Modeling the Impact of Traffic Signals on V2V Information Flow

Jungyeol Kim*, Rohan Saraogi†, Saswati Sarkar‡ and Santosh S. Venkatesh§

Department of Electrical and Systems Engineering, University of Pennsylvania
Philadelphia, PA 19104, United States
Email: *jungyeol@seas.upenn.edu, †rohan.saraogi@gmail.com, ‡swati@seas.upenn.edu, §venkates@seas.upenn.edu

***Abstract*—Information propagation in V2V-enabled transportation networks is highly influenced by both vehicle mobility and wireless communication. The mobility patterns and communication conditions are not only heterogeneous, but also vary both temporally and spatially. In particular, realistic traffic flow changes with time, exhibiting sharp time-triggered transitions, due to external factors such as traffic lights, unpredictable disruptions (e.g., accidents), and planned disruptions (e.g., road-block). More specifically, traffic signals cause traffic synchronization, due to vehicles stopping during the red phase, and starting almost simultaneously during the green phase, which fundamentally alters the dynamics of V2V message propagation in a complex manner. In this paper, we propose a mathematical framework, starting from a continuous-time Markov chain, that characterizes the fraction of vehicles that have received a message over time and space in an arbitrary road network even when the traffic flow exhibits sharp time-triggered transitions. Our framework can accommodate arbitrary traffic synchronization patterns corresponding for example to the presence of an arbitrary number of traffic signals. The stochastic model for V2V message flow converges to a set of differential equations as the number of vehicles increases. The analytical characterization lends itself to a fast computation regardless of the number of vehicles and traffic synchronization patterns, while vehicular network simulators can only realistically simulate small-scale transportation networks. We find that V2V simulations of a statistical model with traffic synchronization and simulation of communications applied on a synthetic traffic trace well match our model solution.**

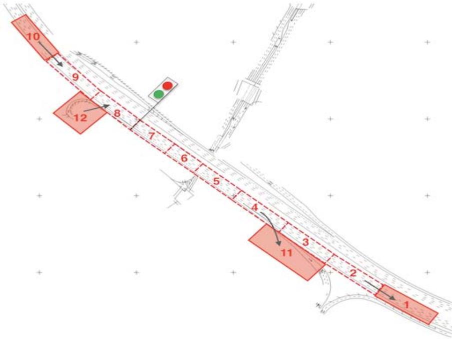
***Index Terms*—V2V communication, information propagation, traffic signals, traffic flow**

I. Introduction

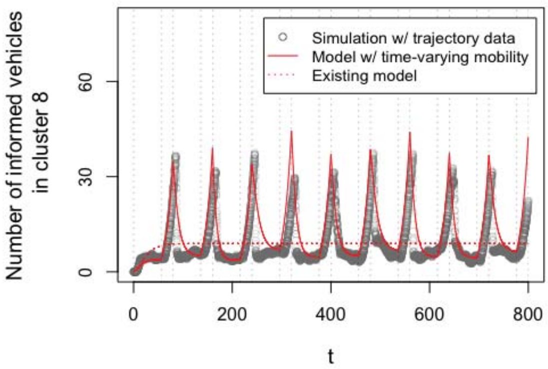
Motivation — Vehicle-to-Vehicle (V2V) is an emerging technology that enables vehicles to exchange information wirelessly, and thereby substantially ameliorate congestion and improve safety. We seek a mathematical framework for characterizing the fraction of vehicles that have received a V2V message under consideration (i.e., the fraction of informed vehicles) in an arbitrary road network. In our recent work [1], we had accomplished the same under the limitation that the mobility process can depend on the congestion levels in the system but do not otherwise depend on time.

However, in real systems, the mobility process changes with time not only through traffic congestion but also through sharp time-triggered transitions due to external factors such as traffic lights, unpredictable disruptions (e.g., accidents), or planned disruptions (e.g., road-block). In particular, traffic lights constitute an integral part of urban transportation networks. Traffic lights can be physical as is today, or can be virtual as envisioned for the future. For physical traffic lights, all vehicles stop when the light turns red, and they start moving when the light turns green. For virtual traffic lights, the simultaneous stopping and the resumption of movement are both accomplished through the exchange of V2V messages amongst the vehicles (refer to [2] for an example of a virtual traffic light protocol). Both physical and virtual traffic lights alter the dynamics of V2V message flow because of the following reasons: 1) all the vehicles simultaneously stop during a certain period and subsequently move during the next period, thus traffic flow becomes synchronized, and 2) different kinds of V2V messages can be quickly propagated amongst vehicles that wait in close proximity to each other at the red traffic light (or virtual equivalent). Ignoring these effects in the model lead to a significant divergence between the message propagation pattern obtained from the model and what arises in practice (e.g., see Figure 1). Note that Figure 1 show that the divergence can be significant even with one traffic light. The divergence may increase significantly more when there are multiple traffic lights due to the correlation between red and green traffic lights. Transportation networks in large cities have a large number of traffic lights (e.g., 2,820 intersections in Manhattan alone are controlled by traffic signals as of June 2011 [3]). Despite this, to the best of our knowledge, there does not exist analytical research on the impact of traffic lights, one or more, on the spatio-temporal dynamics of V2V message propagation. This is the void that we seek to fill in.

Literature review — A genre of work designs traffic signal control to improve traffic flow at intersections. For example, [5], [6] apply deep reinforcement learning to traffic light control problems based on the Markov decision process framework (refer to the technical report [7] for a detailed review). However, to the best of our knowledge this genre does not consider the flow of V2V messages. In addition, the works on MDP and reinforcement learning based control of traffic signals models the flow of vehicles as Markov processes. In principle, Markovian formulation (in presence or absence of V2V messaging capabilities) provides computation approaches only for the steady-state distributions (i.e. , distributions as the time approaches infinity) of traffic or information flow. The computation approaches are computationally intensive



(a)



(b)

Fig. 1. (a) Clustered U.S Highway 101 with single hypothetical traffic light. We assume that there is a traffic light between cluster 7 and 8. (This figure is modified from Figure 11a of [1]) (b) Fraction of informed vehicles in cluster 8 over time. The gray points are the average of 100 simulation runs for V2V message flow. The message flow has been obtained by superimposing a statistical communication process on a synthetic trace data, modified based on an actual trajectory dataset [4], to reflect the pulsed traffic. The dotted red line is the solution of the previous model in [1], which does not incorporate traffic lights. Notice that there is a significant difference between the gray points and the red dotted line. The solid red line is the solution of the model that we will present in this paper. The solid red line tracks the gray points more closely than the the dotted red line. The ratio between the average deviation over time between the gray points and the dotted red line and that between the gray points and the solid red line is approximately 1.7:1.

as these rely on the inversion of large transition probability matrices whose size rapidly increases with the increases in the number of vehicles, number of roads, etc. Computing the relevant distributions at any given time, which is what we seek, is even more hopeless. The technical innovation we introduce in this paper in the form of diffusion-based model for traffic and information flows (1) constitutes the asymptotic limit of a Markov formulation, (2) is computationally tractable (i.e., scales efficiently to a large number of vehicles), (3) and can accommodate arbitrary transportation networks and traffic synchronization patterns.

Contribution — We propose a mathematical framework, based on a continuous-time Markov chain, that characterizes the fraction of informed vehicles as a function of space and time in an arbitrary road network under the condition of pulsed vehicular traffic flow. In Section II, we show that the stochastic model converges to a set of differential equations (diffusion equations) as the number of vehicles increases, despite discontinuous temporal changes in the traffic flow. Our framework scales efficiently to a large number of vehicles in a large-scale vehicular network (i.e., metropolitan city) while most vehicular network simulators (e.g., VEINS) can only realistically simulate small number of vehicles due to large memory usage (Section V). Finally, we validate our model and evaluate the impact of several attributes through 1) simulation of statistical communications superimposed on synthetic traffic trace (Section III) and 2) simulations of statistical models for both communication and mobility (Section IV). We present a direction of future research to address a limitation in our model (Section VI).

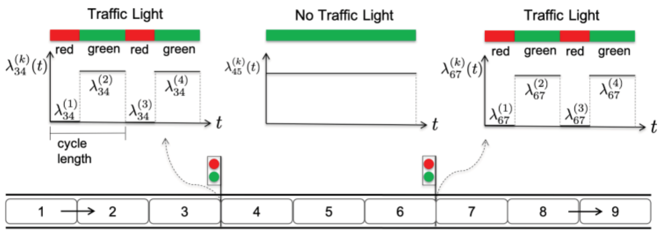


Fig. 2. (This figure is modified from Figure 14 of [1]) Clustered road topologies. The mobility rates between clusters with traffic lights are changed discontinuously depending on the mode switching over time.

II. Model Formulation

We develop a tool to model the spatio-temporal spread of V2V message in general transportation networks with pulsed vehicular traffic flow. We divide an arbitrary road into a set of J clusters. Suppose that N vehicles are located in one of the J clusters corresponding to a road segment (Figure 2 for an example). Vehicles can communicate and move between clusters or within a cluster at a rate determined by the communication environment, vehicle speed, and external factors such as traffic lights, unpredictable disruptions, or planned disruptions. We summarize the mathematical notations that we use in Table I.

We capture time-varying mobility by combining continuous evolution of state variables with discrete, instantaneous changes. The time axis is now subdivided into a sequence of *modes*, such that the mobility rates switch discontinuously to different values at the end of each mode, and remain constant during each mode. For example, the mode can be the time interval between color changes of traffic lights. We illustrate the concept of mode through the following example.

Example 1: Consider a transportation network with two traffic lights (Figure 2). Whenever any one traffic light changes color, the mode changes (Figure 3). If the two traffic lights are synchronized, the mode changes when the signals of both traffic lights are changed at the same time (Case 1 in Figure 3). In this case, the mobility rates discontinuously transition when both signals change simultaneously; the mobility rates in the clusters with traffic lights are 0 if the lights are red, and

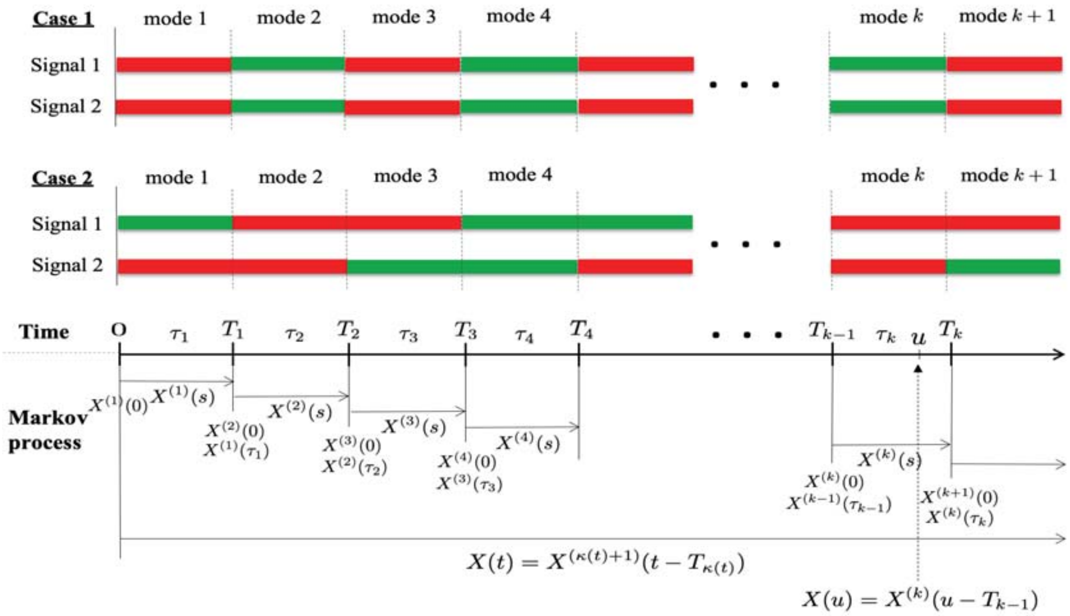


Fig. 3. Evolution of the process. Case 1 is an example of two traffic lights being synchronized, and case 2 is an example of not being synchronized.

TABLE I
MATHEMATICAL NOTATION

N	Total number of vehicles
$n_j^I(t)$	Number of informed vehicles in cluster j at time t
$n_j^S(t)$	Number of non-informed vehicles in cluster j at time t
$n_j^{I:k}(s)$	Number of informed vehicles in cluster j at time s from the beginning of the kth mode
$n_j^{S:k}(s)$	Number of non-informed vehicles in cluster j at time s from the beginning of the kth mode
$X(t)$	State process at time t
$X^{(k)}(s)$	State process at time s from the beginning of the kth mode
$X_N(t)$	Scaled process, $X(t)/N$
$X_N^{(k)}(s)$	Scaled process for the kth mode, $X^{(k)}(s)/N$
$\lambda_{ij}^I(\cdot)$	Mobility rate from cluster i to j for informed vehicles
$\lambda_{ij}^S(\cdot)$	Mobility rate from cluster i to j for non-informed vehicles
$\beta_{ij}^{(N)}$	Communication rate between a vehicle in cluster i and a vechicle in j
p_{jk}	Probability that a vehicle in cluster j move to k

other specified values if they are green. Figure 2 provides an example of transitions in mobility rates with changes in traffic lights. If the two traffic lights are not synchronized, the mode changes when either signal changes, and the mobility rate in the cluster with the corresponding traffic light transitions from 0 to the normal value or vice versa (Case 2 in Figure 3).

Let $\mathbf{n}^{I:k}(s)$ and $\mathbf{n}^{S:k}(s)$ represent J-dimensional vectors, whose j-th elements are the number of informed and non-informed vehicles respectively, in cluster j, at time s from the beginning of the kth mode, $k \in \mathbb{Z}_{>0}$. Consider the state process $X^{(k)}(s) = (\mathbf{n}^{I:k}(s), \mathbf{n}^{S:k}(s))$. The time duration of the k-th mode, τ_k, is defined as the time elapsed between the $(k-1)$-th and k-th signal changes, and T_k denotes the time at which the signal changes for the k-th time, thus τ_k and T_k are given by

$$\tau_k = T_k - T_{k-1}; \quad T_k = \textstyle\sum_{j=1}^{k} \tau_j, \quad T_0 = 0.$$

The last state of the k-th mode is considered the initial state of the next mode $k+1$ (i.e., $X^{(k+1)}(0) = X^{(k)}(\tau_k)$). This process is illustrated in detail in Figure 3.

We now consider the $2J$-dimensional vector $X(t) = (\mathbf{n}^I(t), \mathbf{n}^S(t))$. The state process $X(t) = \big(n_1^I(t), n_2^I(t), \ldots, n_J^I(t);\ n_1^S(t), n_2^S(t), \ldots, n_J^S(t)\big)$ represents the state at time t from the beginning of the whole process, not just from the beginning of each mode; the first J elements represent the number of informed vehicles for each cluster, and the remaining J elements after the semicolon represent the number of non-informed vehicles for each cluster. Naturally, the snapshot of the process at time t from the beginning of the whole process, $X(t)$, is given by

$$X(t) = X^{\kappa(t)+1}\left(t - T_{\kappa(t)}\right),$$

where $\kappa(t) = \max\{k | T_k < t\}$ is the total number of signal changes in the system by time t, regardless of which traffic light it is. The state space is contained in the set of points S^N:

$$S^N := \left\{(\mathbf{n}^I, \mathbf{n}^S) \mid n_j^I \geq 0,\ n_j^S \geq 0,\ j = 1, \ldots, J;\ \textstyle\sum_{j=1}^{J}\left(n_j^I + n_j^S\right) = N\right\}.$$

In each mode, we assume that the time until a vehicle in a cluster moves to a neighboring cluster is exponentially distributed with parameters depending on the states. Similarly, for both intra- and inter-cluster communication, we assume that the time between successful communications from an informed vehicle to a non-informed vehicle is exponentially distributed. Under these assumptions, the process $X^{(k)}(t)$, $k \in \mathbb{Z}_{>0}$, becomes a continuous-time Markov chain (CTMC).

The CTMC has the following three state transitions: 1) an informed vehicle moves from cluster i to cluster j, $j \neq i$; 2) a non-informed vehicle moves from cluster i to cluster j,

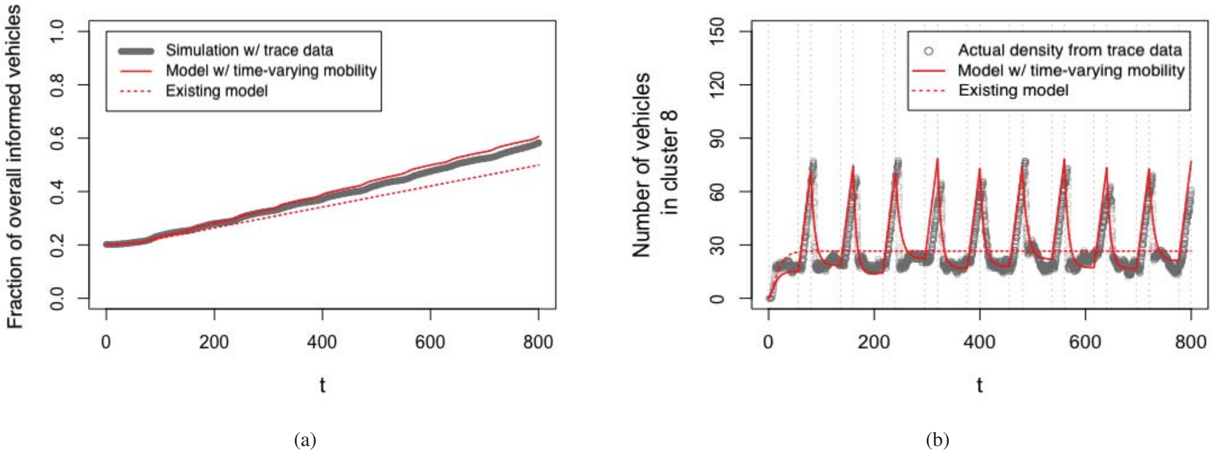


Fig. 4. (a) Fraction of overall informed vehicles over time. The average deviation over time between the gray points and the dotted red line is 0.3, while the average deviation between the gray points and the solid red line is 0.1. (b) Fraction of vehicles in cluster 8. The gray points represent the actual traffic density based on the synthetic trace data. The average deviation over time between the gray points and the dotted red line is 34.2, while the average deviation between the gray points and the solid red line is 16.2. The solid red line is the solution of the model with time-varying mobility, and the red dotted line is the solution of the previous model [1] assuming a consistent exponential mobility process.

$j \neq i$; and 3) a non-informed vehicle in a cluster j receives an information from an informed vehicle located in the same cluster or in a different cluster.

The first two types of state transitions capture the mobility of the vehicle in the system, which can vary depending on the mode. During the k-th mode, an informed and non-informed vehicle migrate along the road departing from the cluster i to j at rate $\lambda_{ij}^{I:k}(\cdot)$ and $\lambda_{ij}^{S:k}(\cdot)$, respectively. We assume that for every mode $k \in \mathbb{Z}_{>0}$, both $\lambda_{ij}^{I:k}(\cdot)$ and $\lambda_{ij}^{S:k}(\cdot)$ are bounded functions of $\frac{1}{N}(\mathbf{n}^{I:k}, \mathbf{n}^{S:k})$ if vehicle movement is possible from cluster i to j and 0 otherwise. Specifically, if the traffic light is located between cluster i and j, the parameters $\lambda_{ij}^{I:k}(\cdot)$ and $\lambda_{ij}^{S:k}(\cdot)$ switch along with the lights: they are 0 if the lights are red in mode k, and at the normal values if they are green in mode k (Figure 2). Finally, the last type of state transition captures the successful communication between an informed vehicle and a non-informed vehicle. A vehicle in cluster i successfully communicates with a vehicle in the same cluster i at rate $\beta_{ii}^{(N)}$. A vehicle in cluster i also successfully communicates with a vehicle in a distinct cluster j at rate $\beta_{ij}^{(N)}$ if the distance between cluster i and j is within communication range and 0 otherwise. The transitions, therefore, capture distinctions in the vehicular routing choices, vehicular speeds, their communication choices etc. based on whether vehicles have the message or not and vehicular congestion in local clusters.

The CTMC can be approximated by a set of ordinary differential equations in the continuum. We define a set $E := \big\{(\mathbf{I},\mathbf{S}) \mid I_i \geq 0, S_i \geq 0, i = 1,2,\ldots,J;\ \sum_{i=1}^{J}(I_i + S_i) = 1\big\}$ with $\big(\mathbf{I},\mathbf{S}\big) = (I_1, I_2, \ldots, I_J;\ \ S_1, S_2, \ldots, S_J)$. Define $\beta_{ij} := N\beta_{ij}^{(N)}$, and suppose β_{ij} is constant. There is implicit assumption that the larger the total number of vehicles, the lower the communication rate $\beta_{ij}^{(N)}$ with limited bandwidth.

We now introduce the formal notation $\lim_{N\to\infty} \frac{\mathbf{n}^I(t)}{N} = \mathbf{I}(t)$ and $\lim_{N\to\infty} \frac{\mathbf{n}^S(t)}{N} = \mathbf{S}(t)$. Note that $\mathbf{I}(t)$ and $\mathbf{S}(t)$ respectively represent the fraction of informed and non-informed vehicles in each cluster. The following theorem provides sufficient conditions for the convergence of the scaled process $X_N(t) = \frac{1}{N}\big(\mathbf{n}^I(t), \mathbf{n}^S(t)\big)$ to a solution of differential equations, $\mathbf{x}(t) = \big(\mathbf{I}(t), \mathbf{S}(t)\big)$, despite the mobility rates changing discontinuously depending on the mode.

Theorem 1: Suppose for $i, j = 1, 2, ..., J$ and $i \neq j$, mobility rate functions $\lambda_{ij}^{I:k} : E \to \mathbb{R}$ and $\lambda_{ij}^{S:k} : E \to \mathbb{R}$ in every mode $k \in \mathbb{Z}_{>0}$ are bounded and Lipschitz continuous on E. Let $\big(\mathbf{I}(0), \mathbf{S}(0)\big) = \lim\limits_{N\to\infty} \frac{1}{N}\big(\mathbf{n}^I(0), \mathbf{n}^S(0)\big)$, and $\big(\mathbf{I}(t), \mathbf{S}(t)\big)$ satisfies the following set of differential equations:

$$\dot{I}_i(t) = -\sum_{j\neq i}^{J} \lambda_{ij}^{I:k}(\mathbf{I},\mathbf{S}) \cdot I_i + \sum_{j=1}^{J} \beta_{ji} I_j S_i + \sum_{j\neq i}^{J} \lambda_{ji}^{I:k}(\mathbf{I},\mathbf{S}) \cdot I_j \quad (i = 1, 2, \ldots, J),$$

$$\dot{S}_i(t) = -\sum_{j\neq i}^{J} \lambda_{ij}^{S:k}(\mathbf{I},\mathbf{S}) \cdot S_i - \sum_{j=1}^{J} \beta_{ji} I_j S_i + \sum_{j\neq i}^{J} \lambda_{ji}^{S:k}(\mathbf{I},\mathbf{S}) \cdot S_j \quad (i = 1, 2, \ldots, J),$$

where $\lambda_{ij}^{I:k}(\cdot)$ and $\lambda_{ij}^{S:k}(\cdot)$ are valid over the time interval $t \in [T_{k-1}, T_k)$, $k \in \mathbb{Z}_{>0}$. Then

$$\lim_{N\to\infty} \sup_{s\leq t} \big| \tfrac{1}{N}\big(\mathbf{n}^I(s), \mathbf{n}^S(s)\big) - \big(\mathbf{I}(s), \mathbf{S}(s)\big) \big| = 0 \ \text{ a.s. for all } t > 0.$$

We provide the proof in the technical report [7]. When sharp transitions between modes occur, mobility functions of the equations are instantaneously changed, but $\mathbf{x}(t)$ is not reset, so executions of the system are always continuous.

III. Empirical validation with traffic trace data

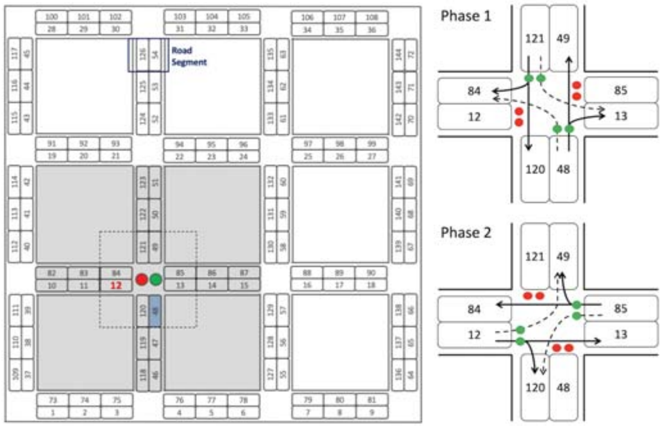
We empirically validate the model considering cases in which our modeling assumptions do not hold. Specifically, the theorem in Section II ensures that actual fraction of informed and non-informed vehicles in all the clusters converge to the corresponding elements of the solution of the differential equations when the number of vehicles approach infinity and vehicles satisfy the exponential sojourn time assumption within each mode. We now consider synthetic trajectory data involving mobility of a finite number of vehicles. The synthetic trajectory data does not satisfy the exponential sojourn time assumption within each mode. This is partly because the traffic lights hold up vehicles for a deterministic duration. Through the empirical validation, we confirm that our mathematical model approximates well the simulation results of V2V message flow involving such trajectory data.

We use the first 831.7 seconds of actual trace data [4] collected from the U.S. Highway 101 in Los Angeles, California. To reflect the pulse traffic caused by a traffic light, we slightly modify the trace data assuming that there is a traffic light on the road. As in [1], we first divide the road into clusters (Figure 1a). We assume that approximately 20% of all vehicles (i.e., 402 out of a total of $N = 1993$ vehicles) in the entry clusters (clusters 10 and 12) already received information before entering the study area (set of clusters {2,3,...,9}). We also assume that the information propagation occurs only in the study area. Suppose there is a traffic light between cluster 7 and 8. The traffic signal has cycle lengths of 60 s, and 30% of the time is spent in red. The signal begins their cycles in green at the start of the simulation. We modify the trajectory of vehicles arriving in cluster 8 during the red light, assuming that these vehicles must be stationary waiting for the signal to turn green. When the light turns green, after the delay during the red light, the vehicle starts to move again according to the trajectory of the original data. Thus, synthetic data with synchronized traffic flow can be generated.

We superimpose the statistical communication process on the trace data (averaged over 100 runs), and then obtain the solutions of both the model with time-varying mobility (introduced in this paper) and the model introduced in prior work [1] (which does not consider sharp time-triggered transitions of mobility patterns). Figures 4a and 4b show that the model with time-varying mobility provides a good match respectively of the fraction of overall informed vehicles and the number of vehicles in an example cluster (cluster 8) as functions of time. The match happens despite the fact that the traffic trace corresponds to a finite number of vehicles and does not satisfy the exponential sojourn time assumption within each mode. Figure 4a and 4b also show that the match is significantly better than what we observe for the model introduced in prior work [1].

IV. Simulation validation

Using statistical simulations, we investigate the impact of the attributes that do not arise in the synthetic dataset used in Section III: (1) two-dimensional road topology, (2) different durations for each phase of a signal cycle, and (3) multiple traffic lights. We show that the solution of our model well matches the simulation result for V2V message flow in the presence of one or more traffic lights in two-dimensional grid topology. We also show that the delay in the propagation of V2V messages across the network significantly depends on the relative durations of red and green lights.



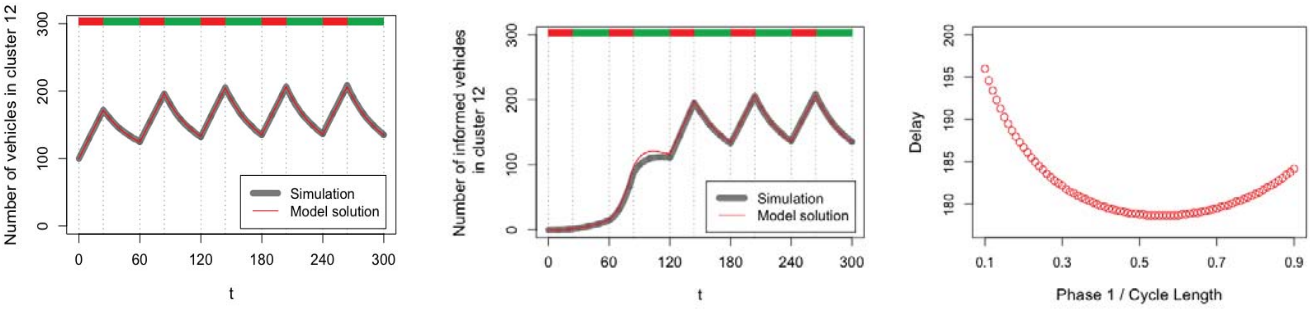
(a) (b)

Fig. 5. (a) Clustered road topology with a traffic signal at one intersection. Since two clusters on the same road segment are adjacent to each other, it is assumed that vehicles in the same segment can communicate with each other. (b) This is an enlarged view of the dotted rectangle around the traffic signal in Figure (a), and represents traffic flow in the 2 phases. Broadly, the phase 1 corresponds to vehicles moving straight in the vertical direction (e.g., cluster 48 to 49), and the phase 2 corresponds to vehicles moving straight in the horizontal direction(e.g., cluster 85 to 84). In addition, vehicles can make left and right turns as indicated by the arrows; for example, vehicles in cluster 48 can move to cluster 13 and 84 during phase 1, and vehicles in cluster 85 can move to cluster 49 and 120 during phase 2.

We first consider the presence of traffic signal at one intersection in a two-dimensional grid road topology (Figure 5a). The traffic signal is located at the intersection indicated by the dotted rectangle. The traffic phase design consists of phase 1 and 2 (Figure 5b) and the cycle length is set to 60 seconds. Refer to Figure 5b and its caption to see what the phases mean.

The mobility rates $\lambda_{ij}^{I:k}(\cdot)$ and $\lambda_{ij}^{S:k}(\cdot)$ for clusters $i \in \{12, 48, 85, 121\}$, located right in front of traffic light, are different depending on the phase; the parameters switch along with the lights: they are 0 if the lights are red in mode k, and are at the normal values if the lights are green in mode k. On the other hand, the mobility rates for the clusters $i \notin \{12, 48, 85, 121\}$ are constant regardless of the modes.

Let the neighborhood of cluster i be the set of clusters $N_G(i)$ adjacent to cluster i; vehicles in cluster i can only directly move to $j \in N_G(i)$. Let p_{ij} be the probability that vehicles in cluster i move to cluster $j \in N_G(i)$; thus $\sum_{j \in N_G(i)} p_{ij} = 1$. For the clusters $i \notin \{12, 48, 85, 121\}$ that are not directly controlled by the traffic light, we assume the routing probability is uniform (i.e., $p_{ij} = 1/|N_G(i)|$). For the clusters $i \in \{12, 48, 85, 121\}$ located right in front of the traffic lights, p_{ij} is determined by whether it is a left turn or not. In order to reflect the opportunistic left turn in the absence of traffic flow conflicts, we assume that the routing



(a) (b) (c)

Fig. 6. (a) The number of total vehicles in cluster 12 over time. (b) The number of informed vehicles in cluster 12 over time. The red and green bars on the top indicate the color of the traffic light over time, from the perspective of vehicles located in cluster 12. For (a) and (b), the gray lines represent the average of 100 simulation runs, and the red lines are the model solutions. The traffic signals have cycle lengths of 60 s, and 40% of the cycle time is spent in phase 1. (c) The time delay until 99% of vehicles in the shaded region receive information, which is obtained from the differential equations. This represents how quickly the V2V message spreads to many vehicles.

probability for the left turn at a green traffic light is half of the probability of the right turn and straight ahead. The mobility rate set in mode k is

$$\lambda_{ij}^{I:k}(\cdot) = \lambda_{ij}^{S:k}(\cdot) = \begin{cases} p_{ij}\lambda, & i \notin \{12, 48, 85, 121\} \\ p_{ij}\lambda, & \text{if } i \in \{12, 48, 85, 121\}, \text{ and green light,} \\ 0, & \text{if } i \in \{12, 48, 85, 121\}, \text{ and red light} \end{cases}$$

where λ is constant. We set the parameters to $\lambda = 0.03$, and $\beta_{ij} = 5$.

Initially, 14400 vehicles are uniformly distributed in the system (i.e., 100 vehicles per cluster). The shaded area with a traffic light in Figure 5a represents a specific region of interest to investigate information propagation near the traffic signals. We assume that the V2V message initially propagates from 10 vehicles in cluster 48, represented by blue in Figure 5a.

Figure 6a shows that the model solution well captures the traffic flow in cluster 12 despite the burstiness caused by the traffic signal; a signal phase 1 results in traffic accumulating in cluster 12 because vehicles stop at the intersection, followed by the lower traffic volume during a signal phase 2. More importantly, Figure 6b shows that the model well captures the impact of the traffic signal on the spatio-temporal dynamics of message propagation (specifically, the number of informed vehicles in cluster 12 over time).

We now show that the delay of the V2V message propagation is significantly affected by the relative durations of the phases in the traffic light, more specifically phase 1 and 2 in Figure 5b. As Figure 6c shows, the time it takes for most vehicles (99%) in the shaded region in Figure 5a to receive information is highly dependent on the time proportion of phase 1.

Recall that the mathematical framework can incorporate an arbitrary number of traffic lights and arbitrary coordination patterns between them. We show how well the model captures the V2V propagation process in the presence of multiple traffic signals. We consider same road topology with traffic signals at two intersections (Figure 7a). The simulation setup is the same as the previous one. The only difference is that with the addition of another traffic light (Signal 2 in Figure 7a), the mobility rate near the added traffic light also changes depending on the phase. We assume that the traffic phase design for the added traffic signal 2 is the same as the phase design for the original traffic signal 1, so both traffic lights have identical traffic phase design, as illustrated in Figure 5b. We also assume that both signals are synchronized, so the mode changes at the same time when the phases of both traffic lights are changed. Figure 7b and 7d show that the model solution captures the pulsed traffic movement caused by the traffic signals well. Further, Figure 7c and 7e show that the temporal change in the number of informed vehicles in the particular locations (clusters 12 and 15, respectively) is also well approximated by the model solution.

V. Computation time

As Theorem 1 states, as the number of vehicles approach infinity, actual fraction of informed and uninformed vehicles in all the clusters converge to the corresponding elements of the solution of the differential equations. And, the computation time taken to solve the set of differential equations does not increase with increase in the number of vehicles. Thus our analytical characterization is computationally tractable regardless of the total number of vehicles in the system. The number of both differential equations and variables are twice the total number of J clusters. An interesting question is how the computation time depends on the number of modes. We answer this considering the case of K traffic lights in a system. The differential equation for a cluster just utilizes the mobility rates in the cluster and in the clusters that feed into the cluster. These mobility rates can be obtained when the states of the corresponding traffic signals are known. If, for example, the signals change states periodically (regardless of any synchronization between them), the states of the corresponding signals at any given time, and hence during a mode, can be obtained in

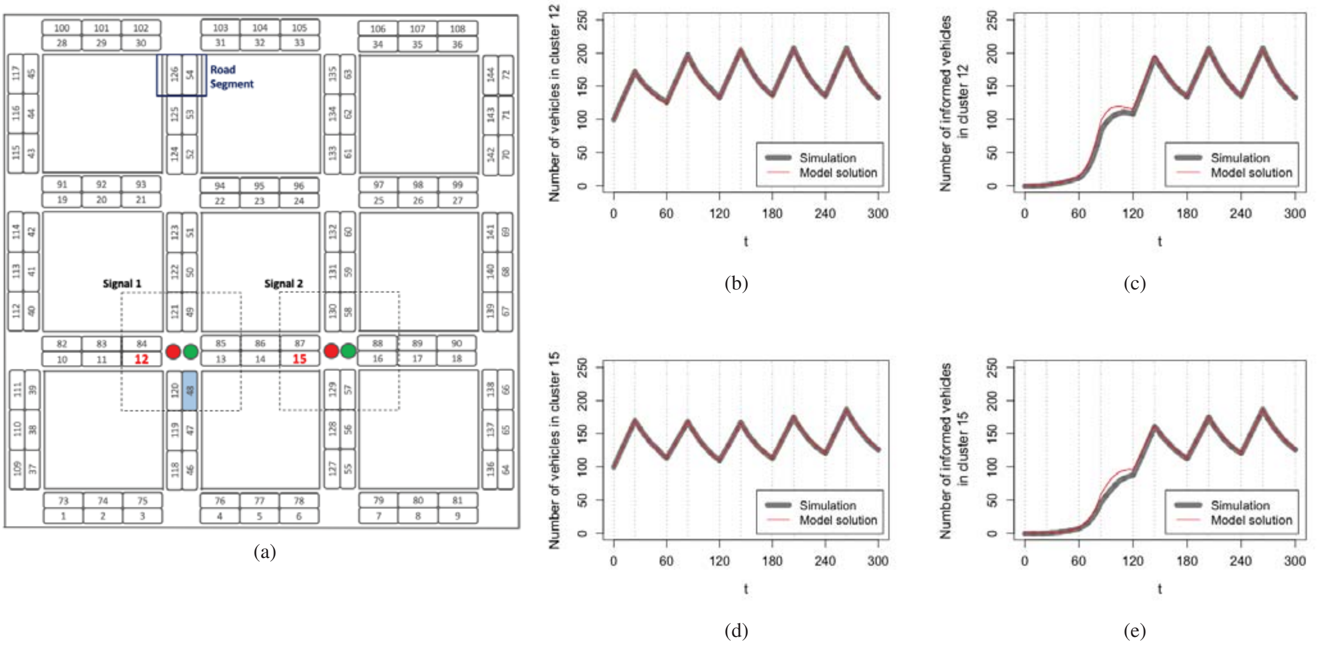


Fig. 7. (a) Clustered road topology with traffic signals at two intersections. (b) The number of vehicles in cluster 12 over time. (c) The number of informed vehicles in cluster 12 over time. (d) The number of vehicles in cluster 15 over time. (e) The number of informed vehicles in cluster 15 over time. The numbers of informed vehicles in each cluster are well approximated by model solutions. For (b), (c), (d), and (e), the gray lines represent the averages of 100 simulation runs, and the red lines represent the model solutions. The traffic signals have cycle lengths of 60 s, and 40% of the cycle time is spent in phase 1. Also, the signal 1 and 2 are synchronized.

$O(K)$ time. Thus the mobility rates involved in the differential equation for a cluster can be obtained in $O(K)$ time. The total complexity becomes $O(JK)$ because the total number of differential equations are linear to the total number of J clusters.

VI. Future Research

The mathematical guarantees for our model holds as the number of vehicles approach infinity and the sojourn times of vehicles in various clusters, intervals between communication of V2V messages are exponentially distributed. An interesting direction for future research is to obtain similar or even somewhat weaker mathematical guarantees, when the above assumptions do not hold. The sojourn times of the vehicles in individual clusters will not in general be exponentially distributed owing to the vehicular queues that build up when for example the vehicles stop at red lights or due to heavy traffic congestion or slow down due to the latter. One idea for generalization of the mathematical guarantees is to choose vehicular mobility rates that appropriately depend on the congestion levels in the clusters, and mathematically prove that the average behavior is captured.

References

[1] J. Kim, S. Sarkar, S. S. Venkatesh, M. S. Ryerson, and D. Starobinski, "An epidemiological diffusion framework for vehicular messaging in general transportation networks," Transportation Research Part B: Methodological, vol. 131, pp. 160–190, 2020.

[2] M. Ferreira, R. Fernandes, H. Conceição, W. Viriyasitavat, and O. K. Tonguz, "Self-organized traffic control," in Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking, 2010, pp. 85–90.

[3] New York City Department of Transportation (NYCDOT). Infrastructure: Traffic Signals. (Accessed Jan 21, 2020). [Online]. Available: https://www1.nyc.gov/html/dot/html/infrastructure/signals.shtml

[4] U.S. Department of Transportation. Next Generation Simulation (NGSIM) vehicle trajectories and supporting data. (Accessed June 10, 2018). [Online]. Available: https://data.transportation.gov/Automobiles/Next-Generation-Simulation-NGSIM-Vehicle-Trajector/8ect-6jqj

[5] X. Liang, X. Du, G. Wang, and Z. Han, "A deep reinforcement learning network for traffic light cycle control," IEEE Transactions on Vehicular Technology, vol. 68, no. 2, pp. 1243–1253, 2019.

[6] T. Tan, F. Bao, Y. Deng, A. Jin, Q. Dai, and J. Wang, "Cooperative deep reinforcement learning for large-scale traffic grid signal control," IEEE transactions on cybernetics, 2019.

[7] J. Kim, R. Saraogi, S. Sarkar, and S. S. Venkatesh. Technical report of "modeling the impact of traffic signals on v2v information flow". [Online]. Available: https://www.dropbox.com/s/tx8tjvn7kt7dzgi/Technical_Report_VTC.pdf?dl=0