data generated by microscopic simulation software. The results show that the method is effective in estimating the back of the queue under oversaturated conditions with a relatively small percentage of probe vehicles. The method produces reasonably accurate results even if some cycles do not contain any probe vehicle observations.

The work presented here can be extended to account for undersaturated conditions. Additional testing can be conducted with field data (e.g., NGSIM data) to evaluate its performance under real-world conditions. The methodology can be revised for real-time prediction purposes. Last but not least, more complex models to account for vehicle platoons and complicated intersections with turn lanes can also be developed.

ACKNOWLEDGMENTS

The author thanks the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

1. *Connected Vehicle Research.* RITA, U.S. Department of Transportation. http://www.its.dot.gov/connected_vehicle/connected_vehicle.htm. Accessed July 27, 2011.
2. Cetin, M., G. F. List, and Y. Zhou. Factors Affecting Minimum Number of Probes Required for Reliable Estimation of Travel Time. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1917,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 37–44.
3. Chen, M., and S. I. J. Chien. Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1719,* TRB, National Research Council, Washington, D.C., 2000, pp. 61–68.
4. Ferman, M. A., D. E. Blumenfeld, and X. Dai. An Analytical Evaluation of a Real-Time Traffic Information System Using Probe Vehicles. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations,* Vol. 9, 2005, pp. 23–34.
5. Srinivasan, K. K., and P. P. Jovanis. Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network. In *Transportation Research Record 1537,* TRB, National Research Council, Washington, D.C., 1996, pp. 15–22.
6. Ban, X. G., P. Hao, and Z. Sun. Real Time Queue Length Estimation for Signalized Intersections Using Travel Times from Mobile Sensors. *Transportation Research Part C,* Vol. 19, No. 6, 2011, pp. 1133–1156.
7. Ban, X., R. Herring, P. Hao, and A. M. Bayen. Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2130,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 109–119.
8. Comert, G., and M. Cetin. Queue Length Estimation from Probe Vehicle Location and the Impacts of Sample Size. *European Journal of Operational Research,* Vol. 197, 2009, pp. 196–202.
9. Comert, G., and M. Cetin. Analytical Evaluation of the Error in Queue Length Estimation at Traffic Signals from Probe Vehicle Data. *IEEE Transactions on Intelligent Transportation Systems,* Vol. 12, 2011, pp. 563–573.
10. Daganzo, C. F. *Fundamentals of Transportation and Traffic Operations.* Emerald Group Publishing Limited, Bingley, United Kingdom, 1997.
11. Lawson, T. W., D. J. Lovell, and C. F. Daganzo. Using Input-Output Diagram to Determine Spatial and Temporal Extents of a Queue Upstream of a Bottleneck. In *Transportation Research Record 1572,* TRB, National Research Council, Washington, D.C., 1997, pp. 140–147.
12. Erera, A. L., T. W. Lawson, and C. F. Daganzo. Simple, Generalized Method for Analysis of Traffic Queue Upstream of a Bottleneck. In *Transportation Research Record 1646,* TRB, National Research Council, Washington, D.C., 1998, pp. 132–140.
13. Lighthill, M. J., and G. B. Whitham. On Kinematic Waves. 1. Flood Movement in Long Rivers. *Proceedings of the Royal Society of London, Series A,* Vol. 229, No. 1178, 1955, pp. 281–316.
14. Lighthill, M. J., and G. B. Whitham. On Kinematic Waves. 2. A Theory of Traffic Flow on Long Crowded Roads. *Proc., Royal Society of London, Series A,* Vol. 229, No. 1178, 1955, pp. 317–345.
15. Richards, P. I. Shock Waves on the Highway. *Operations Research,* Vol. 4, No. 1, 1956, pp. 42–51.
16. van Lint, J. W. C., S. P. Hoogendoorn, and M. Schreuder. FASTLANE: New Multiclass First-Order Traffic Flow Model. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2088,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 177–187.
17. Wong, G. C. K., and S. C. Wong. A Multi-Class Traffic Flow Model: An Extension of LWR Model with Heterogeneous Drivers. *Transportation Research Part A,* Vol. 36, 2002, pp. 827–841.
18. Zhang, H. M., and W. L. Jin. Kinematic Wave Traffic Flow Model for Mixed Traffic. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1802,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 197–204.
19. Liu, H. X., X. Wu, W. Ma, and H. Hu. Real-Time Queue Length Estimation for Congested Signalized Intersections. *Transportation Research Part C,* Vol. 17, 2009, pp. 412–427.
20. Geroliminis, N., and A. Skabardonis. Prediction of Arrival Profiles and Queue Lengths Along Signalized Arterials by Using a Markov Decision Process. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1934,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 116–124.
21. Viti, F., and H. J. van Zuylen. Modeling Queues at Signalized Intersections. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1883,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 68–77.

The Traffic Flow Theory and Characteristics Committee peer-reviewed this paper.