

Stochastic Model and Connectivity Dynamics for VANETs in Signalized Road Systems

Ivan Wang-Hei Ho, *Member, IEEE*, Kin K. Leung, *Fellow, IEEE*, and John W. Polak

Abstract—The space and time dynamics of moving vehicles regulated by traffic signals governs the node connectivity and communication capability of vehicular ad hoc networks (VANETs) in urban environments. However, none of the previous studies on node connectivity has considered such dynamics with the presence of traffic lights and vehicle interactions. In fact, most of them assume that vehicles are distributed homogeneously throughout the geographic area, which is unrealistic. We introduce in this paper a stochastic traffic model for VANETs in signalized urban road systems. The proposed model is a composite of the fluid model and stochastic model. The former characterizes the general flow and evolution of the traffic stream so that the average density of vehicles is readily computable, while the latter takes into account the random behavior of individual vehicles. As the key contribution of this paper, we attempt to approximate vehicle interactions and capture platoon formations and dissipations at traffic signals through a density-dependent velocity profile. The stochastic traffic model with approximation of vehicle interactions is evaluated with extensive simulations, and the distributional result of the model is validated against real-world empirical data in London. In general, we show that the fluid model can adequately describe the mean behavior of the traffic stream, while the stochastic model can approximate the probability distribution well even when vehicles interact with each other as their movement is controlled by traffic lights. With the knowledge of the mean vehicular density dynamics and its probability distribution from the stochastic traffic model, we determine the degree of connectivity in the communication network and illustrate that system engineering and planning for optimizing both the transport (in terms of congestion) and communication networks (in terms of connectivity) can be carried out with the proposed model.

Index Terms—Connectivity, signalized road system, stochastic traffic model, vehicle interaction, vehicular ad hoc network (VANET).

I. INTRODUCTION

IN a vehicular ad hoc network (VANET), vehicles communicate with each other and possibly with roadside infrastructure nodes. Node connectivity and the amount of data that can be exchanged are limited by the duration and quality of the communication links established among nodes, which are determined by the space and time dynamics of moving vehicles.

There are a number of studies on node connectivity in mobile ad hoc networks (MANETs). For instance, [1] shows that

Manuscript received March 02, 2009; revised November 17, 2009 and April 26, 2010; accepted June 21, 2010; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor E. Modiano. Date of publication August 09, 2010; date of current version February 18, 2011.

I. W.-H. Ho and K. K. Leung are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, U.K. (email: wh.ho@imperial.ac.uk; kin.leung@imperial.ac.uk).

J. W. Polak is with the Centre for Transport Studies, Imperial College London, London SW7 2AZ, U.K. (email: j.polak@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNET.2010.2057257

if the radio transmission range of n nodes that are placed uniformly and independently in a disc of unit area is set to $r_0 = [(\log n + c(n))/n\pi]^{1/2}$, the resulting wireless multihop network is asymptotically connected with probability one if and only if $c(n) \rightarrow \infty$. [2] investigates the radio range assignment problem and obtains bounds for the probability that a node is isolated and the network is connected on a one-dimensional (1-D) line. On the other hand, [3] examines the node density threshold for achieving full connectivity in both 1-D and two-dimensional (2-D) ad hoc network. [4] and [5] study the relation between the minimum node degree and k -connectivity in a random graph and explore the minimum radio transmission range r_0 for achieving a fully connected ad hoc network for a given node density.

However, most of the existing studies assume that nodes are uniformly random distributed in an area and they are either stationary or move according to the random waypoint model [5], which are obviously inadequate to capture the spatial distribution of vehicles and their movements. In fact, vehicle movements, particularly in urban environments, are restricted by the road topologies, buildings, etc., and affected by traffic density, which is determined by road capacity, traffic control, and driver behaviors. There are also recent works that aim to model connectivity of vehicles on a one-dimensional highway. [6] assumes the space headway between vehicles is exponentially distributed and introduces a robustness factor to capture the effect of disturbance on VANET connectivity. [7] assumes a continuous-time mobility model where movement of each node is a function of time consisting of a sequence of random intervals that is exponential distributed, and during each interval, a node moves at a constant speed that is independently chosen from a normal distribution. Based on this synthetic mobility model, the author derived the mean cluster size and the probability that the nodes will form a single cluster under the assumption that vehicles arrive according to a Poisson process. These previous studies, however, lack a realistic traffic model to adequately capture node density and its influence on vehicle speed, which are significant in determining connectivity, especially in urban road systems where strong interactions among neighboring vehicles exist.

In light of the inadequacy of existing studies, developing mathematical models to adequately capture the spatial and temporal details, vehicle movement, link condition, and node connectivity is of paramount importance. To capture such dynamics, we propose to investigate a stochastic traffic model for VANETs particularly in signalized urban road systems. The stochastic traffic model is a composite of the deterministic fluid dynamic model and the stochastic model. The densities of the mean number of vehicles and the mean flow rate of the traffic stream are described by or readily computable from the conservation equations of fluid dynamics, which is in the form of partial differential equations (PDEs), while the stochastic

model captures the randomness of individual vehicle and provides the probability distribution.

When we consider vehicle arrivals as a Poisson process, the stochastic model becomes a special version of the Poisson Arrival Location Model (PALM) introduced in [8]. The actual number of vehicles in a given road section at a certain time instance thus has Poisson distribution according to previous PALM results in [8] and [9] given that the arrival of vehicles follow a nonhomogeneous Poisson process. PALM provides a useful framework for representing both the time-dependent behavior and node mobility in wireless communication networks. However, the general PALM introduced in [8] is quite abstract. In view of this, the Markovian highway PALM was presented in [9] and [10] to model the call density and handoff rate of communicating mobiles on a highway.

According to [8], and [9], the Poisson distributional conclusion of PALM is established based on the assumption that different customers do not interact with each other, and the locations of all cars are thus mutually independent, conditional on their arrival times. Therefore, the prior work [9], [10] does not directly model the slowdown decision of a car based on the actions of other cars. Instead, they model it *indirectly* by having all cars slow down in a specific region at a specific time through a velocity profile. The velocity profile is considered and demonstrated in the highway PALM [9], [10] as constant or deterministic velocity fields, neglecting the fundamental interdependence between velocity and vehicular density, as this could ruin the Poisson property of PALM when interactions among vehicles are introduced.

In this paper, to study road conditions more closely, we attempt to capture vehicle interactions *directly* through a *density-dependent velocity profile* and bring the PALM from highway to *generic urban route* by introducing an important transport factor, *traffic signals*. By considering the velocity field as a function of preceding vehicular density, the velocity profile is no longer deterministic; it is constructed iteratively in an automatic manner together with the evolving traffic stream. We evaluate the stochastic traffic model with such approximation in terms of the stochastic independence of the stochastic model as well as the expected value computed from the fluid conservation equations through extensive simulations. The Poisson distributional result of the model is also validated against empirical loop detector data in Central London. In essence, we show that the dependency inherited from vehicular interactions and traffic signals tends to be ignored by the system, and the Poisson property of PALM still holds under such approximation of vehicle interactions. In addition, the solutions from the fluid model provide reasonable estimates for the mean number of vehicles in a given region, and the formation and dissipation of vehicular platoons at traffic signals or stops are well accommodated by the density-dependent velocity profile given the signaling inputs.

With the knowledge (in terms of expected value and distribution) of the vehicular density as a function of space and time from the proposed model, traffic signal management can be carried out and communication connectivity of vehicles along urban routes can be further deduced, which is crucial for a number of network management and protocol design decisions. For instance, we determine the probability that the entire communication network in an urban route segment is connected in a multihop manner, and the problem is further investigated for a

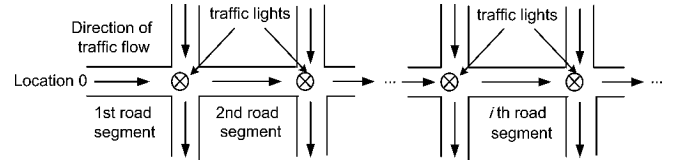


Fig. 1. Road configuration considered in this paper.

general case of a k -connected network. To illustrate the model applicability, we show how vehicular density can be affected by the timing synchronization of traffic signals and how the model can be used to identify the operating point for minimal transport (vehicle) congestion. Furthermore, our model can be applied to identify regions with poor connectivity at such an operating point for the placement of roadside communication nodes or adjustment of transmission ranges of vehicles to boost the overall communication connectivity.

The introduction of vehicle interactions and traffic signals and the model validations are regarded as the key contributions of the paper since none of the previous work has addressed the problem of modeling connectivity in a signalized urban road network with practical vehicle distribution. Our work here provides the methodology and its validations so that node connectivity in practical VANETs can be evaluated in the future with sufficient confidence. The proposed model can also serve as a fundamental building block for constructing more elaborated urban traffic models.

The rest of the paper is organized as follows. Section II presents the set of conservation equations that establish the fluid model. Section III introduces the stochastic model and the distributional results that can be obtained if vehicles' arrivals appear to be a Poisson process. In Section IV, we present an algorithm to approximate vehicle interactions and capture platoon formations and dispersions at traffic signals with a density-dependent velocity profile. We evaluate the stochastic traffic model with approximation of vehicular interactions and traffic signals through numerical results in Section V, and validate the Poisson distributional result of the model against empirical data in Section VI. Analysis of VANET connectivity based on the vehicular density dynamics obtained from the stochastic traffic model is discussed in Section VII. Finally, Section VIII concludes the paper.

II. DETERMINISTIC FLUID DYNAMIC MODEL

In this paper, we consider traffic in a one-way, single-lane, semi-infinite signalized urban road (or route) as shown in Fig. 1. Although the road is fed with traffic from adjacent streets, the one-way road under consideration is the one running from the left to the right in the figure. More complicated road topology can be represented by superposing multiple versions of urban routes. Let our location space be the interval $[0, \infty)$, the boundary point 0 is the spatial origin, and it marks the starting point of the road. The arrival process $\{A(t) - \infty < t < \infty\}$ counts the number of arrivals to the first segment of the route up to time t , which we assume is finite with probability 1, and is characterized by a nonnegative and integrable external arrival rate function $\alpha(t)$.

Furthermore, the route consists of a number of road segments indexed by $i = 1, 2, \dots$, and traffic lights are located at the junctions of road segments, where vehicles can leave and join

the route. We let the location of the i th junction (or traffic light) between road segments i and $i + 1$ be x_i .

The fluid dynamic model is a kind of continuum traffic flow model, which reduces laws of traffic to a PDE that may be studied as elegantly and simply as other physical phenomena that are also governed by PDEs. The reader is referred to [11] for other common continuum models in transport studies, such as the Lighthill–Whitham–Richards (LWR) model.

The major difference between our fluid dynamic model and other continuum models is that we model vehicle motions with a velocity profile, vehicles at location x and time t move forward the route according to a velocity field $v(x, t)$, and it can be deterministic or density-dependent. Vehicles stop at road junctions for a red signal, which can be reflected and modeled by the velocity profile. However, the continuum model alone is unable to capture traffic instability and the randomness of individual vehicle. Therefore, we couple the fluid model with the stochastic model in the next section as a remedy.

To begin, we will describe the fluid dynamic conservation equations and corresponding notations that hold for the general systems. Let $N(x, t)$ be the number of vehicles in location $(0, x]$ at time t , and $n(x, t)$ be the density of vehicles in location $(0, x]$ at time t . Thus

$$n(x, t) = \frac{\partial N(x, t)}{\partial x}. \quad (1)$$

Let $Q(x, t)$ be the number of vehicles moving past position x before time t . Then, the flow rate $q(x, t)$ is defined by

$$q(x, t) = \frac{\partial Q(x, t)}{\partial t}. \quad (2)$$

Let $C^+(x, t)$ and $C^-(x, t)$ be the number of vehicles arriving to and departing from the route in location $(0, x]$ during time interval $(-\infty, t]$, respectively. Then, the associated rate densities are respectively

$$c^+(x, t) = \frac{\partial^2 C^+(x, t)}{\partial x \partial t}$$

and

$$c^-(x, t) = \frac{\partial^2 C^-(x, t)}{\partial x \partial t}. \quad (3)$$

Assuming all traffic moves only from left to right down the positive real line, then the four variables N , Q , C^+ , C^- satisfy the following conservation relation:

$$C^+(x, t) = N(x, t) + Q(x, t) + C^-(x, t). \quad (4)$$

By applying the operator $\partial^2/(\partial x \partial t)$ to (4), we have the partial differential equation

$$\frac{\partial n(x, t)}{\partial t} + \frac{\partial q(x, t)}{\partial x} = c^+(x, t) - c^-(x, t). \quad (5)$$

According to traffic flow theory [11], there is a fundamental relationship that flow = density * velocity, i.e., the rate that vehicles pass position x at time t is equal to the density multiplied by the velocity of vehicles in position x at time t .

$$q(x, t) = n(x, t)v(x, t). \quad (6)$$

By substituting (6) into (5), we have

$$\frac{\partial n(x, t)}{\partial t} + \frac{\partial [n(x, t)v(x, t)]}{\partial x} = c^+(x, t) - c^-(x, t). \quad (7)$$

The resulting partial differential (7) is a one-dimensional version of the generalized conservation law for fluid motion [12]. This equation governs the mean behavior of any stochastic traffic model. For the ease of solving (7), we introduce an additional assumption to convert the PDE to an ordinary differential equation (ODE). Let the location x as a time function $x(t)$ be given by

$$\frac{dx(t)}{dt} = v(x(t), t). \quad (8)$$

By chain rule, $\frac{dn(x(t), t)}{dt} = \frac{\partial n(x, t)}{\partial x} \cdot \frac{dx(t)}{dt} + \frac{\partial n(x, t)}{\partial t}$, and by substituting (7) and (8) into this equation, we have

$$\frac{dn(x(t), t)}{dt} = c^+(x(t), t) - c^-(x(t), t) - \frac{\partial v(x, t)}{\partial x} n(x(t), t). \quad (9)$$

Due to the partial derivative of $v(x, t)$ w.r.t. x in (9), it is ODE if and only if $v(x, t)$ is not a function of $n(x, t)$. Note that though we introduce the use of a density-dependent velocity profile later on, we still solve the ODEs for the mean number of vehicles as an approximation in Section V.

Throughout this paper, we consider that vehicles arrive at the first route segment according to an external arrival rate function $\alpha(t)$ and keep moving along the route without vehicles joining or leaving at junctions (e.g., buses move along a bus route). In this case, $c^+(x, t) = \alpha(t)\delta(x)$ and $c^-(x, t) = 0$, where $\lim_{\varepsilon \rightarrow 0} \int_{x-\varepsilon}^{x+\varepsilon} \delta(y)dy = 1$ if $x = 0$, and 0 otherwise.

III. STOCHASTIC MODEL

In contrast to the deterministic fluid dynamic model, the stochastic model captures the stochastic fluctuations of the quantities of interest. When the two models are coupled with each other to form the stochastic traffic model, the solutions from the PDEs or ODEs describe the expected number of vehicles, and the actual number of vehicles is captured by the additional distribution information from the stochastic model.

From now on, the densities $n(x, t)$ and $q(x, t)$ are defined as the partial derivatives of expected values, that is,

$$n(x, t) = \frac{\partial E[N(x, t)]}{\partial x}$$

and

$$q(x, t) = \frac{\partial E[Q(x, t)]}{\partial t}.$$

Similarly, the rate densities $c^+(x, t)$ and $c^-(x, t)$ are the second partial derivatives of expected values, that is,

$$c^+(x, t) = \frac{\partial^2 E[C^+(x, t)]}{\partial x \partial t}$$

and

$$c^-(x, t) = \frac{\partial^2 E[C^-(x, t)]}{\partial x \partial t}.$$

The general stochastic model can be of any distributions depending on the arrival process of vehicles, and the equations in the fluid model continue to hold regardless the distribution of the stochastic model. In this paper, we specifically consider PALM. Again, the fluid dynamic model is not dependent on the Poisson assumption; it holds as long as the arrival process A is an arbitrary point process with time-dependent arrival-rate function α .

With PALM, the arrival process $\{A(t) | -\infty < t < \infty\}$ for vehicles to arrive at the first road segment of the route is a *non-homogeneous Poisson process* with nonnegative and integrable external arrival rate function $\alpha(t)$. That is, the number of arrivals in the interval $(t_1, t_2]$ is Poisson with mean $\int_{t_1}^{t_2} \alpha(s) ds$.

According to [8], [9], we can construct $N(x, t)$, the random number of vehicles within the range $(0, x]$ at time t , via stochastic integration starting with the Poisson process A , where $A(t)$ counts the number of vehicles arriving to the road segment up to time t .

$$\begin{aligned} N(x, t) &= \int_{\sigma(x, t)}^t \mathbf{1}_{\{L_s(t) \in (0, x]\}} dA(s) \\ &= \sum_{n=A(\sigma(x, t))}^{A(t)} \mathbf{1}_{\{L_{\hat{A}_n}(t) \in (0, x]\}} \end{aligned} \quad (10)$$

where \hat{A}_n is the n th jump time of A , counting backward from time t . $\mathbf{1}_B$ is an indicator function such that it returns 1 if B is true, and 0 otherwise. $L_s(t)$ is the location process, which specifies the position of the vehicle on the road segment at time t that arrived at time s . Let $\sigma(x, t)$ denote the route entrance time for a vehicle to be in region $(0, x]$ at time t . For vehicles that arrive to the route before $\sigma(x, t)$, it will be past position x by time t . On the other hand, for vehicles that arrive after $\sigma(x, t)$, it will be still in the region $(0, x]$ at time t . In simpler terms, (10) counts the number of vehicles that is located in region $(0, x]$ at time t . The condition $\{L_s(t) \text{ belongs to } (0, x]\}$ is the location constraint while the integration from time $\sigma(x, t)$ to t gives the time constraint.

Similarly, we can express $Q(x, t)$, the number of vehicles passed through position x at time t , in terms of Poisson integration as

$$Q(x, t) = \int_{-\infty}^{\sigma(x, t)} \mathbf{1} dA(s) = \sum_{n=1}^{A(\sigma(x, t))} \mathbf{1} = A(\sigma(x, t)). \quad (11)$$

Note that the stochastic integral in (10) and (11) can be treated as a countable sum.

The highway PALM in [9] and [10] did not consider the stop-and-go motions of vehicles in urban routes. We are now on the way to prove that the introduction of traffic signals and stops will not affect the Poisson property of the stochastic model as long as we model vehicle's stop-and-go motions through a deterministic velocity profile as a function of space and time, i.e., we model the stop-and-go behaviors of cars at traffic lights by having all cars slow down, stop, and then go in a specific region (e.g., at road junctions where traffic lights are located) at a specific time (e.g., during the red signal periods).

Proposition 1: For a nonhomogeneous Poisson arrival process with integrable external arrival rate function $\alpha(t)$, if we consider traffic signals via modeling a deterministic velocity profile $v(x, t)$ that is independent of other parameters such as vehicular density, the following results hold.

a) For all real t , $\{N(x, t) | x \geq 0\}$ is a Poisson process with

$$E[N(x, t)] = \int_{\sigma(x, t)}^t \alpha(s) ds. \quad (12)$$

b) For all nonnegative x , $\{Q(x, t) | -\infty < t < \infty\}$ is a Poisson process with

$$E[Q(x, t)] = \int_{-\infty}^{\sigma(x, t)} \alpha(s) ds. \quad (13)$$

Proof: When we capture traffic signals via a deterministic velocity profile function $v(x, t)$, we model vehicle stopping by setting $v(x_i, t_i) = 0$, for all $x_i \in J$, and $t_i \in T_r$, where J is the set of locations of traffic lights, and T_r is the set of red signal periods. In this case, there are still no interactions between vehicles, and the locations of all of the cars are still mutually independent, conditional to their arrival time.

Therefore, the proof is regarded as a consequence of [8, Theorem 2.1]. Given the properties of Poisson integration, it is clear that each $N(x, t)$ has a Poisson distribution and, as a function of x , these random variables have the independent increment property. By the same token, we can see in (11) that $Q(x, t) = A(\sigma)$, which is a Poisson process as a function of t .

Both integrations use $\sigma(x, t)$ because by definition any arrival after this time must still be in the region $(0, x]$ at time t , and any arrival before this time must be past position x by time t given the deterministic velocity profile. Taking expectations of these Poisson integrals gives us the formulas for the expectations of $N(x, t)$ and $Q(x, t)$ as in (12) and (13), respectively. \square

As a result, as long as we model traffic signals through a deterministic velocity field as a function of space and time, and all vehicles do not interact with each other, the Poisson distributional conclusion in [8] and [9] remains valid. Furthermore, It is significant that the mean formulas in Proposition 1 do not depend on the arrival process A being a Poisson process. Of course, the Poisson process conclusions do depend on A being Poisson.

In addition, as a property of the stochastic model, the number of vehicles in two nonoverlapping road regions are independent. Its proof is as follows.

Lemma 1: For a Poisson arrival process, the number of arrivals $M(t_1, t_2)$ and $M(t_3, t_4)$ in two nonoverlapping time intervals $(t_1, t_2]$ and $(t_3, t_4]$, where $t_1 \leq t_2 \leq t_3 \leq t_4$, are independent.

Proof: It is proved by the definition of Poisson process. \square

Proposition 2: For a nonhomogeneous Poisson arrival process $A(t)$, the number of vehicles in two nonoverlapping road regions $N(x_1, x_2, t)$ and $N(x_3, x_4, t)$ (where $x_1 \leq x_2 \leq x_3 \leq x_4$) at time t are independent if we model vehicle motions through a deterministic velocity profile as a function of space and time that is independent of other parameters such as vehicle density.

Proof: According to the definition of the PALM, we have

$$\begin{aligned} N(x_1, x_2, t) &= N(x_2, t) - N(x_1, t) \\ &= \int_{\sigma(x_2, t)}^{\sigma(x_1, t)} \mathbf{1}_{\{L_s(t) \in (x_1, x_2]\}} dA(s). \end{aligned} \quad (14)$$

Therefore, vehicles that arrive in time interval $(\sigma(x_2, t), \sigma(x_1, t)]$ will be in region $(x_1, x_2]$ at time t [i.e., $N(x_1, x_2, t) = M(\sigma(x_2, t), \sigma(x_1, t))$]. Similarly, vehicles that arrive in time interval $(\sigma(x_4, t), \sigma(x_3, t)]$ will be in region $(x_3, x_4]$ at time t . Since $x_1 \leq x_2 \leq x_3 \leq x_4$ and we assume there are no vehicles overtaking, we have $\sigma(x_4, t) \leq \sigma(x_3, t) \leq \sigma(x_2, t) \leq \sigma(x_1, t)$, where

$(\sigma(x_4, t), \sigma(x_3, t))$ and $(\sigma(x_2, t), \sigma(x_1, t))$ are two nonoverlapping time intervals. Thus, it is proved by Lemma 1 that $N(x_1, x_2, t)$ and $N(x_3, x_4, t)$ are independent. \square

In the next section, we approximate vehicular interactions through a density-dependent velocity profile. In that case, Propositions 1 and 2 in principle are invalid, and the Poisson property no longer holds as we cannot determine if a car arrives at time s will have passed location x at time t with the density-dependent velocity profile. Although we expect that the distribution of the number of vehicles in a given region is no longer ideal Poisson distribution, we are still interested in knowing the deviations incurred in order to quantify the goodness of the stochastic traffic model with the vehicle interactions approximation. Numerical results are provided in Section V for such purpose.

IV. TRAFFIC SIGNALS MODELING AND APPROXIMATION OF VEHICLE INTERACTIONS

A practical way to represent the stop-and-go motion of vehicles at traffic signals is to introduce the concept of *shock wave* to model the platoon formation (compression of vehicles) and dissipation (release of vehicles) at traffic signals. Based on a basic traffic flow model that relates vehicular density to velocity, we present an automatic approach to capture shockwave propagation by approximating vehicle interactions with a density-dependent velocity profile.

The general model connecting speed, flow, and density is proposed by Greenshield [13], in which velocity depends on vehicular density, with velocity decreasing as the density increases as shown in (15). We assume the Greenshield's model here for its simplicity and explicit relationship between velocity and vehicular density, which are respectively the input and output of our system.

$$v(x, t) = v_f \left(1 - \frac{n(x, t)}{k_j} \right) \quad (15)$$

where v_f represents the mean free speed and k_j denotes the jamming density. If we define l_c as the average space occupied by a car at stationary, then $k_j = 1/l_c$.

We can actually capture the propagation of shockwave through the fundamental relationship between vehicular density and velocity. In the following, we propose the front-density-dependent velocity field based on the Greenshield's model [13] for such purpose:

$$v(x, t) = v_f \left(1 - \frac{n(x + \Delta x, t)}{k_j} \right). \quad (16)$$

Equation (16) illustrates our front-density-dependent velocity profile: Instead of having the velocity at certain location x as a function of density at location x , we have it as a function of density Δx in front, and the velocity decreases as the density in front increases. It is analogous to the car-following mechanism [14] in transport studies: When the vehicular density in front becomes high (i.e., cars in front decelerate), we should decelerate as well. Such general traffic flow model is applicable to most of the urban route scenarios, no matter if there are vehicles slowing down, stopping, or starting to move or not.

With the front-density-dependent velocity profile, we can approximate the interactions between vehicles in the fluid dynamic model iteratively. The pseudocode is shown in Algorithm 1. The

Algorithm 1: Iterative Approximation of Vehicle Interactions with the front density-dependent velocity profile

```

1: for ( $t_i = 0; t_i < t_{end}; t_i = t_i + dt$ ) {
2:   if ( $t_i = 0$ )
3:     // initialize the vehicular density profile along the route
4:      $n(x, t_i) = 0$  for all  $x$  belongs to  $X$ ;
5:     // obtain inputs for solving the differential equations in the fluid model
6:     obtain the partial derivative of  $n$  w.r.t.  $x$  from  $n(x, t_i)$ ;
7:     obtain the front-density-dependent velocity field  $v(x, t_i)$  from  $n(x, t_i)$ 
       according to Eq. (16);
8:     // solve the differential equations
9:     by solving the differential equations, we obtain  $n(x, t_i + dt)$  for all  $x$ 
       belong to  $X$ ;
10: }
```

general idea is, initially, we assume there are no vehicles on the road, i.e., $n(x, 0) = 0$ for all x belongs to X , where X is the location space. Therefore, the initial velocity will be the mean free speed v_f according to (16). Based on these initial conditions, we solve the differential equations in the fluid dynamic model for the vehicular density along the route. We can then compute the new velocity profile for the next time slot from the vehicular density profile according to (16), and so on. We perform this algorithm iteratively until we reach some defined end time of the system.

Strictly speaking, by introducing vehicle interactions through the front-density-dependent velocity profile, the stochastic independence or Poisson property of the stochastic traffic model should no longer be valid, but the fluid dynamic model still holds as there is no such interaction restriction in the conservation equations. It will be interesting to explore how good or how poor our model is by incorporating such approximation of vehicle interactions in terms of the independence property. In the following section, we conduct extensive simulations to evaluate the stochastic traffic model.

V. NUMERICAL ANALYSIS

In this section, we present numerical results to illustrate the space and time dynamics captured by the fluid dynamic model and the Poisson property captured by the stochastic model in a signalized urban route. Specifically, we evaluate the quality of the stochastic traffic model with the iterative approximation algorithm in terms of two metrics: 1) the mean values computed from the fluid dynamic model; and 2) the distributional property from the stochastic model.

A. Evaluation of the Stochastic Traffic Model With Approximation of Vehicle Interactions

First of all, let us consider a simple stop-and-go motion example to evaluate 1) and 2) above. In this example, we assume there are no cars joining and leaving the urban route at junctions, and cars only arrive at location 0 at a constant rate (denoted by α) of 30 cars/min, with the mean free speed $v_f = 1$ km/min, using the front-density-dependent velocity profile with $\Delta x = 0.02$ km as described in (16) unless otherwise specified. This complies with the "3-second rule" [15] of safe following distance when vehicle speed is about 20 mi/h, which is the average uncongested speed of vehicles in London [16]. Assume the space occupied by a stationary car as 4 m, so the jamming density $k_j = 250$ cars/km.

Consider that there is a traffic signal at location 2 km on the route, which stops vehicles from time 4 to 4.5 min for 30 s. We allow a further 0.012 km distance behind the traffic signal as the length of the junction. Thus, when $4 \leq t < 4.5$, $v = 0$ if

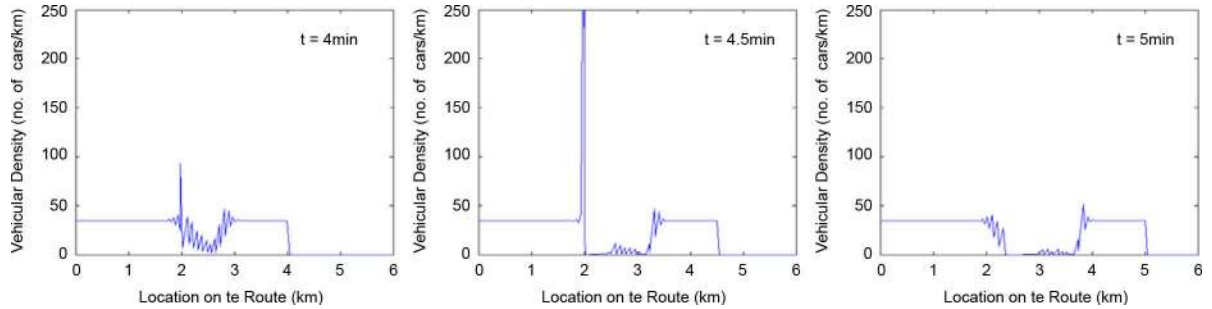


Fig. 2. Platoon formation and dissipation at time instances 4, 4.5, and 5 min as described by the fluid dynamic model.

$2 \leq x < 2.012$ as shown in (17). In addition, we also include extra 0.02 km in front and behind the zero-velocity region so as to ensure that the partial derivative of v w.r.t. x in (9) is finite. For other regions during the stopping period, the front-density-dependent velocity field v applies.

$$v(x, t) = \begin{cases} v, & \text{if } x < 1.98 \\ (v/0.02) * (2 - x), & \text{if } 1.98 \leq x < 2 \\ 0, & \text{if } 2 \leq x < 2.012 \\ (v/0.02) * (x - 2.012), & \text{if } 2.012 \leq x < 2.032 \\ v, & \text{if } x \geq 2.032. \end{cases} \quad (17)$$

When $t \geq 4.5$, we instantaneously set the velocity in the junction $(2, 2.012]$ to be the mean free speed v_f (for 3 s). This is a realistic assumption in a noncongested network since the first car in the queue does not need to follow any cars when the signal turns green, it moves down to the next road segment with the mean free speed until it catches up with preceding vehicles. The following cars move according to the front-density-dependent velocity field.

To model the stopping motion of vehicles more rigorously, we introduce an amber signal period for $3.95 \leq t < 4$, such that the velocity within the junction region $(2, 2.012]$ uniformly decreases to zero during the 3-s amber signal period. Thus, for $t < 3.95$, the system is in equilibrium state.

Fig. 2 depicts the evolution of vehicular density along the signalized route as represented by the fluid dynamic model through solving the ODEs in (8) and (9). We can see that the density pulse starts growing at 4 min due to the amber signal, and the jamming density pulse established and propagated backwards at 4.5 min due to the red signal. Lastly at 5 min, the density pulse was dissipated by the green signal. Our results verify that the modeling of traffic signals and the platoon formation and dissipation process are achievable in the fluid model through the front-density-dependent velocity field plus additional signaling information.

To evaluate the expected values computed from the fluid dynamic model, we compare the solutions of the ODEs in (8) and (9) to the results from simulation. In the simulation, we have the same set of parameters described including the vehicle arrival rate, velocity profile, and signaling inputs. We run the simulation for 5000 times with different random seeds and record the number of vehicles in a given region at a given time. Thus, we have 5000 samples from the simulation for the mean and distribution evaluation. Specifically, we compare the mean number of vehicles in the road region $(1.5, 2]$ at time instances 3.9, 4, 4.2, 4.5, and 5 min. As the solution from the differential equations is the vehicular density instead of the number of vehicles,

TABLE I
COMPARISON OF THE MEAN NUMBER OF VEHICLES IN THE ROAD REGION $(1.5, 2]$ BETWEEN THE ODE'S SOLUTIONS AND SIMULATION RESULTS

Time instance (min)	ODE's (# of cars)	Simulation (# of cars)	Error (%)	D-statistic
3.9	16.3874	15.3554	6.30	0.1581
4	16.9526	16.8446	0.64	0.1527
4.2	24.1911	22.6240	6.48	0.1331
4.5	30.9754	31.2382	0.85	0.1197
5	16.5939	15.3282	7.63	0.1638

we acquire it through integrating the density profile. For example, the mean number of vehicles within the region $(x_1, x_2]$ is $E[N(x_1, x_2, t)] = \int_{x_1}^{x_2} n(x, t) dx$.

Table I compares the mean number of vehicles in the region $(1.5, 2]$ on the urban signalized route at several time instances, from which we can see that the deviations between the ODEs' solutions and simulation results are less than 7.7%, and at time instances 4 and 4.5 min, we can even achieve an error less than 1%. Note that for computation simplicity, we solve the ODEs in (8) and (9) instead of the PDE in (7). However, as we are using a density-dependent velocity profile, $v(x, t)$ is a function of $n(x, t)$, and (9) is no longer an ODE theoretically, but we solve it as an ODE as a way to approximate the mean value.

To evaluate if the distributions of the number of vehicles from simulations can be approximated with a Poisson distribution, we conduct the Kolmogorov–Smirnov test (K–S test) [17] on their cumulative distribution functions (CDFs). In the K–S test, given $F_e(x)$, the empirical cumulative distribution function (ECDF) obtained by simulations, we compare it to $F_h(x)$, the hypothesized CDF. The result of the K–S test is based on the value of the greatest discrepancy between the empirical and hypothesized cumulative distribution, which is called the D-statistic. It is formally defined as

$$D = \max_x |F_e(x) - F_h(x)|. \quad (18)$$

As we aim to approximate the ECDF by certain hypothesized CDF (and not to prove that they are identical), we only examine the D-statistic value. If the value is small, it indicates that the ECDF can be approximated by the hypothesized CDF. In our case, the hypothesized CDF will be the CDF of Poisson distribution with the parameter as the mean of the simulation results.

We treat the case with deterministic velocity profile ($v = 1$ km/min) as the benchmark for this goodness-of-fit study, of which we found its D-statistic is about 0.1. With regard to Table I, the D-statistics of the simulation results with vehicle interactions considered remain small relative to the benchmark (at about 0.14 on average) for all time instances,

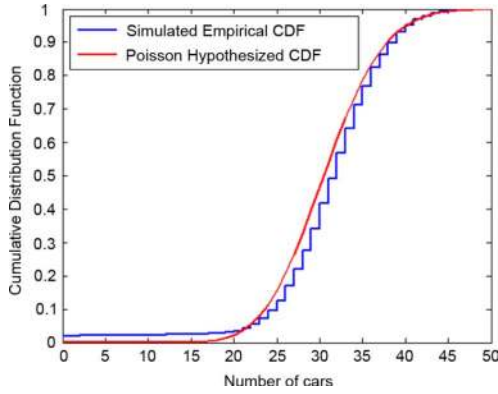


Fig. 3. Empirical and hypothesized CDF at time instance 4.5 min.

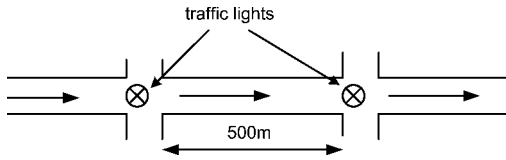


Fig. 4. Street configuration test network.

while Fig. 3 plots the empirical CDF of the number of vehicles within the road region $(1.5, 2]$ at time instance 4.5 min for the 5000 simulation trials and the Poisson hypothesized CDF with parameter as the mean of the simulated results. They appear to be close to each other, and it is interesting to observe that the dependency that is brought upon by the density-dependent velocity profile, and traffic signals do not seem to be destroying the Poisson property and stochastic independence of the stochastic model.

B. Evaluation With Cascaded Traffic Signals

We now evaluate the robustness of our model with more complicated traffic signal systems and demonstrate the potential of our model for modeling vehicular density in urban routes with the ramifications of consecutive traffic signals for system engineering and planning in both the transport and communication networks.

Fig. 4 depicts the new street configuration for simulations. We consider that there are two traffic signals that are 500 m apart, located at the 2 and 2.5 km locations, respectively. Both of them have a 1-min cycle time, with 30-s green, 3-s amber, and 27-s red signal periods. The velocity profile is modeled in the same way as in the previous example according to the signal cycle. Assuming that cars only enter the route at location 0 with a rate of 30 cars/min and no cars join or depart at junctions in this example, we explore how the number of vehicles in the road segment $(2, 2.5]$ varies with the phase shift between the two signals. We consider three cases here. They are respectively: 1) the first and second signals are totally in phase; 2) the first signal is 15 s lag behind the second one; and 3) the first signal is 30 s lag behind the second one.

Fig. 5(a) plots the simulation results of the mean number of vehicles in the road region $(2, 2.5]$, and Fig. 5(b) the corresponding D-statistics during the time interval $(3, 5]$ for the three cases. First, we can see from Fig. 5(b) that the D-statistics remain small (at about 0.15) even with more complicated and oscillating traffic signal systems, except when the mean number of vehicles is small. The empirical and Poisson CDFs at time

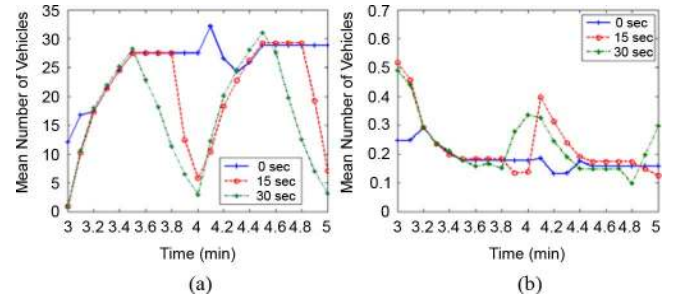


Fig. 5. (a) Mean number of vehicles in the road region $(2, 2.5]$. (b) Corresponding D-statistics against time with different phase shift between the two traffic signals.

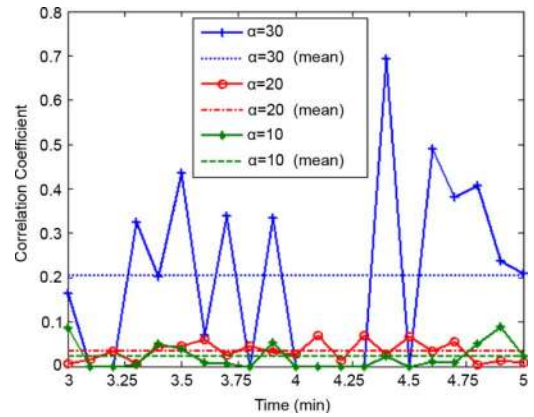


Fig. 6. Correlation coefficient in time interval $(3, 5]$ in the two-traffic-light scenario with different arrival rate (phase shift = 15 s).

instance 4.8 min for the three cases are plotted in Fig. 7, which again shows a very good match.

To evaluate the stochastic independence of the model with vehicle interactions, we examine the dependency of the number of cars in two nonoverlapping regions with vehicle interactions, raise the arrival rate α of vehicles to the urban route in steps from 10 to 30 cars/min, and examine the correlation between the number of vehicles in two consecutive regions $(2, 2.2]$ and $(2.2, 2.4]$ in time interval $(3, 5]$. We can see from Fig. 6 that as the arrival rate decreases, the correlation decreases. For $\alpha = 30$ cars/min, the average correlation is 0.2041 (the closer the coefficient is to 1, the stronger the correlation between the variables), while for 20 and 10 cars/min, it drops to 0.0334 and 0.00214, respectively. The reason behind this is that there are less interactions between vehicles with lower traffic load, and thus the stochastic independency of the PALM is better maintained. Therefore, given that the arrival rate is not too high, the degree of dependency between the number of vehicles in two nonoverlapping regions is low.

From the numerical results, we can see that the signaling system plays an important role in distributing vehicles in both space and time dynamics. For example, it appears in Fig. 5(a) that the signaling combination with no phase shift has a more stable mean number of vehicles during the time interval $(3.5, 5]$ than the other two combinations with phase shift. Therefore, such observation motivates the future research direction of optimizing node connectivity in signalized roads by controlling the signaling system. For instance, we can have the signaling cycles or phase shift between consecutive traffic signals change according to vehicular density in order to maintain a sufficient

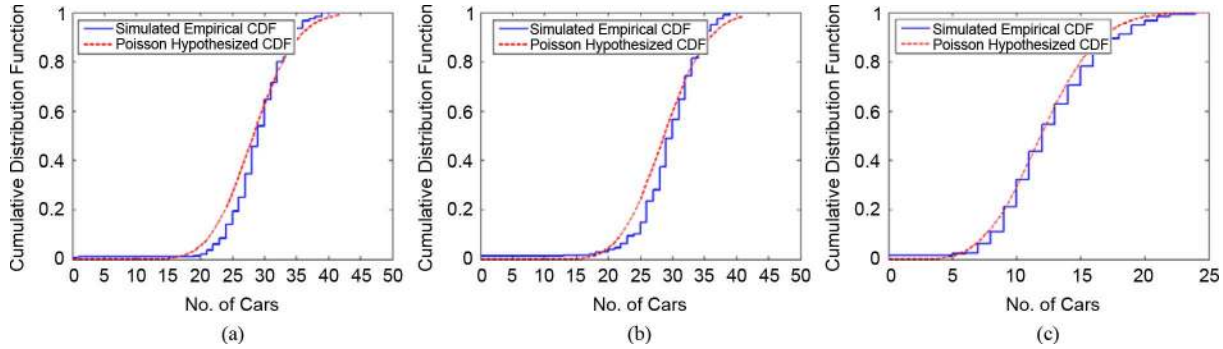


Fig. 7. Empirical and Poisson hypothesized CDFs of the number of vehicles in the road region $[2, 2.5)$ with (a) 0 s phase shift, (b) 15 s phase shift, and (c) 30 s phase shift at time instance 4.8 min.

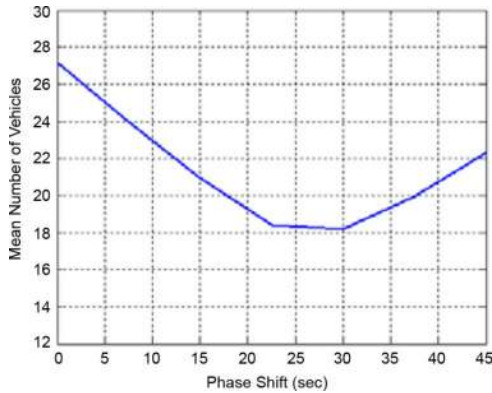


Fig. 8. Space-time averaged number of vehicles in the road region $(2, 2.5]$ from 3.4 to 4.4 min as a function of traffic signals' phase shift.

density level on each road segment for better connectivity in the communication network.

On the other hand, from the transport network point of view, we can control the traffic load on a road segment in order to reduce transport congestions. Fig. 8 shows how the mean number vehicles in the road region $(2, 2.5]$ within the time interval $(3.4, 4.4]$ varies with the phase shift between the two traffic signals, which appears that the minimal traffic congestion can be achieved if we set the phase shift to be about 30 s. Therefore, through the stochastic traffic model, we can identify the operating point for minimal transport congestion, and at such an operating point, our model can be further applied to identify regions with poor connectivity for the placement of roadside infrastructure nodes to boost the overall communication connectivity, which will be discussed in the following section. In this way, we can optimize both the transport (in terms of congestion) and communication networks (in terms of connectivity) based on the information provided by our model.

VI. MODEL VALIDATION WITH EMPIRICAL DATA

In this section, we utilize Inductive Loop Detector (ILD) data in central London to validate the key result of the stochastic traffic model that the number of vehicles within a road region (or vehicular density) at certain time-space instance is a Poisson process. The selected site is on Marylebone Road, London, which is a three-lane dual carriage-way with a dedicated bus lane in each direction. We consider the eastbound traffic at this site.

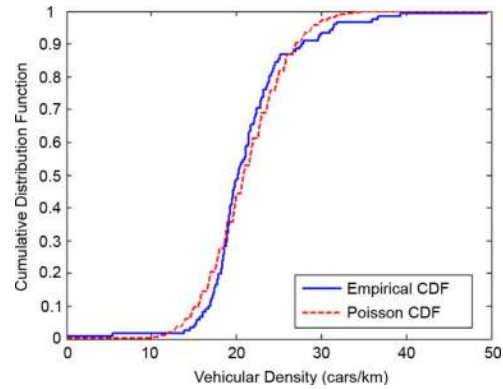


Fig. 9. Poisson CDF and empirical CDF of weekday vehicular density at Marylebone Road, London.

Flow and occupancy data for each lane were obtained from the ILD for each 15-min sampling period for six months, from August 1, 2008, to January 31, 2009. By assuming that the traffic patterns at a certain weekday hour are similar and independent, we derive the vehicular density from the flow, occupancy, and vehicle length data in the time window 8:30 to 8:45 (on a 24-h scale) on every weekday (Monday to Friday excluding public holidays to represent the peak-hour case). Due to the space limit, the reader is referred to [18, Sec. 4.6] for the detailed derivation of the vehicular density from empirical data.

Fig. 9 plots the CDF of the vehicular density from empirical data and the hypothesized Poisson CDF at the detector in time window 8:30 to 8:45 on weekdays, which on average has a traffic load of 33 cars per min. Interestingly, we can see from the figure that the key result of the stochastic traffic model (i.e., vehicular density as a Poisson process) represents a valid approximation even for signalized urban road networks. The D-statistic of K-S test for this case is only 0.1056, with a cutoff value of 0.1264 at a significance level of 0.05, which means the null hypothesis that the empirical distribution is Poisson cannot be rejected. We have also examined the time windows 23:30 to 23:45 to represent the off-peak-hour traffic, which corresponds to a traffic load of about 22 cars/min. The D-statistic for this case is 0.1074, with a cutoff value of 0.1149 at a significance level of 0.05, and again the null hypothesis cannot be rejected.

VII. ANALYSIS OF VANET CONNECTIVITY

In this section, we demonstrate that connectivity dynamics of vehicular networks can be characterized by the stochastic traffic

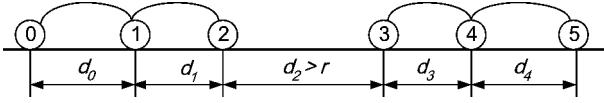


Fig. 10. Disconnection in a one-dimensional network.

model as well. Specifically, with the knowledge of the vehicular density dynamics from the proposed model, we determine the probability that the network within an urban road segment is connected and extend the investigation to k -connectivity of the network. In addition, we demonstrate how the traffic model can be utilized to identify poorly connected regions and facilitate the infrastructure node placement and transmission range adjustment processes for improving network connectivity.

To model wireless transmission between vehicles, a radio link model is assumed in which each vehicle has a transmission range r , and two vehicles are able to communicate (or are connected) directly with each other via a wireless link if the Euclidean distance between them is less than or equal to r . In a one-dimensional ad hoc network, a node is said to be *disconnected from the forward network* if it is not connected to any forward neighbors in the network (e.g., node 2 in Fig. 10), where “*forward neighbors*” here represents nodes that are on the right of the considered node assuming that the stream of traffic moves from left to right.

A network is said to be *connected* if, for every pair of nodes, there exists a path (that is composite of one or more number of communication links) between them, and otherwise it is *disconnected*. For the connectivity of a one-dimensional network, the following definition holds.

Definition 1: For a one-dimensional ad hoc network, it is connected if and only if there does not exist any nodes in the network that are not connected to any of the forward nodes.

In other words, the network is disconnected if the separation between any two adjacent nodes is greater than r , as illustrated in Fig. 10. Therefore, the probability that the network in a road segment is connected is equivalent to the probability that no nodes got isolated from the forward network. The reader is referred to the proof of Proposition 3, which considers the general case of a k -connected network.

A. Connectivity Without Vehicle Interactions

For analytical simplicity, we cease to focus on time dynamics here, and assume the system has reached a steady state with respect to time. Thus, all system variables and parameters (including the velocity profile) become independent of time, and we simply drop the variable t from our previously defined notation.

We define $N(x_a, x_b) = N(x_b) - N(x_a)$, where $x_b > x_a$, as the number of vehicles in the region $(x_a, x_b]$. According to the stochastic model, the actual number of vehicles distributed within a road region is Poisson-distributed. Thus, we can acquire from it the probability that a specific number of vehicles are located within a road region, which is given by

$$P(N(x) = n) = \frac{E[N(x)]^n}{n!} e^{-E[N(x)]} \quad (19)$$

$$P(N(x_a, x_b) = n) = \frac{E[N(x_a, x_b)]^n}{n!} e^{-E[N(x_a, x_b)]} \quad (20)$$

where $E[N(x)] = \int_0^x n(u) du$.

Consider a road segment of length L . Cars arrive at location 0 as a homogeneous Poisson process with mean rate α . Since we assume the system has reached a steady state with respect to time, the velocity profile is only a function of space, but independent of time. Thus, for any two vehicles that pass through the road with an interarrival time of I , the velocity–time graph of the second vehicle is the same as that of the first vehicle, but lags behind by time I . Let us denote such velocity–time function experienced by vehicles that pass through the road segment as $v(s)$. At time t , the physical separation between the i th and the $(i + 1)$ -th cars is

$$d_{I_i}(t) = \int_{t-I_i}^t v(s) ds \quad (21)$$

where I_i is the interarrival time between the i th and $(i + 1)$ -th arrivals. We are interested in finding the *critical interarrival time* of the road T_c , such that if $I_i \leq T_c$, the i th and $(i + 1)$ -th cars will remain connected throughout the whole journey. On the other hand, if $I_i > T_c$, the maximum separation between the two cars in the road segment will be greater than r . We have

$$\max_{t \in \Omega} \{d_{T_c}(t)\} = \max_{t \in \Omega} \left\{ \int_{t-T_c}^t v(s) ds \right\} = r \quad (22)$$

where Ω is the set of time instances such that the i th and $(i + 1)$ -th cars coexist in the road segment. Given the velocity profile, we are able to find T_c as a function of r . For Poisson arrival process with parameter α , $P(I_i \leq T_c) = 1 - e^{-\alpha T_c}$, let it be p_c . Therefore, the probability that the entire network remains connected (conditioning on that the population size of the road is nonzero, i.e., $N(L) > 0$) is

$$P(\text{net con}) = \frac{1}{1 - P(N(L) = 0)} \sum_{j=1}^{\infty} p_c^{j-1} P(N(L) = j). \quad (23)$$

Since the number of cars in the road segment has a Poisson distribution with parameter $E[N(L)]$, which can be computed from the fluid model, by substituting (19) into (23), we have

$$P(\text{net con}) = \frac{(e^{p_c E[N(L)]} - 1)}{p_c (e^{E[N(L)]} - 1)}. \quad (24)$$

As an illustrative example, consider a road segment of length 10 km. Given the velocity profile $v(x)$ as in Fig. 11(a), we can find that $T_c = r$, and we can compute from the fluid model that $E[N(L)] = 481$ for $\alpha = 30$ cars/min. Thus, according to (24), we plot the probability that the network in the road segment is connected as a function of the transmission range r in Fig. 11(b). The reader is referred to Section VII-E for additional examples and applications regarding the tradeoff between transmission range and connectivity. In real practice, the velocity profile can be obtained through empirical measurements, and more complicated velocity profiles can be handled by our models.

B. Connectivity With Vehicle Interactions

Section VII-B gives us an exact treatment of modeling the spatial separations between vehicles and connectivity of the network. Strictly speaking, when vehicle interactions are considered, the stochastic model is no longer valid. Therefore, we can only do approximation here and use simulation to evaluate its validity.

For velocity profile that is a function of both the space and time, and with vehicle interactions, we approximate the probability of connectivity based on the results of the fluid model,

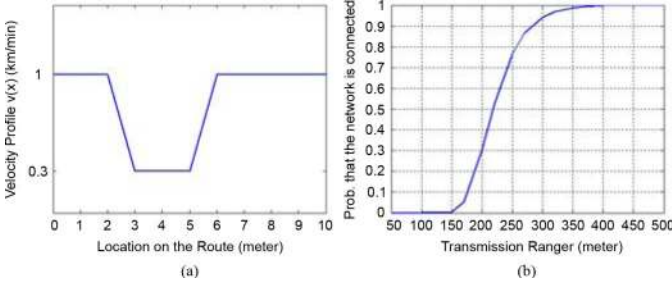


Fig. 11. (a) Velocity profile. (b) Corresponding probability that the network is connected as a function of transmission range r .

which captures the time and space dynamics as well as vehicular interactions. We now consider a specific time instance t_0 , and thus the variable t is again simply dropped from our notation. The probability over a period of time can be obtained by taking the time-average of multiple time instances.

Again, we consider a road segment with length L in region $(0, L]$. With the knowledge of the mean density profile $n(x)$ from the fluid model, we can derive the pdf

$$f_L(x) = n(x)/E[N(L)] \quad (25)$$

such that $f_L(x)\Delta x$ represents the probability that a random node in the road segment is located in the small region $(x, x + \Delta x]$. Given that there are j nodes located in the road segment at a time instance and assuming that their locations are independent (even when there are vehicle interactions, we have shown that the degree of dependency remains low if the traffic load is not too high), we now regard a randomly chosen node located in the road segment. The probability that this node is disconnected from the forward network is given by the weighted sum of the probability that the other $j - 1$ nodes are not located within the transmission range in front over all possible locations of the node in the road segment.

$$\begin{aligned} &P(\text{node discon} | j \text{ nodes}) \\ &= \int_0^{L-r} P(\text{the other } j-1 \text{ nodes are not in } (x, x+r]) \\ &\quad \times f_L(x) dx \\ &= \int_0^{L-r} \left(1 - \int_x^{x+r} f_L(x) dx\right)^{j-1} f_L(x) dx. \end{aligned} \quad (26)$$

Note that we define a node is not disconnected if it is located in region $(L - r, L]$ in the road segment.

According to [5], the events that an individual node is isolated or disconnected from the forward network are almost independent from node to node with the assumptions that the number of nodes in the road segment $N \gg 1$ and $r \ll L$. Thus, the probability that there are no disconnected nodes in the road segment or the network is connected is

$$\begin{aligned} &P(\text{net con}) \\ &= P(\text{no discon node}) \\ &\simeq \sum_{j=0}^{\infty} (1 - P(\text{node discon} | j \text{ nodes}))^j P(N(L) = j). \end{aligned} \quad (27)$$

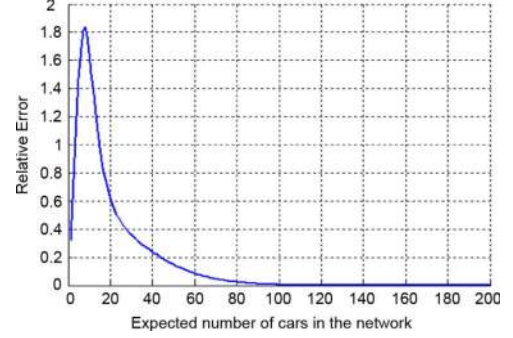


Fig. 12. Relative error = $|\tilde{P}(\text{net con}) - P(\text{net con})|/P(\text{net con})$ as a function of the expected number of cars in the network for homogenous Poisson arrival with constant velocity profile.

By substituting (19) and (26) into (27), we are able to find this probability based on the input $n(x)$ computed from the fluid model.

When the expected number of cars in the network is large, for the simplicity of calculation, we can approximate the probability above with

$$\begin{aligned} &\tilde{P}(\text{no discon node}) \\ &= \sum_{j=0}^{\infty} (1 - P(\text{node discon}))^j P(N(L) = j) \end{aligned} \quad (28)$$

where

$$\begin{aligned} &P(\text{node discon}) \\ &= \int_0^{L-r} P(\text{node discon} | x) f_L(x) dx \\ &= \int_0^{L-r} P(N(x + \Delta x, x + r) = 0) f_L(x) dx. \end{aligned}$$

Substitute (20) into it, and we have

$$P(\text{node discon}) = \int_0^{L-r} e^{-E[N(x+\Delta x, x+r)]} f_L(x) dx. \quad (29)$$

By substituting (19) and (29) into (28), we have

$$\begin{aligned} &\tilde{P}(\text{net con}) \\ &= \tilde{P}(\text{no discon node}) \\ &= \sum_{j=0}^{\infty} \left(1 - \int_0^{L-r} e^{-E[N(x+\Delta x, x+r)]} f_L(x) dx\right)^j \\ &\quad \times \frac{E[N(L)]^j}{j!} e^{-E[N(L)]} \\ &= \exp\left(-E[N(L)] \int_0^{L-r} e^{-E[N(x+\Delta x, x+r)]} f_L(x) dx\right). \end{aligned} \quad (30)$$

To verify the tightness of approximating (27) with (30), we plot the relative error $|\tilde{P}(\text{net con}) - P(\text{net con})|/P(\text{net con})$ as a function of the expected number of cars in the network in Fig. 12 for a homogeneous Poisson arrival with constant speed. From that, we can see that the error of the approximation converge to zero as the expected number of cars in the network is large. For instance, when there are 100 expected cars in the network, the error is less than 1%. The expected number of cars in the network is a function of the arrival rate, velocity, and the length of the network. In general, given that the length of the road segment L considered

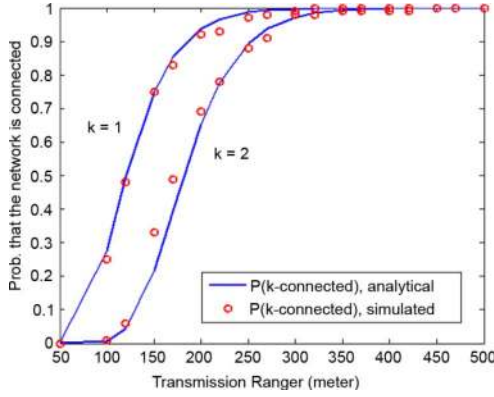


Fig. 13. Analytical and simulated probability that the network in road region (1.5, 2.5] at time instance 4.5 min in the two-traffic-light scenario is k -connected as a function of transmission range r .

is long enough, we can have a sufficient expected number of cars, and thus $P(\text{net con})$ can be well approximated by (30). Further evaluation of the analytical results with consideration of vehicle interactions is shown later in Fig. 13.

C. k -Connectivity

In Sections VII-B and VII-C, we consider specifically the 1-connectivity of the network, we now proceed to investigate the general case of k -connectivity. Once again, the result is not applicable in principle to cases with vehicle interactions because of the existence of dependency. However, our previous evaluations show that the Poisson property and stochastic independence of the model can still be primarily retained when the traffic load is not too high even in the presence of vehicle interactions. Therefore, we treat the results for cases with vehicle interactions as an approximation and verify its validity through simulations here.

For connected one-dimensional ad hoc network, we can characterize the degree of connectivity by examining the *forward node degree* of nodes, which represents the number of direct single-hop forward neighbors of a node. Given that a node is located at x , we define that the forward node degree of the node as the number of nodes in the region $(x + \Delta x, x + r]$. Let $K_F(x)$ denote the forward node degree of a given node located at x . With the PALM assumptions, we have

$$\begin{aligned} P(K_F(x) = k) &= P(N(x + \Delta x, x + r) = k) \\ &= \frac{E[N(x + \Delta x, x + r)]^k}{k!} e^{-E[N(x + \Delta x, x + r)]}. \end{aligned} \quad (31)$$

To associate forward node degree to connectivity in a one-dimensional communication network, we use a geometric graph $G = G(V, E)$ to represent the ad hoc network, which consists of a set of nodes (vertices) and a set of communication links (edges). There is an edge between two vertices i and j if and only if the Euclidean distance between them $|i - j| \leq r$.

Let us label the vertices in the one-dimensional geometric graph from left to right as nodes 1, 2, 3, and so on. If there is an edge (communication link) between vertices (nodes) i and j , where $j > i$, it directly implies there are edges between vertex i and vertices located in between nodes i and j since those node pairs have shorter Euclidean distances than that between nodes i and j .

With reference to [4], a graph is said to be k -connected if for each node pair, there exist at least k mutually independent paths connecting them. Also, graph is k -connected if and only if no set of $(k - 1)$ nodes exists whose removal would disconnect the graph. Then, we have the following proposition.

Proposition 3: In a one-dimensional geometric graph G , let $K_{F\min}(G)$ denotes the *minimum forward node degree* of graph G , then

$$P(G \text{ is } k\text{-connected}) = P(K_{F\min}(G) \geq k). \quad (32)$$

Proof: We divide the proof into two parts. First, we prove that $P(G \text{ is } k\text{-connected}) \leq P(K_{F\min}(G) \geq k)$, followed by proving $P(G \text{ is } k\text{-connected}) \geq P(K_{F\min}(G) \geq k)$. The combination of them proves Proposition 3. \square

Proposition 3.1: $P(G \text{ is } k\text{-connected}) \leq P(K_{F\min}(G) \geq k)$

Proof: This is equivalent to proving (G is k -connected) implies $(K_{F\min}(G) \geq k)$, which can be proven by contradiction. Assume on the contrary that $(K_{F\min}(G) = k - 1 < k)$, that is, every node only connects to at least $k - 1$ forward neighbors. In this case, if we remove $k - 1$ nodes that are forward neighbors of a specific node i that has only $k - 1$ forward neighbors, node i will be disconnected from the forward network, and the network is disconnected. Hence, the network (graph) is not k -connected. \square

Proposition 3.2: $P(G \text{ is } k\text{-connected}) \geq P(K_{F\min}(G) \geq k)$

Proof: This is equivalent to proving $(K_{F\min}(G) \geq k)$ implies (G is k -connected). In a one-dimensional network, assume that every node connects to at least k forward neighbors, i.e., $(K_{F\min}(G) \geq k)$, then it is trivial that no set of $k - 1$ nodes whose removal will disconnect the network. For example, in the worst case, if we remove $k - 1$ nodes whose are forward neighbors of a specific node i which has only k forward neighbors, node i is still connected to the forward network with one forward neighbor. Hence, the network (graph) is k -connected. \square

Consider the case when $k = 1$ in Proposition 3, we have $P(G \text{ is connected}) = P(K_{F\min}(G) \geq 1)$, that is, the network is connected if and only if every node has at least one forward neighbor, which is equivalent to Definition 1, and the probability can be found with (30). The generalized version of (30) for k -connected network is as follows.

For a randomly chosen node in the road segment, the probability that it has less than k forward neighbors is

$$\begin{aligned} P(K_F < k) &= \int_0^{L-r} P(K_F(y) < k | x) f_L(x) dx \\ &= \int_0^{L-r} \sum_{i=0}^{k-1} \frac{E[N(x + \Delta x, x + r)]^i}{i!} \\ &\quad \times e^{-E[N(x + \Delta x, x + r)]} f_L(x) dx. \end{aligned} \quad (33)$$

By Proposition 3 and following similar steps in the derivation of (30), we have

$$\begin{aligned} P(k\text{-connected}) &= P(K_{F\min} \geq k) \\ &\simeq \exp(-E[N(L)]) P(K_F < k) \end{aligned}$$

$$\begin{aligned}
&= \exp\left(-E[N(L)] \int_0^{L-r} \sum_{i=0}^{k-1} \frac{E[N(x+\Delta x, x+r)]^i}{i!}\right. \\
&\quad \left. \times e^{-E[N(x+\Delta x, x+r)]} f_L(x) dx\right). \quad (34)
\end{aligned}$$

Therefore, given the density profile $n(x)$ from the fluid model and the communication range r , the probability that the network in a one-dimensional road segment is k -connected can be found.

To verify the analytical results in (30) and (34), we compare them to simulated results. For the cascaded traffic lights scenario described in Section V-B, we plot the simulated and analytical results in Fig. 13 for the probability that the network is k -connected (for $k = 1$ and 2) as a function of the transmission range r . We consider the network in the road region (1.5, 2.5] at time instance 4.5 min. From the figure, we can see that the analytical and simulated results are close to each other even when vehicle interactions are considered by (16), and the k -connected probability grows to 1 when the transmission range is large enough.

D. Applications of the Stochastic Traffic Model

Through the stochastic traffic model, many system engineering and network management issues can be investigated. As an illustrative example, we demonstrate here with numerical results how the knowledge of connectivity dynamics from the proposed model facilitates infrastructure node assignment and transmission range adjustment of mobile nodes for boosting the overall network connectivity.

1) *Infrastructure Node Assignment*: Although we can identify road segments that are likely to be disconnected with (30), the range of a road segment is still too large for us to determine the optimal position for infrastructure node placement. To further reduce the potential range, we can examine the probability $P(K_F(x) > k)$ for all x belong to the road segment. The regions with low probability denote the desired locations for infrastructure nodes.

First, we need to determine a probability threshold $\varepsilon > 0$, then we maximize k (which is an integer) such that $\{\min_{x \in (x_a, x_b]} P(K_F(x) \geq k) > \varepsilon\}$ holds.

We define the maximum k value that satisfies the inequality above as the *maximum targeted connectivity* of the network in the road segment $(x_a, x_b]$, which is the maximum degree of connectivity that the network can achieve subjected to the probability threshold ε . Let us denote it as K_{\max} .

For all x belonging to the road segment $(x_a, x_b]$, we define the region with $\{P(K_F(x) > k) \leq \varepsilon\}$ as the *k -connectivity gap*. After knowing K_{\max} of the road segment, we can easily identify the connectivity gaps as the regions with

$$\{P(K_F(x) > K_{\max}) \leq \varepsilon\} \forall x \in (x_a, x_b].$$

Specifically, if $K_{\max} = 0$, the connectivity gap represents the region that is likely to be disconnected in the network, where infrastructure nodes should be placed to maintain connectivity.

To illustrate how the stochastic traffic model can be used to identify a poorly connected region and facilitate the infrastructure node placement process, we explore connectivity in the road segment (2, 2.5] time-averaged over the time interval (3, 5] in the two-traffic-light scenario described in Section V-B. We consider vehicles arrive at constant rates (denoted by α) of

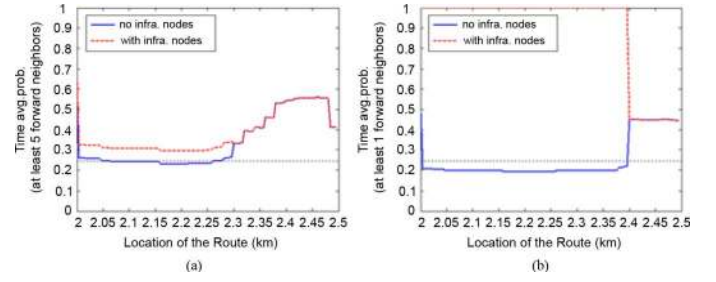


Fig. 14. Connectivity ($\varepsilon = 0.25$) with and without infrastructure nodes in road region (2, 2.5] averaged over the time interval (3, 5] in the two-traffic-light scenario with phase shift = 30 s, transmission range = 100 m. (a) Arrival rate = 30 cars/min. (b) Arrival rate = 3 cars/min.

30 and 3 cars/min respectively represent the peak-hour and off-peak-hour traffic. Transmission range of vehicles is assumed to be 100 m unless otherwise specified. We set the phase shift between the two traffic lights to be 30 s, which corresponds to the operating point with minimal congestion according to Fig. 8, so as to emphasize the potential applications of the proposed model on optimizing both the transport and communication networks.

For the peak-hour case ($\alpha = 30$ cars/min), we find that $K_{\max} = 4$ with the probability threshold ε set to be 0.25. Fig. 14(a) plots the time-averaged $P(K_F(x) > 4)$ for location points within the road segment (2, 2.5], from which we can see that the connectivity gaps (where $P(K_F(x) > 4) \leq 0.25$) are region (2.05, 2.28]. In order to further boost the connectivity of the network, we place infrastructure nodes to cover the connectivity gaps. Assuming that infrastructure nodes have the same communication range as mobile nodes (i.e., 100 m), we place three infrastructure nodes at locations 2.1, 2.2, and 2.3 km. With the introduction of infrastructure nodes, we can see from the dashed line in the figure that $\min_{x \in (2, 2.5]} P(K_F(x) > k)$ has been raised above the probability threshold, and K_{\max} now becomes 5.

For the off-peak-hour case, we find that $K_{\max} = 0$, since we can see from Fig. 14(b) that $P(K_F(x) > 0) \leq 0.25$ in region (2, 2.4]. Note that the inverse of the solid line in Fig. 14(b) denotes $P(K_F(x) = 0)$, which characterizes the probability that the location point x is isolated from the forward network, thus the connectivity gap (where $P(K_F(x) = 0) > 0.75$) actually represents the likely disconnected region in the network. We add four infrastructure nodes at locations 2.1, 2.2, 2.3, and 2.4 km to prevent disconnection of the network, which is shown by the dashed line in the figure.

Since infrastructure nodes are static, the regions that are covered by one newly added infrastructure nodes in the backward direction will have K_F be increased by 1 deterministically. Therefore, for all x that is covered by one infrastructure node, the curve $P(K_F(x) > K_{\max})$ is increased by $E[N(x+\Delta x, x+r)]^{K_{\max}} e^{-E[N(x+\Delta x, x+r)]} / K_{\max}!$, which is sufficient to rise above the threshold. To minimize the number of infrastructure nodes used for bringing the whole curve above the threshold, it is a one-dimensional coverage problem.

2) *Transmission Range Adjustment*: Other than adding road-side infrastructure nodes, we can also adjust the transmission ranges of vehicles to achieve an almost surely connected network.

As illustrated in Fig. 13, given the transmission range r , we can calculate from (34) the probability that the network in the

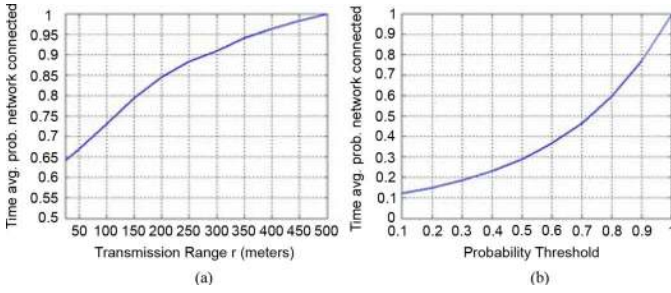


Fig. 15. Probability that the network in road region $(2, 2.5]$ is connected time-averaged over interval $(3, 5]$ as a function of (a) transmission range r and (b) the probability threshold ξ .

road segment is k -connected at a time instance. We can then take the time average of the probabilities to determine the critical transmission range for achieving certain degree of connectivity with high probability over a period of time. For instance, for the off-peak-hour case, we plot in Fig. 15(a) the time-averaged probability that the network is connected over time interval $(3, 5]$ as a function of transmission range, from which we can see that as the transmission range of vehicles increases beyond 370 m, the probability the network is connected is greater than 95%.

Instead of having all vehicles use large transmit power and thus transmission range, we can adjust the transmission range of an individual vehicle dynamically according to its location and the corresponding vehicular density $n(x, t)$ so as to prevent disconnection of the network.

At a time instance, given $f_L(x)$, which is derived from $n(x)$, we can adjust r to $r^*(x)$ such that $P(N(x + \Delta x, x + r^*(x)) \geq 1) \geq \xi$ for all x , where ξ is a certain probability threshold that is close to 1, to ensure connectivity. Hence, we have

$$e^{-E[N(x+\Delta x, x+r^*(x))]} \leq 1 - \xi. \quad (35)$$

Substitute (35) into (30), the probability that the network in the road segment $(0, L]$ is connected after such manipulation is

$$P(\text{net con}) \geq \exp\left(-E[N(L)](1 - \xi) \int_0^{L-r} f_L(x) dx\right). \quad (36)$$

Fig. 15(b) depicts the probability that the network is connected for the off-peak-hour case averaged over time interval $(3, 5]$ as a function of ξ . For instance, we can see from the figure that with $\xi = 0.95$, the time-averaged probability of connectivity is great than 0.9.

VIII. CONCLUSION AND FUTURE WORK

Vehicles in urban road networks do not distribute homogeneously primarily due to road layout and the presence of traffic lights. However, most of the existing studies on node connectivity have the implicit assumption that nodes are distributed homogeneously in the geographic area, which is inappropriate in VANETs and may lead to erroneous results. In addition, connectivity modeling of VANET in practical urban environment (with vehicle interactions and traffic signals) remains to be an unknown issue in the literature.

In light of the inadequacy, we have presented in this paper a stochastic traffic model for VANETs in signalized road systems to capture the time and space dynamics of moving vehicles in order to model mobility and connectivity in practical VANETs.

The fluid dynamic model treats vehicles as a continuous fluid, and the expected vehicular density is readily computable by solving the PDEs, while the stochastic model considers the behavior of an individual vehicle and provides probability distribution knowledge. As the key contribution, we have incorporated vehicle interactions in the model through a density-dependent velocity profile to closely approximate the practical road condition. With additional signaling inputs, it automatically captures the shockwave propagation of vehicles' stop-and-go motions at traffic signals.

Furthermore, we have validated the Poisson distributional result of the stochastic traffic model against real-world empirical inductive loop detector measurements in central London, and have provided numerical results to evaluate the quality of the model with approximation of vehicular interactions in terms of the expected value and distribution of its solution. In essence, we have shown that the fluid model well captures the evolution of the traffic stream, while the Poisson distributional conclusion of the stochastic model is still a good approximation for the distribution of the actual number of vehicles even when vehicles interact with each other as their movement is controlled by traffic lights. These validation results guarantee the robustness of connectivity analysis in this paper.

With the density knowledge acquired from the stochastic traffic model, we have derived the probability that the network in a road segment is k -connected and illustrated that the proposed model can be utilized for identifying the operating point for minimal transport congestion and regions with poor connectivity. In general, we have shown that the stochastic traffic model can serve as a useful tool for system engineering and planning for both the communication networks and transport systems.

Nowadays, most of the vehicles are installed with navigation systems that can collect velocity information. Such information can serve as an initial input to the stochastic traffic model for computing the vehicular density. The future evolution of the traffic can then be approximated by the density-dependent velocity profile, and the propagation of shockwave, compression, and rarefaction of the traffic stream due to traffic signals can also be automatically captured given the corresponding signaling information.

Our proposed model can be extended in several areas to study more complicated traffic scenarios, and it serves as the fundamental building block for constructing more elaborate urban traffic models—for example, routes with a greater number of segments and traffic signals and with arrival and departure of vehicles at road junctions. As other extensions, roads with multiple lanes, bidirectional traffic, and more complicated urban road topology (e.g., 2-D road network) can be represented by superposing multiple versions of urban routes.

ACKNOWLEDGMENT

The work reported in this paper forms part of the MESSAGE project. MESSAGE is a three-year research project that started in October 2006 and is funded jointly by the U.K. Engineering and Physical Sciences Research Council and the U.K. Department for Transport. The project also has the support of 19 nonacademic organizations from public sector transport operations, commercial equipment providers, systems integrators, and technology suppliers. More information is available

from the Web site <http://www.message-project.org>. The views expressed in this paper are those of the authors and do not represent the view of the Department for Transport or any of the nonacademic partners of the MESSAGE project.

REFERENCES

- [1] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Proc. IEEE CDC*, 1998, vol. 1, pp. 1106–1110.
- [2] P. Santi, D. Blough, and F. Vainstein, "A probabilistic analysis for the radio range assignment problem in ad hoc networks," in *Proc. ACM MobiHoc*, 2001, pp. 212–220.
- [3] L. Wang, B. Liu, D. Goeckel, D. Towsley, and C. Westphal, "Connectivity in cooperative wireless ad hoc networks," in *Proc. ACM MobiHoc*, 2008, pp. 121–130.
- [4] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM MobiHoc*, 2002, pp. 80–91.
- [5] C. Bettstetter, "On the connectivity of ad hoc networks," *Comput. J.*, vol. 47, no. 4, pp. 432–447, 2004.
- [6] S. Ukkusuri and L. Du, "Geometric connectivity of vehicular Ad Hoc networks: Analytical characterization," *Transp. Res. C*, vol. 16, pp. 615–634, 2008.
- [7] M. Khabazian and M. K. M. Ali, "A performance modeling of connectivity in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2440–2450, Jul. 2008.
- [8] W. A. Massey and W. Whitt, "Network of infinite server queues with nonstationary Poisson input," *Queueing Syst.*, vol. 13, pp. 183–250, 1993.
- [9] W. A. Massey and W. Whitt, "A stochastic model to capture space and time dynamics in wireless communication systems," *Prob. Eng. Inf. Sci.*, vol. 8, pp. 541–569, 1994.
- [10] K. K. Leung, W. A. Massey, and W. Whitt, "Traffic models for wireless communication networks," *IEEE J. Sel. Areas. Commun.*, vol. 12, no. 8, pp. 1353–1364, Oct. 1994.
- [11] C. F. Daganzo, *Fundamentals of Transportation and Traffic Operations*. New York: Pergamon Press–Elsevier Science, 1997.
- [12] K. R. Symon, *Mechanics*, 3rd ed. Reading, MA: Addison-Wesley, 1971.
- [13] C. J. Khisty and B. K. Lall, *Transportation Engineering*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [14] M. Brackstone and M. McDonald, "Car-following: A historical review," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 2, no. 4, pp. 181–196, 1999.
- [15] "Maintain a safe following distance (The 3 second rule)," 2010 [Online]. Available: <http://www.smartmotorist.com/traffic-and-safety-guideline/maintain-a-safe-following-distance-the-3-second-rule.html>
- [16] "TfL Central London congestion charging impacts monitoring sixth annual report," Transportation for London, London, U.K., 2008 [Online]. Available: <http://www.tfl.gov.uk/assets/downloads/sixth-annual-impacts-monitoring-report-2008-07.pdf>
- [17] "Kolmogorov-Smirnov Test," 2007 [Online]. Available: <http://www.physics.csbsju.edu/stats/KS-test.html>
- [18] I. W.-H. Ho, "Joint transport and communication models for vehicular ad-hoc networks," Ph.D dissertation, Imperial College London, London, U.K., 2010.



Ivan Wang-Hei Ho (M'10) received the B.Eng. and M.Phil. degrees in information engineering from the Chinese University of Hong Kong, Hong Kong, in 2004 and 2006, respectively, and the Ph.D. degree in electrical and electronic engineering from Imperial College London, London, U.K., in 2010.

He worked on the Mobile Environmental Sensing System Across a Grid Environment (MESSAGE) project and the International Technology Alliance (ITA) project during his Ph.D. studies. In 2007, he spent a summer working at the IBM T. J. Watson

Research Center, Hawthorne, NY. He is currently a Research Associate with the Department of Electrical and Electronic Engineering, Imperial College

London, working on the optimization of traffic control and management strategies with the aid of vehicular ad hoc network for minimizing vehicle emissions and fuel consumption in urban environment. His publications can be found at <http://www.commsp.ee.ic.ac.uk/~wanghei>. His research interests include the capacity of wireless ad hoc networks, power control in wireless networks, and joint transport and communication modeling and optimization in vehicular ad hoc networks.



Kin K. Leung (M'86–SM'93–F'01) received the B.S. degree from the Chinese University of Hong Kong, Hong Kong, in 1980, and the M.S. and Ph.D. degrees in computer science from the University of California, Los Angeles, in 1982 and 1985, respectively.

He joined AT&T Bell Labs, Holmdel, NJ, in 1986 and worked at its successor companies, AT&T Labs and Bell Labs of Lucent Technologies, until 2004. Since then, he has been the Tanaka Chair Professor at Imperial College London, London, U.K., where he serves as the Head of Communications and Signal Processing Group and as the Deputy Director for the University Defense Research Center in Signal Processing in the Electrical Engineering Department. His research interests include networking, protocols, optimization and modeling issues for wireless broadband, and sensor and ad hoc networks.

Prof. Leung received the Distinguished Member of Technical Staff Award from AT&T Bell Labs in 1994 and was a co-recipient of the 1997 Lanchester Prize Honorable Mention Award. He has received the Royal Society Wolfson Research Merits Award from 2004 to 2009. He has actively served on many conference committees. He is a member of the IEEE Fellow Evaluation Committee for the Communications Society for 2009 to 2011. He was a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC), *IEEE Wireless Communications*, and *Mobile Networks and Applications*, and an Editor for the JSAC: Wireless Series and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. Currently, he is an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the *International Journal on Sensor Networks*, and *ACM Computing Surveys*.



John W. Polak received the B.Sc. degree in mathematics from University College London, London, U.K., in 1979; the M.Sc. degree in transportation and traffic engineering from the University of Birmingham, Birmingham, U.K., in 1981; and the M.A. degree from the University of Oxford, Oxford, U.K., in 1991.

He is a Professor of transport demand and Director of the Centre for Transport Studies at Imperial College London, London, U.K., and is Research Director of the UK Transport Research Centre. He is a mathematician by background with over 30 years experience in transport research and teaching, specializing in the areas of mathematical and statistical transport modeling and analysis.

He has served as an advisor to central and local government and industry on a wide range of transport issues, both in the U.K. and overseas. He has been in the forefront of innovative transport model development in the U.K. for a number of years and has published extensively on a number of aspects of travel demand modeling, network performance estimation, network control and traffic management, and intelligent transport systems. Much of his recent research has been concerned with the collection, analysis, and interpretation of very large-scale real-time data sets related to operational, behavioral, attitudinal, and environmental aspects of transport.

Prof. Polak is a past President of the International Association for Travel Behaviour Research, a past Council Member of the Association for European Transport, and a past member of a number of TRB Committees. He serves on the editorial advisory boards of a number of leading international journals.