

Article

Short-Range Prediction of the Zone of Moving Vehicles in Arterial Networks

Rouzbeh Forouzandeh Jonaghani ^{1,*}, Sepehr Honarparvar ² and Navid Khademi ³

¹ School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran

² Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran 1996715433, Iran; sep.honarparvar@mail.kntu.ac.ir

³ School of Civil Engineering, College of Engineering, University of Tehran, Tehran 1417613131, Iran; navid.khademi@ut.ac.ir

* Correspondence: r_forouzandeh@alumni.ut.ac.ir; Tel.: +98-937-086-2917

Received: 29 October 2017; Accepted: 18 January 2018; Published: 22 January 2018

Abstract: In many moving object databases, future locations of vehicles in arterial networks are predicted. While most of studies apply the frequent behavior of historical trajectories or vehicles' recent kinematics as the basis of predictions, consideration of the dynamics of the intersections is mostly neglected. Signalized intersections make vehicles experience different delays, which vary from zero to some minutes based on the traffic state at intersections. In the absence of traffic signal information (red and green times of traffic signal phases, the queue lengths, approaching traffic volume, turning volumes to each intersection leg, etc.), the experienced delays in traffic signals are random variables. In this paper, we model the probability distribution function (PDF) and cumulative distribution function (CDF) of the delay for any point in the arterial networks based on a spatiotemporal model of the queue at the intersection. The probability of the presence of a vehicle in a zone is determined based on the modeled probability function of the delay. A comparison between the results of the proposed method and a well-known kinematic-based method indicates a significant improvement in the precisions of the predictions.

Keywords: moving objects prediction; probability distribution function of delay; spatiotemporal models; arterial transportation networks; traffic signals

1. Introduction

Predicting the location of vehicles has been the focus of many researchers in recent years. In many cases, vehicles send their locations to a database so that a variety of systems utilize the information. Location-aware advertising [1], driving safety support systems [2], and vehicular ad hoc networks (VANET) [3] are examples of systems that consider the location of vehicles in their services. In some systems (e.g., driving safety systems), predicting the future location of vehicles is essential, because the systems should send some information (e.g., warnings) to vehicles before they reach specific locations. In some other systems (e.g., VANET), only the real-time location of a vehicle is required. Although the future location of vehicles is not required in the latter category, limitations in position updating make such systems predict the location of the vehicles.

Different statistical and nonparametric methods have been applied in moving object (especially vehicles) predictions [4]. Ying et al. [5] introduced a prediction model based on a cluster-based prediction approach. They predicted the future location of users based on the similarities between the attributes of users' semantic trajectories. Jeung et al. [6] proposed a hybrid method that combines an object's pattern information with motion functions to predict the future location of objects. In their proposed approach, a motion function is extracted from an object's recent movement.

They reported that their method surpassed the approaches in which an object's trajectory pattern is the only determinant factor in predictions.

If an object travels within a network, its movement is constrained by the geometry and dynamics of the network. In most metropolitan cities, highways are supervised by dedicated sensing systems, such as traffic cameras and loop detector networks. However, in many cases, especially in developing countries, arterial networks (unlike highways) are not thoroughly covered by traffic control infrastructure [7]. Moreover, private sectors (e.g., Internet taxi providers and delivery services) sometimes do not have access to real-time traffic information due to financial or technological limitations. Using information collected by mobile sensors could be an inexpensive alternative for estimating traffic states and predicting the future locations of vehicles. Civilis et al. [8] proposed a tracking and update method for vehicles moving within road networks. They used the kinematics of vehicles as the basis for location prediction. In their work, routes were divided into several segments with constant accelerations. The movement of moving objects was then reconstructed by applying a pre-computed acceleration profile and the speed of objects. Mo et al. [3] introduced an updating strategy by using Kalman filter prediction. They could enhance the selection of the data communication mode by improving the prediction of the location of vehicles. Their results also indicated a reduction in the data updating frequency, which was reported as a consequence of successful predictions. Reza et al. [9] applied the Dirichlet multinomial model to capture movement patterns in a metropolitan network. They could reduce the number of roadside units in the tracking operation.

By exploring the literature, it is determined that most research tries to predict a vehicle's location, presenting methods that estimate a single point as the future location of a vehicle. Although this might be the ultimate goal of any prediction model to determine the exact location of a moving vehicle, in most cases, especially in arterial networks, it is far too unrealistic to predict a single point where a vehicle will be, even in the near future. In arterial networks, the different behaviors of drivers, varying lengths of vehicle queues at intersections, the stochastic arrival time of vehicles, and random turning at intersections are some of the variables that make location prediction stochastic.

Instead of predicting a single point as the location of a vehicle, we estimate the probability that a vehicle will be in a specific zone after a time interval (which is called a prediction horizon). In this paper, a zone is defined as a continuous area that divides arterial links into smaller parts. Predicting a group of points (the segments located in a zone) as the future location of a vehicle, instead of a single point, might, at first glance, seem to be a step backward in the prediction of moving vehicles; however, in many applications, the main challenge is in predicting the probable areas where a vehicle might be present. For example, in vehicular ad hoc networks (VANETs), messages are communicated between roadside units (RSUs) and onboard units (OBUs). The number of triggered messages by different OBUs is limited by identifying the probable areas in which a target vehicle may exist [9]. In a more general case, in wireless sensor networks (WSNs), positioning the moving targets is vital to energy-efficient communication [10]. A sensor node in a WSN covers some adjacent areas (faces of the network) efficiently, so it is of great importance to determine the face that covers a moving target efficiently. As a further example, warning systems operating based on location-based services notify vehicles based on their probable presence in danger zones [11]. Re-identifying objects using cameras with non-overlapping coverage is also an area that would leverage zone predictions. In such problems, an object is tracked by more than one camera. When an object disappears from a camera, other cameras must re-identify the object. Some researchers have tried to reduce the number of candidates in the re-identifying process by taking the kinematics of objects into account [12]. Qualifying the search area using images, by considering the predicted zones, although not examined heretofore, seems to be a promising approach.

In this paper, the probability of the presence of a vehicle in different zones in the future is estimated by taking the dynamics of intersections into account. Vehicles experience some delays at intersections, which are intrinsically uncertain [13]. For example, a vehicle that arrives at an intersection during the green phase is prone to experience a shorter delay than a vehicle that arrives during the red phase. While some research has attempted to handle this uncertainty by introducing the log-normal or other families of distributions as the delay distribution function [14], the physical dynamics of intersections are almost neglected.

The sparseness of historical data is also addressed in this paper. The limited number of the observations is a challenge in modeling vehicles' movements. The proposed method is flexible to the sparseness of data by aggregating observations gathered over different days, and also by adopting some simplifying assumptions. The data used in this study were gathered using 320 GPS devices from taxi cabs in Roma, Italy [15]. The dataset contains the trajectories established by the GPS points, which were recorded every 15 s.

Section 2 presents more details on the dimensions of the problem. In Section 3, some introductory concepts are discussed. The proposed method is introduced in Section 4, and the results are discussed in Section 5. Section 6 contains a summarized conclusion of the research.

2. Problem Definition

A city can be divided into zones, either virtually or physically (Figure 1). A vehicle traveling in a road network in a city moves from one zone to another. In many systems, the future zone in which a vehicle will be located has to be predicted.

While a vehicle travels in a road network, its future location is affected by the dynamics of the network. In arterial networks, vehicles experience delays due to the presence of signalized intersections. In other words, the future locations of vehicles are affected by signalized intersections. The delay experienced by a vehicle at an intersection can be estimated in the presence of the dynamics of the intersections (i.e., the length of the green and red phases, length of the queue at the intersection, and the number of the lanes) and start time of the signal cycle in a particular time system. However, the mentioned parameters are not available in most cases because of technological or administrative constraints. In such a situation, the experienced delay by different vehicles remains stochastic, while the arrival times of vehicles at intersections, and the queue lengths, are unknown. The stochastic delay should be modeled somehow so that the future location of vehicles is predicted realistically.

In different research, the forecasting horizon varies from a few seconds (short-range) to tens of minutes (long-range). In long-range predictions, the kinematics of the vehicles and dynamics of the network are not the only determinative factors in the location of the vehicles. In such cases, the context of passengers is also required in the predictions [16]. This study primarily aims to predict the location of vehicles in the near future, which is defined to be in the order of tens of seconds. In addition, as we are not going to predict the final destination of a vehicle, we do not consider the turning probabilities in our model. In other words, we assume that the paths of vehicles are known.

This paper presents a method to estimate the probability of the presence of a vehicle in different zones in the future. This is done by modeling the probability distribution function (PDF) and cumulative distribution function (CDF) of the future location of vehicles, at any point on the links in arterial networks. These functions are modeled in terms of the dynamics of the intersections and their parameters are estimated using historical GPS data. While we are seeking a short-range prediction, the probability of the presence of a vehicle is found in zones that intersect the current and adjacent links of a vehicle.

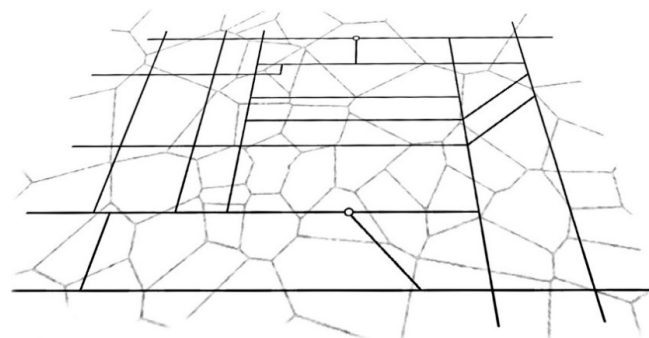


Figure 1. The links of the road network are divided by zones.

3. Preliminaries

3.1. Queue Model at Signalized Intersections

A classic queue model at signalized intersections was used in this study [13]. According to this model, the queue grows at speed q during the signal cycle (red and green phases), and vehicles forming the queue move at speed s during the green phase. It means that the queue length increases at speed q during the red phase and decreases at speed $(s - q)$ during the green phase. If the queue is fully removed at the end of the green phase, the intersection is deemed undersaturated. Otherwise, the intersection is deemed oversaturated. In an oversaturated intersection, vehicles have to stop for more than one signal cycle so that they can pass the intersection. The queue length is:

$$l = Q(t) = \begin{cases} n_0 + qt & \text{if } t \leq t_r \\ n_0 + qt_r - (s - q)(t - t_r) & \text{if } t > t_r \end{cases} \quad (1)$$

where $t = 0$ is the start time of the red phase, t_r is the signal red time, and n_0 is the initial queue length at the beginning of the red phase (which is referred as initial queue in this paper). In an undersaturated intersection, the initial queue is equal to zero.

3.2. Delay Distribution at Signalized Intersections

Van Zuylen et al. [13] modeled the analytical delay distribution at signalized intersections. According to their study, the delay distribution function for an undersaturated intersection is:

$$P(w) = \alpha \delta(w) + \beta \quad (2)$$

in which $\alpha = 1 - t_r \{C(1 - \frac{q}{s})\}^{-1}$, $\beta = \{C(1 - \frac{q}{s})\}^{-1}$, C is the cycle time (sum of red and green time), and $\delta(w)$ is the Dirac delta function.

The delay distribution function for an oversaturated intersection is also calculated as:

$$P(w|n_0) = \{C(1 - \frac{q}{s})\}^{-1} \left\{ B(w, C + \frac{(n_0 + t_g s)}{q}, t_r + \frac{n_0}{s}) + B(w, 2t_r + \frac{n_0}{s} - C(1 - \frac{q}{s}), C + \frac{n_0 - t_g s}{q + t_r}) \right\} \quad (3)$$

in which t_g is the green time and B is the block function, which is defined as:

$$B(x, x_1, x_2) = \begin{cases} 0 & \text{if } x < x_1 \\ 1 & \text{if } x_1 \leq x \leq x_2 \\ 0 & \text{if } x > x_2 \end{cases}$$

Equations (2) and (3) are the probability distribution functions of the total delay experienced by vehicles that enter an intersection. A vehicle experiences the total delay only if it enters the intersection. Otherwise, the vehicle does not experience the total delay. For example, a vehicle that is in the queue has not yet experienced the total delay. In this paper, the probability distribution of the total delay is used to determine the parameters of the proposed model.

3.3. Movement Modeling Assumptions

I. Time discretization: Arterials are assumed to have discrete dynamics in time. This means that we divided the daytime into intervals (e.g., 10:00 to 10:30, 10:30 to 11:00, and so on) and assumed that the dynamics of our model were stationary during these intervals on different days (work days and holidays are modeled separately). Based on this assumption, parameters of the queue model are constant during the estimation intervals. Furthermore, a queue exhibits a periodic behavior with period C (length of the light cycle) in each time interval. Although the assumption is not fully compatible with the reality of arterial traffic, it is still acceptable since our method is proposed to deal with the

sparseness of data. Moreover, a time discretization assumption is widely accepted by researchers in the traffic modeling area [17].

II. Neglected overtaking: Under this assumption, all vehicles travel at the same speed. The vehicles travel at a free flow speed when they are out of the queues, and follow queue dynamics when they are part of a queue. In addition, acceleration and decelerations are neglected. These assumptions are considered good approximations in arterials.

III. Continuous changes in the queue length: Although the queue length increases in a discrete manner (by joining a vehicle to it), we assumed that the queue length increases continuously. While we are looking for movement patterns in the order of tens of seconds, subtle differences between these two models can be ignored.

4. Methodology

4.1. Probability Distribution Function (PDF) of the Future Location of a Vehicle

PDF of the future location of vehicle $X_{x_0, \tau}(x)$ determines the probability that a vehicle located at x_0 (at time t_0) will be at x (Figure 2) after prediction horizon τ . In this paper, the PDF of a future location of a vehicle was determined based on the PDF of the delay of the vehicle, which indicates the probability that a vehicle located at x_0 will have a delay w in time window t_0 to $t_0 + \tau$. The delay of a vehicle is defined as the difference between the travel time of a real vehicle and the travel time of a hypothetical vehicle that travels at free flow speed v_{ff} . According to this definition, delay w is:

$$delay(x) = w = \tau - \frac{x_0 - x}{v_{ff}} \quad (4)$$

Since there is a one-to-one relationship between the future location of vehicle x and experienced delay w , the PDF of the future location of vehicle $p_{x_0, \tau}(x)$ is determined based on the PDF of the (future) delay of vehicle $d_{x_0, \tau}(x)$:

$$p_{x_0, \tau}(x) = d_{x_0, \tau}(delay(x)) \quad (5)$$

For notational simplicity, the indices x_0 and τ are omitted in the next sections.

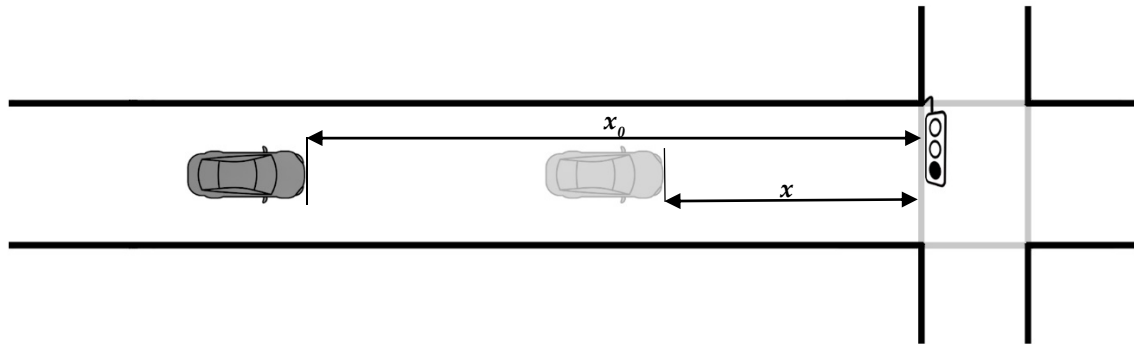


Figure 2. The distance between the vehicle and the downstream intersection is denoted by x .

4.2. PDF of the Delay

Although Equations (2) and (3) can be used as the probability distribution of the total delay that a vehicle experiences in an intersection, a vehicle does not necessarily experience the total delay at an intersection during time window t_0 to $t_0 + \tau$. Instead, in this time interval, a vehicle may still be waiting to enter the intersection or experiences a delay in more than one intersection. In the next section, the distribution of delay between t_0 to $t_0 + \tau$ for one intersection is modeled. Based on the modeled distribution, the probability of a certain delay for one or two intersections is estimated.

A vehicle experiences a delay if it joins a queue before entering the intersection. The vehicle may join a growing or a removing queue, which cause different delays. The delay is also dependent on the location at which a vehicle joins a queue. The PDF and CDF of the delay are the sum of the following terms:

1. The PDF and CDF of the delay of a vehicle, which is not in a queue at x_0 and joins a removing queue at x_j (Case 1).
2. The PDF and CDF of the delay of a vehicle, which is not in a queue at x_0 and joins a growing queue at x_j (Case 2).
3. The PDF and CDF of the delay of a vehicle, which is not in a queue at x_0 and does not join a queue (Case 3).
4. The PDF and CDF of the delay of a vehicle, which is already in a removing queue at x_0 (Case 4).
5. The PDF and CDF of the delay of a vehicle, which is already in a growing queue at x_0 (Case 5).

We model these cases for both undersaturated and oversaturated intersections. The PDF and CDF of the delay between t_0 and $t_0 + \tau$ for one intersection is the sum of the PDF and CDF (respectively) of Cases 1 to 5. The proof of the derivations for Case 1 is presented in the next section so that readers can understand the logic behind the derivations, but the details of the derivations of other cases (for the undersaturated intersections) can be found in Appendix A. The PDF of joining the queue at x_j is presented in the next section as a prerequisite for our model.

4.2.1. The PDF of Joining the Queue at x_j

Since the signal timing (start of the cycles) is not known with respect to the time system of vehicles (t_0), the queue length at t_0 remains a random variable. In other words, the probability that a vehicle joins a queue at x_j is independent of the location of the vehicle (x_0).

The probability that a vehicle joins a queue at x_j is proportional to the probability that queue length l is equal to x_j . We model the probability distribution of the queue length based on the queue model at intersections (Equation (1)).

In Figure 3, dl is a differential part of the queue. While it takes time dt as a growing queue to reach from l to $l + dl$ (or from $l + dl$ to l in a removing queue), the probability that the queue length is between l and $l + dl$ is $\frac{dl}{C}$, in which C is the total cycle duration. The differentiation of Equation (1) is:

$$dl = \begin{cases} q \times dt & \text{if } t \leq t_r \\ (s - q)dt & \text{if } t > t_r \end{cases} \quad (6)$$

Then dt is:

$$dt = \begin{cases} \frac{dl}{q} & \text{if } t \leq t_r \\ \frac{dl}{(s - q)} & \text{if } t > t_r \end{cases} \quad (7)$$

Based on Equation (7), the probability of queue length l for a growing queue ($t \leq t_r$) is $\frac{dl}{qC}$, which is a constant value. To model the PDF of this case, we used a uniform distribution function with support $[n_0, l_{\max}]$. The weight of this uniform distribution function is equal to the probability of the presence of a growing queue in a cycle, which is the ratio of the growing time ($\frac{l_{\max} - n_0}{q}$) to the cycle time (C). The PDF of the length of a growing queue, denoted by η^g , reads:

$$\eta^g(l) = U_{[n_0, l_{\max}]} \frac{l_{\max} - n_0}{qC} \quad (8)$$

in which $U_{[n_0, l_{\max}]}$ is the uniform distribution function with support $[n_0, l_{\max}]$.

Similarly, the PDF of the length of a removing queue ($t > t_r$), denoted by η^r , reads:

$$\eta^r(l) = U_{[n_0, l_{\max}]} \frac{l_{\max} - n_0}{(s - q)C} \quad (9)$$

The probability that the length of a growing queue is between l_1 and l_2 reads:

$$H_{l_1, l_2}^g = \int_{l_1}^{l_2} \eta^g = \frac{l_2 - l_1}{qC} \quad (10)$$

Similarly, the probability that the length of a removing queue is between l_1 and l_2 reads:

$$H_{l_1, l_2}^r = \frac{l_2 - l_1}{(s - q)C} \quad (11)$$

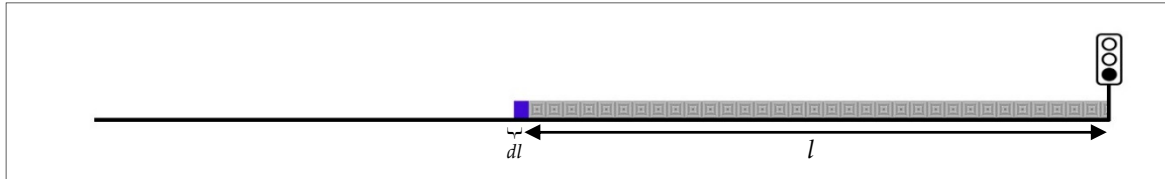


Figure 3. The differential part of a queue.

4.2.2. The Vehicle Joins a Removing Queue (Case 1)

In this case, the vehicle located at x_0 joins a removing queue at x_j (Figure 4). While the vehicle travels at the speed of v_{ff} , the vehicle reaches the queue after $\frac{x_0 - x_j}{v_{ff}}$. The remaining time of prediction horizon (τ) is denoted by τ_{rem} :

$$\tau_{rem} = \tau - \frac{x_0 - x_j}{v_{ff}} \quad (12)$$

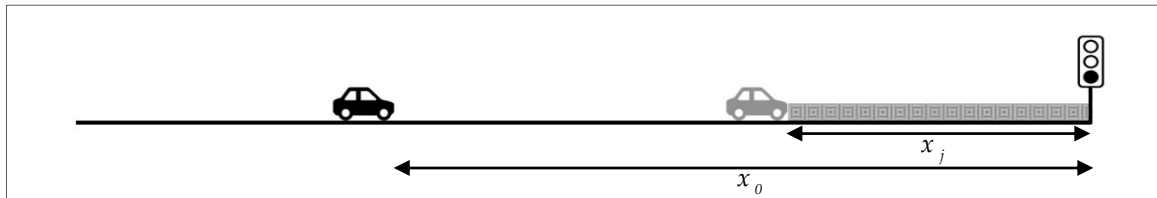


Figure 4. Illustration of Case 1. This case describes a vehicle that is located at x_0 at time t_0 and then joins the queue at x_j .

If τ_{rem} is smaller than zero, it means that a vehicle does not reach the queue before finishing prediction horizon τ , which means that it does not experience delays. After joining a growing queue, a vehicle continues its movement inside the removing queue at s speed. The time in which the vehicle travels inside the queue is denoted by τ_Q :

$$\tau_Q = \min \left\{ \tau_{rem}, \frac{x_j}{s} \right\} \quad (13)$$

in which $\frac{x_j}{s}$ is the time the vehicle needs to reach the signal ($x = 0$).

When a vehicle travels at speed s for time τ_Q , it experiences a delay compared with a vehicle traveling at the free flow speed. In this situation, the delay caused by traveling inside the queue is determined as:

$$\tau_Q - \frac{s\tau_Q}{v_{ff}} = \tau_Q \left(1 - \frac{s}{v_{ff}} \right) \quad (14)$$

in which is the difference between τ_Q and the time a hypothetical vehicle spends at the free flow speed v_{ff} to traverse the same distance.

According to the explanations, the delay w in the Case 1 is determined as:

$$w = \begin{cases} 0 & \text{if } \tau_{rem} \leq 0 \\ \tau_Q \left(1 - \frac{s}{v_{ff}}\right) & \text{if } 0 < \tau_{rem} \end{cases} \quad (15)$$

In Equation (8), w is a function of x_j . The location at which a vehicle joins a queue (x_j) is a random variable with a determined probability distribution (Equations (8) and (9)). We determined x_j as a function of w so that we could model the PDF of w based on the PDF of x_j . To have x_j as a function of w , Equation (15) is expanded using Equations (13) and (14):

$$w = W(x_j) = \begin{cases} 0 & \text{if } x_j \leq x_0 - \tau v_{ff} \\ \tau_{rem} \left(1 - \frac{s}{v_{ff}}\right) & \text{if } x_0 - \tau v_{ff} \leq x_j \leq \frac{s(x_0 - \tau v_{ff})}{s - v_{ff}} \\ \frac{x_j}{s} \left(1 - \frac{s}{v_{ff}}\right) & \text{if } \frac{s(x_0 - \tau v_{ff})}{s - v_{ff}} \leq x_j \leq l_{max} \end{cases} \quad (16)$$

Figure 5 shows a diagram of Equation (16):

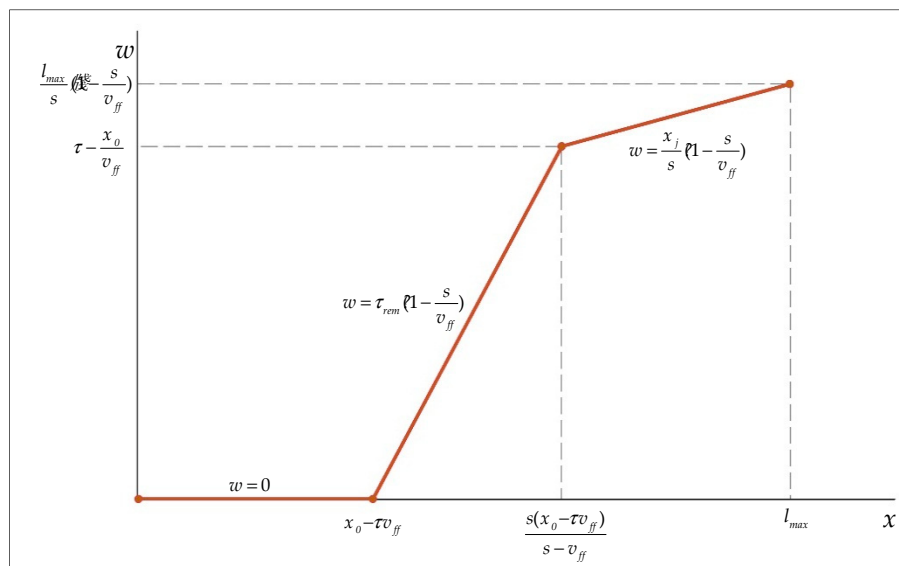


Figure 5. Delay w as a function of x_j in Case 1.

In Equation (16) (Figure 5), delay w is a hybrid function of x_j . While w is monotonic, x_j is also a function of w . x_j is derived by simply inverting the sub-functions in Equation (16):

$$x_j = X_j(w) = \begin{cases} [0, x_0 - \tau v_{ff}] & \text{if } w = 0 \\ x_0 - \tau v_{ff} - \frac{w v_{ff}^2}{s - v_{ff}} & \text{if } 0 < w \leq \tau - \frac{x_0}{v_{ff}} \\ \frac{s w v_{ff}}{v_{ff} - s} & \text{if } \tau - \frac{x_0}{v_{ff}} \leq w \leq \frac{l_{max}}{s} \left(1 - \frac{s}{v_{ff}}\right) \end{cases} \quad (17)$$

Equation (17) determines x_j for any delay w . Accordingly, the probability that a vehicle does not experience any delay ($w = 0$) during the prediction horizon is equal to the probability that the vehicle joins the queue at point $x_j \in [0, x_0 - \tau v_{ff}]$, which is a mass probability with a value of $H_{0, x_0 - \tau v_{ff}}^r$. We modeled this probability using a Dirac delta function (centered at 0) and its corresponding weight $H_{0, x_0 - \tau v_{ff}}^r$. The probability distribution of other defined values of w ($0 < w \leq \frac{l_{max}}{s} (1 - \frac{s}{v_{ff}})$) is

determined by the PDF of joining the queue at x_j (Equation (9)). The probability of undefined values of w is 0. The PDF of delay in Case 1 reads:

$$r^o(w) = \begin{cases} \delta(w)H_{0,x_0-\tau v_{ff}}^r & \text{if } w = 0 \\ \eta^r(X_j(w)) & \text{if } 0 < w \leq \frac{l_{\max}}{s}(1 - \frac{s}{v_{ff}}) \\ 0 & \text{if } w > \frac{l_{\max}}{s}(1 - \frac{s}{v_{ff}}) \end{cases} \quad (18)$$

Since w is monotonic (Figure 5), the cumulative distribution function (CDF) of w is equal to the cumulative probability of $x_j = X_j(w)$. The CDF of the delay in Case 1 reads:

$$R^{out}(w) = H_{0,X_j(w)}^r \quad (19)$$

4.3. CDF of the Delay for Two Consecutive Intersections

We do not present any solution for deriving the PDF of the delay for two consecutive intersections since it is not required in zone prediction. However, we introduce a method for modeling the CDF of the delay for two consecutive intersections.

Our future work addresses modeling the delay for multiple intersections. However, in this paper modeling the delay for two intersections is sufficient because we are looking for a short-range prediction. To model the CDF of the delay for two intersections, we assume that the experienced delays in two consecutive intersections are independent. This assumption is widely adopted in modeling the travel time of arterial links [17]. By considering the delay independence assumption, the CDF of a delay at two consecutive intersections is a convolution between the CDF of delays at individual intersections. The CDF of delay at intersections i and $i + 1$ is denoted as $D^i(w)$ and $D^{i+1}(w)$, respectively. The CDF of delay at two consecutive intersections, denoted as $D^{i,i+1}(w)$, reads:

$$D^{i,i+1}(w) = \int_0^w (D^i(w_1) * D^{i+1}(w - w_1))dw_1 \quad (20)$$

The CDF of the delay at an intersection is the sum of some linear terms (e.g., Equations (10) and (11)). We did not derive an analytical presentation for Equation (20); instead, we applied a numerical integration method. We used rectangular integration, which is an efficient method for the integration of polynomial functions [18].

4.4. Probability of Zones

The arterial links are segmented by zones (Figure 6 illustrates an example). The probability of a zone is defined as the probability that a vehicle located at x_0 falls into a segment within a zone after prediction horizon τ . It is obvious that the probability of a zone is not a constant quantity, and that its value is dependent on the location (x_0) of the vehicle. According to the definition, the probability of a zone (in this case, zone c in Figure 6) is:

$$P(\text{zone } c) = \int_{x_{c2}}^{x_{c1}} p(x)dx \quad (21)$$

in which $p(x)$ is the PDF of the future location of the vehicle (Equation (5)). Based on the Equations (4) and (5):

$$P(\text{zone } c) = \int_{x_{c2}}^{x_{c1}} p(x)dx = \int_{w_{c2} = \text{delay}(x_{c2})}^{w_{c1} = \text{delay}(x_{c1})} d(w)dw = D^{i,i+1}(w_{c1}) - D^{i,i+1}(w_{c2}) \quad (22)$$

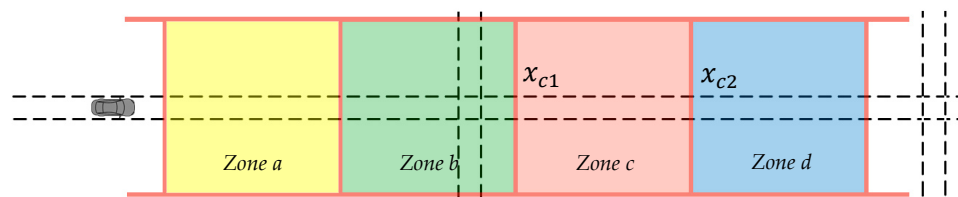


Figure 6. Links of the arterials are segmented by zones. The probability of zones is determined for all the zones a vehicle has ahead.

For a vehicle, the probability of all the zones ahead is determined. Based on the estimated probabilities, a set of priorities for the presence of the vehicle in the zones (after prediction horizon, τ) is defined. For example, in Figure 6 if $P(\text{zone } b) > P(\text{zone } c) > P(\text{zone } a) > P(\text{zone } d)$, then the priorities of the zones b , c , a , and d are 1, 2, 3, and 4, respectively. These priorities can be used in some applications; for example, in a wireless network, after an unsuccessful search for a vehicle in the zone with the highest priority (or the greatest probability), the system looks for the vehicle with the second priority.

4.5. Parameter Estimation

Zheng and Van Zuylen [19] indicated that the parameters of a delay distribution could be properly estimated for both undersaturated and oversaturated intersections using the trajectories of vehicles. They used delay distribution functions at intersections (Equations (2) and (3)) to determine the parameters of the dynamic model of the traffic at intersections. Zheng and Van Zuylen revealed that, compared with the minimum square estimator, the maximum likelihood estimator gets better results when estimating parameters. Based on their findings we used the maximum likelihood estimator to determine the parameters required in modeling the future location of vehicles. Based on the time discretization assumption, we estimated parameters t^s , t^r , n_0 , q , s , v_{ff} , and the state of an intersection (undersaturated or oversaturated) for time intervals. We aggregate the trajectories that are collected on different days but in the same time intervals (e.g., all the trajectory observations collected between 13:00 and 13:30 on work days). Similar to Zheng and Van Zuylen, in the absence of an analytical solution, we applied the genetic algorithm (GA) to estimate the best parameter set by applying a maximum likelihood estimator. Details of the parameter estimation method can be found in Zheng and Van Zuylen's paper.

5. Results and Discussion

The proposed method for predicting the location of vehicles can be utilized in a variety of relevant location-based systems (e.g., advertising services and driver safety systems). While we are not going to limit the results to a specific application, we adopted some different zonings to explore the precision of the proposed method in different scenarios. In many service networks (e.g., wireless networks), space is divided into Voronoi polygons around the service points [10]. Applying Voronoi diagrams or grids are also common approaches for indexing moving objects in spatial databases [20,21]. Although the proposed method is not built on a presumed zoning approach, constant interval grids with cell sizes of 25 m and 50 m (Figure 7a) and Voronoi polygons (Figure 7b) are the basis of our analysis so that some usual zonings are taken into consideration. In this paper, the centers of the Voronoi polygons are random points with an average distance of 100 m. In selecting the size of the Voronoi polygons and grids, we chose relatively small sizes (between 25 m and 100 m) since we are going to examine the precision of a short-range prediction method.

To explore the efficiency of the proposed method, the results should be compared with existing methods of predicting the future location of the vehicles. While some relevant methodologies are found in the literature [22], most of them are dependent on the additional data sources (e.g., flow information derived from inductive detectors). Additionally, the existing methodologies that only use the positional information are not designed for short-range prediction of the location of vehicles in arterial networks. In this paper, the results are compared with a kinematics-based method in which

the future location of a vehicle is estimated regarding the kinematics of the vehicle derived from its recent trajectory [8]. Applying the kinematics of the vehicles in location prediction is a conventional approach in moving object databases [23,24], but the accuracy of its results is limited in the arterial networks, as it neglects the dynamics of the intersections.

The data used in this study were collected by 320 GPS devices in taxi cabs in Roma, Italy. The dataset contains a one-month period of trajectories, established by GPS points, which were recorded in 15-s intervals. We chose an area containing seven arterial links and eight traffic signals (Figure 8) on Angelico and Delle Milizie streets. The lengths of the links are between 180 m and 440 m. We applied 18,000 GPS records in the study area. The data collected in the first three weeks were applied in the parameter estimation (the historical data), and the prediction precision was estimated using the rest of the data, which were the real-time data in our experiments. The future zone of vehicles was predicted for 15-, 30-, 45-, and 60-s prediction horizons in 300 randomly-chosen sample points, collected from the real-time trajectories (the selected prediction horizons were multiples of 15 so we could extract the correct future location of the vehicles from the trajectories and estimate the precision of the predictions).

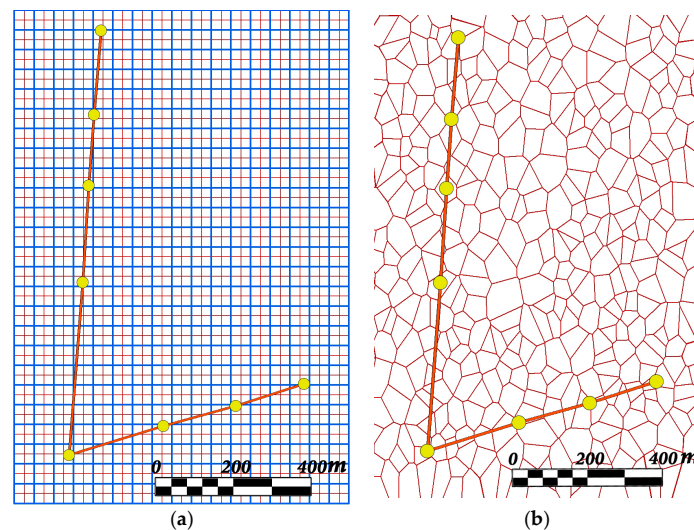


Figure 7. (a) Indicates a 25-m and a 50-m grid; (b) indicates a Voronoi diagram that divides the links into 30- to 100-m segments.



Figure 8. Study area on Google Maps.

A proper method to assess the correctness of a modeled probability distribution function is to compare it with the empirical distribution of the random variable. However, this is not a feasible method in our case, because it is not possible to collect the future location of vehicles at any given point (x_0) and prediction horizon (at least by using the sparse trajectories). We define some metrics to examine the efficiency of the proposed method. First, we compared the results of our method with a common prediction approach (kinematics-based). Second, we looked for a relationship between the estimated probability of the zones and the true zones extracted from the trajectories.

For all 300 sample points, we determined the probability of the zones vehicles had ahead. We considered the zone with the greatest probability as the predicted future zone of a vehicle. At the same time, we determined the future zone of a vehicle using the kinematics-based method. Since the prediction horizons are on the order of tens of seconds, we considered the mean speed of vehicles as the kinematics of their movement. Tables 1 and 2 indicate the results of both zoning methods described in Figure 7.

Table 1. Results of the proposed method and kinematics-based method for the grid zoning described in Figure 7a.

Prediction Horizon (s)	Percentage of Correct Predictions			
	Kinematics-Based Method		Proposed Method	
	25-m Grid	50-m Grid	25-m Grid	50-m Grid
15	42%	53%	47%	62%
30	30%	41%	42%	55%
45	26%	39%	33%	41%
60	19%	31%	30%	40%

Table 2. Results of the proposed method and kinematics-based method for the Voronoi zoning described in Figure 7b.

Prediction Horizon (s)	Percentage of Correct Predictions	
	Kinematics-Based Method	Proposed Method
15	57%	63%
30	48%	56%
45	47%	54%
60	35%	50%

Tables 1 and 2 indicate that, in different zoning scenarios and prediction horizons, the proposed method had significantly better results than the kinematics-based method. While the proposed method takes the physics of arterials into account, this can be proposed as the main reason behind the improvement in the results.

The precision of the prediction also has a meaningful relationship with the initial location of the vehicle (x_0). Figure 9 illustrates that when the prediction horizon is 15 s and 30 s, the precision of the prediction reduces for vehicles that are closer to the intersection. It seems that vehicles have a more predictable movement when they are far from the intersection. In other words, while the movement of the vehicles is affected by the queues at intersections, the precision of the predictions decreases near the intersections. However, when the prediction horizon is greater (in this case 45 s and 60 s), the vehicles reach the intersections and their movements are affected by intersections. This can be proposed as the reason why the precision of predictions is not related to the initial location of a vehicle in greater prediction horizons.

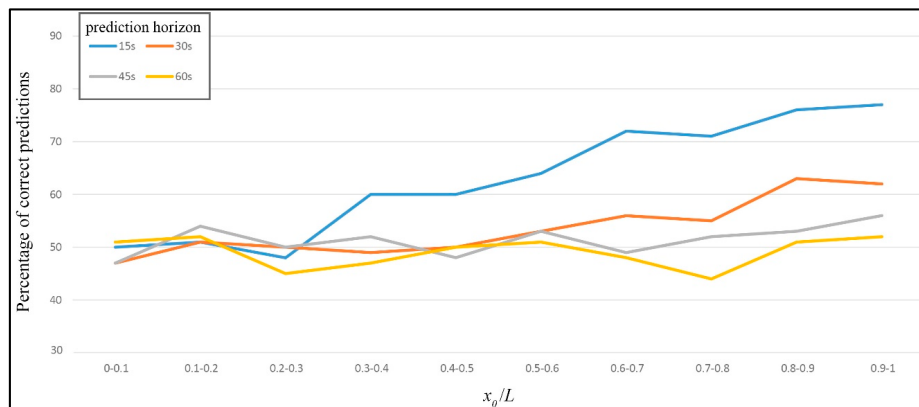


Figure 9. Changes in the percentage of the correct predictions with respect to the initial location of the vehicle.

Figure 10 indicates the average of estimated probabilities and the percentage of realizations in zones for different priorities (priority 1 to 5 are presented in the graphs). These graphs reveal a relationship between the percentage of the realizations and average of the probabilities of zones with specific priorities. Although we were not able to validate the modeled probability functions directly (by comparing the modeled functions with empirical distributions), this relationship indicates that our model follows the real-world movement of vehicles in arterials.

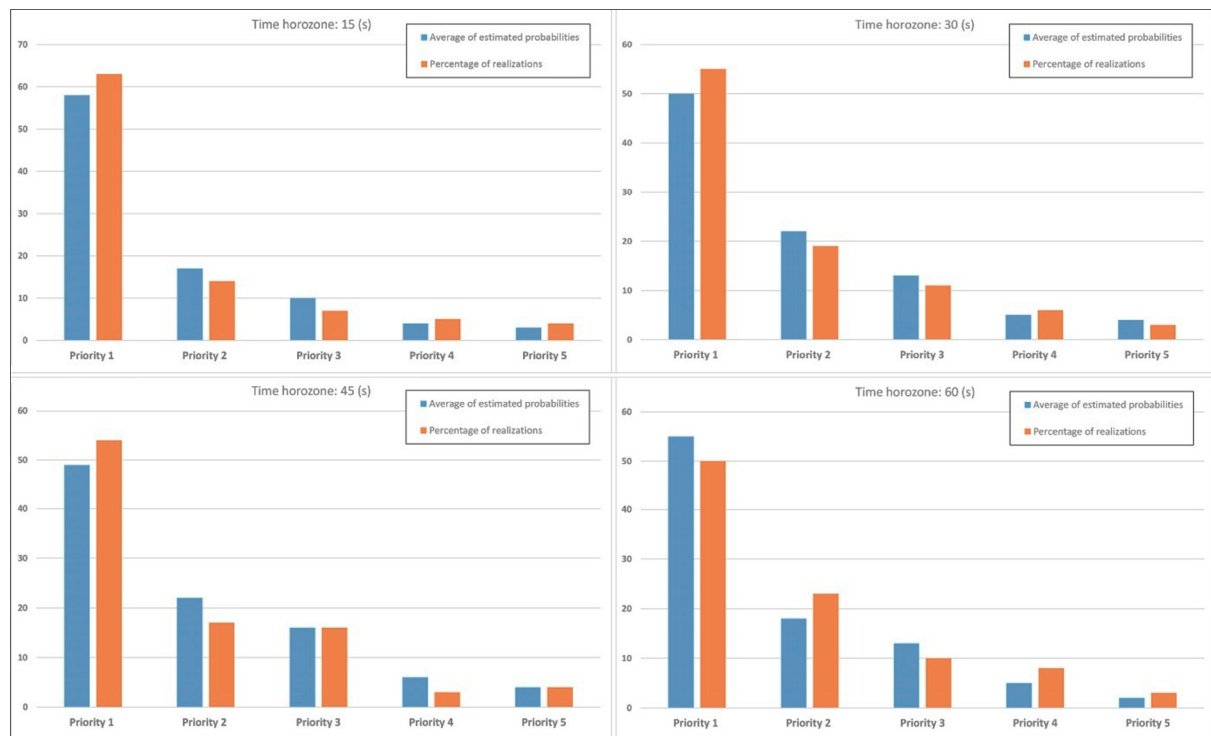


Figure 10. The average of estimated probabilities and the percentage of realizations in zones in different priorities.

6. Conclusions

In this paper, a new method for predicting the location of vehicles in arterials is proposed and validated. The method estimates the future zone of a vehicle based on the CDF of the delay of the

vehicles. The results indicate a promising improvement compared with a kinematics-based method that is regularly used in the databases. Especially, when the prediction horizon is longer, the results indicate a more significant improvement in the precision of the predictions in comparison with the kinematic-based method. The observed improvement may be related to the result of considering the dynamics of the intersections in the prediction model.

The proposed method not only predicts the most probable zone of vehicles but also determines the priorities for the presence of vehicles in different zones. While the geometry of the zones is not a determinant in the proposed model, our model can also be used in dynamic scenarios in which the zones change over time.

Author Contributions: Rouzbeh Forouzandeh Jonaghani conceived and designed the methodology; Sepehr Honarparvar analyzed and prepared the data; Rouzbeh Forouzandeh Jonaghani and Sepehr Honarparvar implemented the methodology; Rouzbeh Forouzandeh Jonaghani wrote the paper; and Navid Khademi and Sepehr Honarparvar revised the methodology and the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. PDF and CDF of Delay at Undersaturated Intersections

Appendix A.1. The Vehicle Joins a Growing Queue (Case 2)

A vehicle that is initially located at x_0 joins the growing queue at x_j . While the vehicle travels at a speed of v_{ff} , the vehicle reaches the queue after $\frac{x_0 - x_j}{v_{ff}}$. The remaining time is denoted as τ_{rem} :

$$\tau_{rem} = \tau - \frac{x_0 - x_j}{v_{ff}} \quad (A1)$$

If τ_{rem} is smaller than zero, it means that the vehicle does not reach the queue before the end of the prediction horizon τ , which means that it does not experience any delays. After joining a growing queue, a vehicle completely stops until the queue reaches its maximum length l_{max} . The stopping time is equal to $\frac{l_{max} - x_j}{q}$. If the stopping time is greater than τ_{rem} , the vehicle is completely stopped before the end of the prediction horizon; thus, the delay equals τ_{rem} . Otherwise, if the remaining time is greater than the stopping time, the vehicle starts moving with the removing queue (at the speed s) after the stopping time. The vehicle travels inside the queue until it reaches the intersection. The time in which the vehicle travels inside the queue (at the speed s) is denoted by τ_Q :

$$\tau_Q = \min \left\{ \underbrace{\tau_{rem} - \frac{l_{max} - x_j}{q}}_{\text{remained time after end of stopping}}, \frac{x_j}{s} \right\} \quad (A2)$$

in which $\frac{x_j}{s}$ is the time it takes for the vehicle to reach the intersection. If the remaining time after the end of the stopping time is greater than $\frac{x_j}{s}$, the vehicle continues its travel at the free flow speed in the next link and does not experience any more delay.

When a vehicle travels at speed s for time τ_Q , it experiences a delay compared with a vehicle traveling at the free flow speed. In this situation, the delay caused by traveling in the queue is $\tau_Q - \frac{s\tau_Q}{v_{ff}}$, which is the difference between τ_Q and the time a hypothetical vehicle spends at free flow speed v_{ff} to traverse the same distance.

According to the explanations, the delay w is determined as:

$$w = \begin{cases} 0 & \text{if } \tau_{rem} \leq 0 \\ \tau_{rem} & \text{if } 0 < \tau_{rem} \leq \frac{l_{\max} - x_j}{q} \\ \frac{l_{\max} - x_j}{q} + \left(\tau_Q - \frac{s\tau_q}{v_{ff}} \right) & \text{if } \tau_{rem} > \frac{l_{\max} - x_j}{q} \end{cases} \quad (A3)$$

In Equation (A3), w is a function of x_j . The location in which the vehicle joins the queue (x_j) is a random variable with a determined PDF. We determine x_j in terms of w so that we can model the PDF of w based on the PDF of x_j . To have x_j in terms of w , Equation (A3) is expanded based on Equation (A2):

$$w = W(x_j) = \begin{cases} 0 & \text{if } x_j \leq x_0 - \tau v_{ff} \\ \tau_{rem} & \text{if } x_0 - \tau v_{ff} < x_j \leq \frac{l_{\max} v_{ff} + qx_0 - \tau q v_{ff}}{q + v_{ff}} \\ \left(\tau_{rem} - \frac{l_{\max} - x_j}{q} \right) \left(1 - \frac{s}{v_{ff}} \right) & \text{if } \frac{l_{\max} v_{ff} + qx_0 - \tau q v_{ff}}{q + v_{ff}} < x_j \leq \frac{s(l_{\max} v_{ff} + qx_0 - \tau q v_{ff})}{qs - qv_{ff} + sv_{ff}} \\ \frac{x_j}{s} \left(1 - \frac{s}{v_{ff}} \right) & \text{if } \frac{s(l_{\max} v_{ff} + qx_0 - \tau q v_{ff})}{qs - qv_{ff} + sv_{ff}} < x_j \leq l_{\max} \end{cases} \quad (A4)$$

Figure A1 indicates a general diagram of Equation (A4).

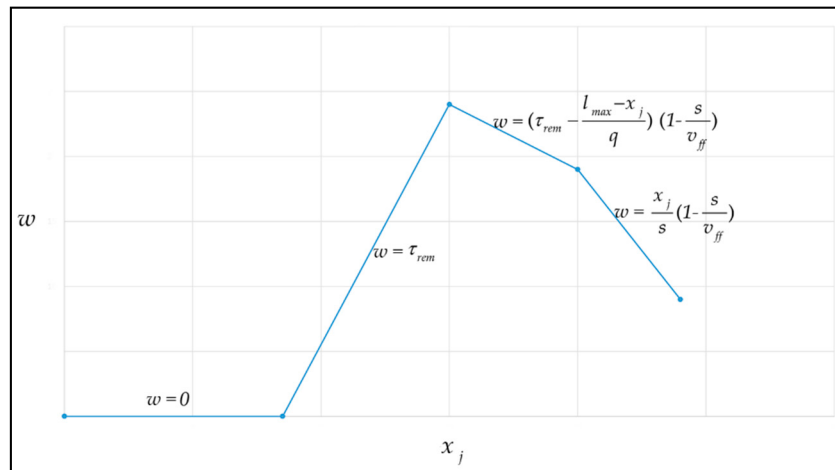


Figure A1. Delay w as a function of x_j in Case 2.

Equation (A4) (Figure A1) indicates w as a hybrid function of x_j . Based on Equation (A4) (Figure A1), we constitute the relation (note that this relation is not generally a function, while w may be non-monotonic) in which x_j is in terms of w . The relation is derived by inverting the sub-functions in Equation (A4):

$$x_j = X_j(w) = \begin{cases} [0, x_0 - \tau v_{ff}] & \text{if } w = 0 \\ x_0 + v_{ff}(w - \tau) & \text{if } 0 < w \leq W(l_{\max}) \\ \left\{ x_0 + v_{ff}(w - \tau), \frac{swv_{ff}}{v_{ff} - s} \right\} & \text{if } W(l_{\max}) < w \leq \frac{(1 - \frac{s}{v_{ff}})(l_{\max} v_{ff} + qx_0 - \tau q v_{ff})}{qs - qv_{ff} + sv_{ff}} \\ \left\{ x_0 + v_{ff}(w - \tau), \frac{l_{\max} v_{ff} + qx_0 - \tau q v_{ff}}{q + v_{ff}} - \frac{qwv_{ff}^2}{(q + v_{ff})(s - v_{ff})} \right\} & \text{if } \frac{(1 - \frac{s}{v_{ff}})(l_{\max} v_{ff} + qx_0 - \tau q v_{ff})}{qs - qv_{ff} + sv_{ff}} < w \leq \frac{l_{\max} - x_0 + \tau v_{ff}}{q + v_{ff}} \end{cases} \quad (A5)$$

Equation (A5) determines x_j for any delay w . Accordingly, the probability that a vehicle does not experience any delay ($w = 0$) during the prediction horizon is equal to the probability that the vehicle joins the queue at a point in $[0, x_0 - \tau v_{ff}]$, which is a mass probability with value $H_{0, x_0 - \tau v_{ff}}^g$. We model this probability with a Dirac delta function (centered at 0) and its corresponding weight $H_{0, x_0 - \tau v_{ff}}^g$. The probability distribution of other defined values of w is determined by the PDF of joining

a growing queue at x_j (Equation (8)). The probability of undefined values of w is 0. The PDF of delay in Case 2 reads:

$$g^{in}(w) = \begin{cases} \delta(w)H_{0,x_0-\tau v_{ff}}^g & \text{if } w = 0 \\ \eta^g(X_j(w)) & \text{if } 0 < w \leq W(l_{\max}) \\ \sum \eta^g(x_j), x_j \in X_j(w) & \text{if } W(l_{\max}) < w \leq \frac{l_{\max} - x_0 + \tau v_{ff}}{q + v_{ff}} \\ 0 & \text{if } w > \frac{l_{\max} - x_0 + \tau v_{ff}}{q + v_{ff}} \end{cases} \quad (A6)$$

Accordingly, the CDF of the delay in Case 2 reads:

$$G^{in}(w) = \begin{cases} H_{0,x_0-\tau v_{ff}}^g & \text{if } w = 0 \\ H_{0,X_j(w)}^g & \text{if } 0 < w < W(l_{\max}) \\ H_{0,\min\{X_j(w)\}}^g + H_{\max\{X_j(w)\},l_{\max}}^g & \text{if } W(l_{\max}) \leq w \leq \frac{l_{\max} - x_0 + \tau v_{ff}}{q + v_{ff}} \\ H_{0,l_{\max}}^g & \text{if } \frac{l_{\max} - x_0 + \tau v_{ff}}{q + v_{ff}} \leq w \end{cases} \quad (A7)$$

Appendix A.2. The Vehicle Does Not Join the Queue (Case 3)

In this case, a vehicle does not experience any delay ($w = 0$) at the intersection (even after the prediction horizon). It happens if vehicles are not in a queue at t_0 (Cases 4 or 5) and do not join the queue at the intersection (Cases 1 or 2). Accordingly, the probability that a vehicle does not experience any delay at the intersection is a mass probability with value:

$$1 - \underbrace{H_{x_0,l_{\max}}^r}_{\substack{\text{The probability that} \\ \text{the vehicle is in a} \\ \text{removing queue}}} - \underbrace{H_{x_0,l_{\max}}^g}_{\substack{\text{The probability that} \\ \text{the vehicle is in a} \\ \text{growing queue}}} - \underbrace{H_{0,x_0}^r}_{\substack{\text{The probability that} \\ \text{the vehicle joins a} \\ \text{removing queue}}} - \underbrace{H_{0,x_0}^g}_{\substack{\text{The probability that} \\ \text{the vehicle joins a} \\ \text{growing queue}}} \quad (A8)$$

We model this case with a Dirac delta function (centered at 0) and a weight equal to Equation (A8). The PDF of delay in Case 3 denoted as $o(w)$ reads:

$$o(w) = \delta(w)(1 - H_{x_0,l_{\max}}^r - H_{x_0,l_{\max}}^g - H_{0,x_0}^r - H_{0,x_0}^g) \quad (A9)$$

The CDF of the delay in Case 3 reads:

$$O(w) = \begin{cases} 0 & \text{if } w < 0 \\ \left(1 - H_{x_0,l_{\max}}^r - H_{x_0,l_{\max}}^g - H_{0,x_0}^r - H_{0,x_0}^g\right) & \text{if } w \geq 0 \end{cases} \quad (A10)$$

Appendix A.3. The Vehicle Is in the Queue at t_0 (Case 4 and Case 5)

If $x_0 > l_{\max}$, this case never happens. Thus, the probability distribution of the delay is determined only in the case $x_0 < l_{\max}$. The probability that a vehicle is in the queue at (x_0, t_0) is equal to the probability that the queue is removing and its length is between x_0 and $l_{\max}$.

In a removing queue (Case 4) the delay w reads:

$$w = \tau_Q \left(1 - \frac{s}{v_{ff}}\right) \quad (A11)$$

in which $\tau_Q = \min\left\{\tau, \frac{x_0}{q}\right\}$.

As indicated by Equation (A11), in this case, the delay is a constant independent value. The PDF of Case 4 is a mass probability with a value $w = \tau_Q(1 - \frac{s}{v_{ff}})$, and its weight is equal to the probability that the queue is removing, and its length is between x_0 and $l_{\max}$. We model this case with a Dirac delta function (centered at $w = \tau_Q(1 - \frac{s}{v_{ff}})$) and the weight $H_{x_0, l_{\max}}^r$. The PDF of the delay in Case 4 reads:

$$r^{in}(w) = \delta(w - (\tau_Q(1 - \frac{s}{v_{ff}})))H_{x_0, l_{\max}}^r \quad (A12)$$

Accordingly, the CDF of the delay in Case 4 reads:

$$R^{in}(w) = \begin{cases} 0 & \text{if } w < \tau_Q(1 - \frac{s}{v_{ff}}) \\ \tau_Q(1 - \frac{s}{v_{ff}})H_{x_0, l_{\max}}^r & \text{if } \tau_Q(1 - \frac{s}{v_{ff}}) \leq w \end{cases} \quad (A13)$$

Case 5 is somewhat similar to the Case 2, but the delay is dependent on the queue length. In a growing queue (Case 5), the delay is determined as a function of the queue length l . In this case, the delay w reads:

$$w = W(l) = \begin{cases} \tau & \text{if } \tau \leq \frac{l_{\max} - l}{q} \\ \frac{l_{\max} - l}{q} + \tau_Q(1 - \frac{s}{v_{ff}}) & \text{if } \tau > \frac{l_{\max} - l}{q} \end{cases} \quad (A14)$$

where $\tau_Q = \min\{\tau - \frac{l_{\max} - l}{q}, \frac{x_0}{q}\}$.

Then w is expanded by τ_Q :

$$w = W(l) = \begin{cases} \tau & \text{if } l \leq \tau q - l_{\max} \\ \tau - \frac{s}{v_{ff}}(\tau + \frac{l - l_{\max}}{q}) & \text{if } \tau q - l_{\max} < l \leq l_{\max} + x_0 - \tau q \\ \frac{l_{\max} - l}{q} + \frac{x_0}{q}(1 - \frac{s}{v_{ff}}) & \text{if } l_{\max} + x_0 - \tau q < l \leq l_{\max} \end{cases} \quad (A15)$$

Using Equation (A15), the queue length l is constituted as a function of w . The queue length is:

$$l = L(w) = \begin{cases} [x_0, l_{\max} - \tau q] & \text{if } w = \tau \\ l_{\max} - \tau q + \frac{qv_{ff}(\tau - w)}{s} & \text{if } \tau - \frac{sx_0}{qv_{ff}} \leq w < \tau \\ l_{\max} + x_0 - qw - \frac{sx_0}{v_{ff}} & \text{if } \frac{x_0}{q}(1 - \frac{s}{v_{ff}}) \leq w < \tau - \frac{sx_0}{qv_{ff}} \end{cases} \quad (A16)$$

According to Equation (A16), the probability that a vehicle experiences delay $w = \tau$ during the prediction horizon is equal to the probability that the queue length is $l \in [x_0, l_{\max} - \tau q]$, which is a mass probability with value $H_{x_0, l_{\max} - \tau q}^g$. We model this probability with a Dirac delta function (centered at τ) and its corresponding weight $H_{x_0, l_{\max} - \tau q}^g$. The probability distribution of other defined values of w is determined by the PDF of the length of the growing queue $\eta^g(l)$ (Equation (8)). The probability of undefined values of w is 0. The PDF of delay in Case 5 reads:

$$g^{in}(w) = \begin{cases} \delta(w - \tau)H_{x_0, l_{\max} - \tau q}^g & \text{if } w = \tau \\ \eta^g(L(w)) & \text{if } \frac{x_0}{q}(1 - \frac{s}{v_{ff}}) \leq w < \tau \\ 0 & \text{if } w < \frac{x_0}{q}(1 - \frac{s}{v_{ff}}) \end{cases} \quad (A17)$$

Therefore, the CDF of delay in Case 5 reads:

$$G^{\text{in}}(w) = \begin{cases} 0 & \text{if } w < \frac{x_0}{q} \left(1 - \frac{s}{v_{ff}}\right) \\ H_{L(w), L(\frac{x_0}{q}(1 - \frac{s}{v_{ff}}))}^g & \text{if } \frac{x_0}{q} \left(1 - \frac{s}{v_{ff}}\right) \leq w < \tau \\ H_{x_0, L(\frac{x_0}{q}(1 - \frac{s}{v_{ff}}))}^g & \text{if } \tau \leq w \end{cases} \quad (\text{A18})$$

References

- Jensen, C.S.; Friis-Christensen, A.; Pedersen, T.B.; Pfoer, D.; Saltenis, S.; Tryfona, N. Location-based services: A database perspective. *ScanGIS* **2001**, 59–68. Available online: http://www.dieter.pfoer.org/publications/jensen_lbs01.pdf (accessed on 20 September 2017).
- Trivedi, M.M.; Gandhi, T.; McCall, J. Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety. *IEEE Trans. Int. Transp. Syst.* **2007**, *8*, 108–120. [CrossRef]
- Mo, Y.; Yu, D.; Song, J.; Zheng, K.; Guo, Y. Vehicle position updating strategy based on Kalman filter prediction in VANET environment. *Discret. Dyn. Nat. Soc.* **2016**, 2016. [CrossRef]
- Yuan, G.; Sun, P.; Zhao, J.; Li, D.; Wang, C. A review of moving object trajectory clustering algorithms. *Artif. Intell. Rev.* **2017**, *47*, 123–144. [CrossRef]
- Ying, J.J.-C.; Lee, W.-C.; Weng, T.-C.; Tseng, V.S. Semantic trajectory mining for location prediction. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, IL, USA, 1–4 November 2011; ACM: New York, NY, USA, 2011; pp. 34–43.
- Jeung, H.; Liu, Q.; Shen, H.T.; Zhou, X. A hybrid prediction model for moving objects. In Proceedings of the 2008 IEEE 24th International Conference on Data Engineering, ICDE 2008, Cancun, Mexico, 7–12 April 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 70–79.
- Hofleitner, A.; Herring, R.; Abbeel, P.; Bayen, A. Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 1679–1693. [CrossRef]
- Civilis, A.; Jensen, C.S.; Pakalnis, S. Techniques for efficient road-network-based tracking of moving objects. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 698–712. [CrossRef]
- Reza, T.A.; Barbeau, M.; Alsubaihi, B. Tracking an on the run vehicle in a metropolitan VANET. In Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast, QLD, Australia, 23–26 June 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 220–227.
- Bhuiyan, M.Z.A.; Wang, G.; Vasilakos, A.V. Local area prediction-based mobile target tracking in wireless sensor networks. *IEEE Trans. Comput.* **2015**, *64*, 1968–1982. [CrossRef]
- Li, Y.; Wang, J.; Zhang, L. LBS-based dilemma zone warning system at signalized intersection. In *Progress in Location-Based Services 2014*; Springer: Berlin, Germany, 2015; pp. 223–237.
- Pham, N.T.; Leman, K.; Chang, R.; Zhang, J.; Wang, H.L. Fusing appearance and spatio-temporal features for multiple camera tracking. In Proceedings of the International Conference on Multimedia Modeling, Dublin, Ireland, 6–10 January 2014; Springer: Berlin, Germany, 2014; pp. 365–374.
- Van Zuylen, H.J.; Zheng, F.; Chen, Y. Using probe vehicle data for traffic state estimation in signalized urban networks. In *Traffic Data Collection and Its Standardization*; Springer: Berlin, Germany, 2010; pp. 109–127.
- Uno, N.; Kurauchi, F.; Tamura, H.; Iida, Y. Using bus probe data for analysis of travel time variability. *J. Intell. Transp. Syst.* **2009**, *13*, 2–15. [CrossRef]
- Bracciale, L.; Bonola, M.; Loreti, P.; Bianchi, G.; Amici, R.; Rabuffi, A. CRAWDAD Dataset Rome/Taxi (v. 2014-07-17). 2014. Available online: <https://crawdad.org/roma/taxi/20140717/> (accessed on 5 January 2017).
- Cho, S.-B. Exploiting machine learning techniques for location recognition and prediction with smartphone logs. *Neurocomputing* **2016**, *176*, 98–106. [CrossRef]
- Hofleitner, A.; Herring, R.; Bayen, A. Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning. *Transp. Res. Part B Methodol.* **2012**, *46*, 1097–1122. [CrossRef]
- Davis, P.J.; Rabinowitz, P. *Methods of Numerical Integration*; Courier Corporation: Chelmsford, MA, USA, 2007.
- Zheng, F.; van Zuylen, H. Reconstruction of delay distribution at signalized intersections based on traffic measurements. In Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC), Funchal, Portugal, 19–22 September 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 1819–1824.

20. Li, C.; Gu, Y.; Qi, J.; Yu, G.; Zhang, R.; Yi, W. Processing moving k NN queries using influential neighbor sets. *Proc. VLDB Endow.* **2014**, *8*, 113–124. [[CrossRef](#)]
21. Šidlauskas, D.; Šaltenis, S.; Christiansen, C.W.; Johansen, J.M.; Šaulys, D. Trees or grids?: Indexing moving objects in main memory. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 4–6 November 2009; ACM: New York, NY, USA, 2009; pp. 236–245.
22. Mazloumi, E.; Rose, G.; Currie, G.; Sarvi, M. An integrated framework to predict bus travel time and its variability using traffic flow data. *J. Intell. Transp. Syst.* **2011**, *15*, 75–90. [[CrossRef](#)]
23. Ding, Z.; Yang, B.; Güting, R.H.; Li, Y. Network-matched trajectory-based moving-object database: Models and applications. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 1918–1928. [[CrossRef](#)]
24. Zheng, Y. Trajectory data mining: An overview. *ACM Trans. Intell. Syst. Technol.* **2015**, *6*, 29. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).