Vehicle Mobility and Communication Channel Models for Realistic and Efficient Highway VANET Simulation

Nabeel Akhtar, *Member, IEEE*, Sinem Coleri Ergen, *Member, IEEE*, and Oznur Ozkasap, *Member, IEEE*

***Abstract*—Developing real-time safety and nonsafety applications for vehicular ad hoc networks (VANETs) requires understanding of the dynamics of the network topology characteristics since these dynamics determine both the performance of routing protocols and the feasibility of an application over VANETs. Using various key metrics of interest, including node degree, neighbor distribution, number of clusters, and link duration, we provide a realistic analysis of the VANET topology characteristics over time and space for a highway scenario. In this analysis, we integrate real-world road topology and real-time data extracted from the freeway Performance Measurement System (PeMS) database into a microscopic mobility model to generate realistic traffic flows along the highway. Moreover, we use a more realistic, recently proposed, obstacle-based channel model and compare the performance of this sophisticated model to the most commonly used and more simplistic channel models, including unit disk and lognormal shadowing models. Our investigation on the key metrics reveals that both lognormal and unit disk models fail to provide realistic VANET topology characteristics. We therefore propose a matching mechanism to tune the parameters of the lognormal model according to the vehicle density and a correlation model to take into account the evolution of the link characteristics over time. The proposed method has been demonstrated to provide a good match with the computationally expensive and difficult-to-implement obstacle-based model. The parameters of the proposed model have been validated to depend only on the vehicle traffic density based on the real data of four different highways in California.**

***Index Terms*—Channel model, mobility model, simulation, vehicular ad hoc networks (VANETs).**

I. Introduction

A VEHICULAR Ad Hoc Network (VANET) is a promising intelligent transportation system technology that offers many applications such as safety message dissemination [2]–[4], dynamic route planning [5], content distribution, gaming, and entertainment [6]. The majority of the VANET research effort on protocol design has relied on simulations due to the prohibitive cost of deploying real-world test-beds. Building a realistic simulation environment for VANETs is therefore essential in judging the performance of the protocols proposed at various layers.

A VANET simulation environment should be realistic, requiring an accurate representation of the vehicular mobility and signal propagation among the vehicles, and efficient, necessitating a reasonable amount of simulation time. Realistic representation of the vehicle mobility requires using real-world road topology, accurate microscopic mobility modeling, and real database traffic demand modeling, whereas a realistic representation of the signal propagation among the vehicles requires reproducing the actual physical radio propagation process for a given environment. On the other hand, an efficient representation of the vehicle mobility and signal propagation model requires analyzing both the closeness to the realistic representations in terms of the key metrics, summarizing the dynamics of the VANET topology in time and space, and the runtime of the simulations. As summarized in Table I, the literature on VANET topology characteristics focuses on realistic channel models tested on simplistic vehicle mobility models [7], [8], realistic mobility models without considering realistic signal propagation models [9], [11], [12], [14], or simplistic models for both vehicle mobility and communication channel [16]–[22]. Moreover, these models are often compared using classical metrics such as node degree, number of clusters, link duration and quality [9], [16]–[18], or related metrics such as packet loss probability and connectivity probability [7], [8], [14], [19], [20], [22]. They include only a small subset of important network-wide metrics summarizing the state of the network, such as closeness centrality, measuring how long it takes for the information to spread in a network, and a clustering coefficient, giving information about the degree to which nodes tend to cluster and the size of the largest group of connected vehicles in the network [11], [12], [17].

The goal of this paper is to analyze VANET topology characteristics on a large highway section by integrating realistic microscopic mobility traces generated using real-world road topology and real database traffic demand with realistic channel models, taking into account the effect of vehicles on the received signal power. We compare the performance of this realistic scenario to the most commonly used and more

Manuscript received June 28, 2013; revised November 15, 2013 and March 15, 2014; accepted March 20, 2014. Date of publication April 21, 2014; date of current version January 13, 2015. This work was supported by Turk Telekom under Grant 11315-07. The work of S. C. Ergen was supported by Bilim Akademisi-The Science Academy, Turkey, under the BAGEP program. This paper was presented in part at the IEEE Wireless Communications and Networking Conference, Shanghai, China, April 2013. The review of this paper was coordinated by Prof. Y.-B. Lin.

N. Akhtar was with the Department of Computer Engineering, Koc University, 34450 Istanbul, Turkey. He is now with the Department of Computer Science, Boston University, MA 02215 USA (e-mail: nabeel@bu.edu).

S. C. Ergen is with the Department of Electrical and Electronics Engineering, Koc University, 34450 Istanbul, Turkey (e-mail: sergen@ku.edu.tr).

O. Ozkasap is with the Department of Computer Engineering, Koc University, 34450 Istanbul, Turkey (e-mail: oozkasap@ku.edu.tr).

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/TVT.2014.2319107

TABLE I
RELATED WORK ON TOPOLOGY CHARACTERISTICS IN VANETS

Ref	Topology	Channel Model	Vehicle Mobility	Database for Vehicle Mobility	Performance Criteria
[7]	Urban	Ray-tracing-derived models	Intelligent Driver Model on Manhattan-grid	No database used	Packet loss ratio, Average end-to-end delay
[8]	Urban	Two ray ground	Chain topology with nodes arranged at a regular interval	No database used	Packet loss distribution
[9]	Urban	Unit disc	Microscopic road traffic	TAPASCologne Dataset by German Aerospace Institute of Transportation Systems [10]	No. of clusters, Node degree
[11]	Urban	Unit disc	Microscopic road traffic	TAPASCologne Dataset by German Aerospace Institute of Transportation Systems [10]	Node degree, betweenness centrality, No. of clusters, Size of the largest cluster
[12]	Urban	Unit disc	Multi-agent microscopic traffic	Zurich traffic data used [13]	Node degree, Link duration, Diameter, Centrality, No. of clusters, Clustering coefficient
[14]	Urban, Highway, Suburban	Unit Disk, Two ray Model	Microscopic road traffic	DRIVE-IN Project [15]	Packet delivery rate, delay
[16]	Urban	Unit disc with LOS and NLOS	Cellular Automata based traffic mobility model on Manhattan-grid	No database used	Link duration, Link re-healing time
[17]	Mostly Urban	Unit disc	Constant Speed Motion, Manhattan, Fluid Traffic Motion, Intelligent Driver Model	No database used	Link duration, Node degree, No. of cluster, Cluster coefficient, Vehicular density distribution
[18]	Urban	Unit disc, Quasi-Unit disc, Log-normal shadowing	Microscopic road traffic	Map of city of Porto, no database for real traffic	Node degree, Link duration
[19]	Urban	Unit disc, Rayleigh Fading, Ricean Fading	Microscopic multi-model traffic	Map of Berlin and Frankfurt, no database for real traffic	Packet reception probability
[20]	Highway	Unit disc, Log-normal shadowing	Randomly generated	No database used	Access probability, Connectivity probability
[21]	Urban	Unit disc	Manhattan Model	No database used	Number of clusters
[22]	Urban	Log-normal shadowing, Multi-path fading	Microscopic road traffic	No database used	Packet delivery ratio

simplistic channel models using various metrics of interest. The original contributions of the paper are listed as follows.

1) We incorporate real-world road topology and real-time data from the Performance Measurement System (PeMS) database [23] into the microscopic mobility model provided by Simulation of Urban Mobility (SUMO) [24]. The PeMS database allows modeling realistic traffic flows on the highway by adjusting the parameters of the SUMO simulator to match the real data. This is the first work to analyze VANET topology characteristics over a large-scale highway using an open-source database for measuring vehicle traffic and speed.
2) We incorporate a more realistic, recently proposed, obstacle-based channel model into the analysis of VANET topology characteristics and compare its performance to the most commonly used and more simplistic channel models, including unit disk and lognormal shadow fading models. This is the first work to analyze the effect of using the obstacle-based channel model on the VANET topology characteristics.
3) Since it is hard to integrate the obstacle-based model into the modern simulators due to its high complexity and computational cost, we propose a matching mechanism to tune the parameters of the lognormal model according to the vehicle density and a correlation model to take into account the evolution of the link characteristics over time. We demonstrate the performance of the proposed method and validate the dependence of the parameters of the proposed model on the vehicle traffic density over four different highways in California. This is the first work to propose a method to adjust the parameters of the lognormal model and introduce time correlation depending on the vehicle density for more realistic and efficient VANET simulation.
4) This is the first work to perform an extensive analysis of the VANET topology characteristics based on the realistic vehicle mobility and channel models. This analysis includes not only node degree, link duration, and number of clusters but also neighbor distribution, closeness centrality, size of the largest cluster, and clustering coefficient.

The rest of this paper is organized as follows. Section II discusses the related work on the vehicle mobility and communication channel models, as well as their usage in the analysis of VANET topology characteristics. Section III describes the generation of the realistic vehicle mobility using the PeMS database. Section IV defines different radio channel models, including the unit disk, lognormal shadowing, and obstacle-based model. Section V describes the performance metrics used in the comparison of different channel models. Section VI introduces the matching mechanism used for tuning the parameters of the lognormal shadowing model. Simulation results and the validation of the proposed matching mechanism are presented in Sections VII and VIII, respectively. Finally, concluding remarks and future work are given in Section IX.

II. RELATED WORK

A. Vehicle Mobility Models

Vehicular mobility simulators have been growing their complexity and features over time, encompassing realistic road topologies and microscopic vehicular models, where each vehicle is represented as a separate entity and the behavior of

vehicles depends on the neighboring vehicles [25]. SUMO [24], VISSIM [26], Development of Inter-Vehicular Reliable Telematics (DIVERT) [27], and Transportation Analysis Simulation System (TRANSIMS) [28] are examples of such simulators.

The simulations are usually performed on small portions of a road with user-generated traffic flows in these simulators and only recently have been extended to larger areas while incorporating a realistic model for the macroscopic mobility of the vehicles where the traffic flows are determined based on the real data. The real data used for this purpose in the recent studies on the analysis of the VANET topology characteristics in urban areas include mobility traces gathered through various measurement campaigns [9], [11] and statistics performed by the urban planning and traffic engineering communities [12], [14]. None of these studies, however, analyze the VANET topology characteristics on a large-scale highway considering the real database traffic demand of the vehicles.

B. Communication Channel Models

Realistic representation of the signal propagation among vehicles requires reproducing the actual physical radio propagation process for a given environment based on the ray-tracing method [7], [29]. The ray-tracing approach generates the complex impulse response of the channel by determining possible paths or rays from the transmitter to the receiver according to the rules of geometrical optics. Such a model, however, is impractical since it requires a detailed description of the site-specific propagation environment.

Stochastic models, on the other hand, determine the physical parameters of the vehicular channel in a completely stochastic manner without presuming any underlying geometry [30]. The distance-dependent path-loss, large-scale, and small-scale fading distributions are the parameters to be estimated in these stochastic models as a result of extensive measurement campaigns. The path loss represents the local average received signal power relative to the transmit power as a function of the distance between the transmitter and the receiver. The path-loss exponent of $n = 1.8$–2.7 was observed on highways in [31]–[34]. Large-scale fading models the effect of the surrounding obstacles on the mean signal attenuation at a given distance. The surrounding obstacles may be mobile (e.g., other vehicles) or static (e.g., buildings in urban environments). Most of the channel modeling activities aims at averaging the additional attenuation due to these obstacles, resulting in a lognormal distribution around the mean received signal power in urban areas [34], [35] and highways [35]. Although some of these models estimate different values for the variance of this large-scale fading distribution at high and low traffic densities [35], only recently a mechanism for incorporating the effect of the vehicles and static obstacles on the received signal power in highways has been proposed in [36] and [37], respectively. Finally, small-scale fading models the effect of the reception of multiple replicas of the transmitted signal at the receiver. Various distributions have been proposed for small-scale fading, including Rice [38], Nakagami [34], and Weibull [35], [39] distributions.

Although the signal propagation has a great impact on the performance of the communication protocols, most of the recent work on the analysis of VANET topology characteristics use a unit disk as the signal propagation model, where the vehicles can communicate with each other if they are within a threshold distance and cannot communicate otherwise [9], [11], [12], [14], [16]–[21]. Although recently such analysis employs more sophisticated stochastic signal propagation models including both large-scale fading [18], [20], [22] and small-scale fading [19], [22], none of these models incorporate the effect of the vehicles on the signal propagation.

III. Vehicle Mobility Model

Realistic representation of the vehicle mobility requires using accurate microscopic mobility modeling, real-world road topology, and real database traffic demand modeling. The input and parameters of the microscopic mobility simulator are determined based on the real traffic flow and speed values measured by the road sensors deployed along the highway, as detailed next.

A. Microscopic Mobility Modeling

SUMO [24] is used to simulate the microscopic mobility of the vehicles. SUMO, which is generated by the German Aerospace Center, is an open-source space-continuous discrete-time traffic simulator capable of modeling the behavior of individual drivers. The path of each driver is determined based on the origin/destination matrix provided as an input to the simulator. The movement of each driver is implemented using the surrounding vehicles via Krauss' car-following model, which regulates its acceleration, and Krajzewicz's lane-changing model, which regulates its overtaking decisions [40]. The parameters of the simulator that determine the driver's acceleration and overtaking decisions include the distance to the leading vehicle, the traveling speed, the acceleration and deceleration profiles, and dimension of the vehicles.

B. Traffic Demand Modeling

PeMS collects historical and real-time data from highways in the state of California with the goal of providing a comprehensive assessment of highway performance [23]. PeMS was developed by the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, in cooperation with the California Department of Transportation, California Partners for Advanced Transit and Highways, and Berkeley Transportation Systems. The flow and speed data are collected in real time from more than 25 000 individual road sensors located over all major metropolitan areas in the state of California. The sampling period of the flow and speed data ranges from 30 s to 5 min. We used road I-880S in Alameda County, Bay Area, California, for our simulation. Fig. 1 shows the road sensors located on I-880S.

C. Realistic Mobility Generation

The first step in generating the realistic mobility model is to determine the input of SUMO using the real database vehicular

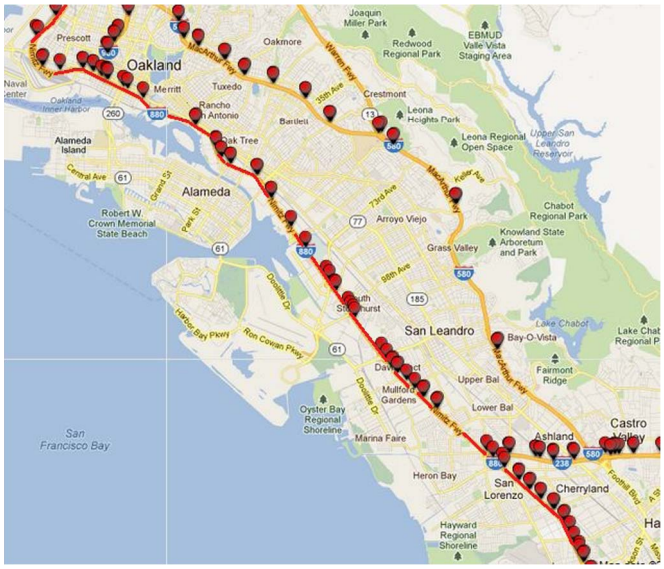

Fig. 1. Road sensors located on I-880S in Alameda County, Bay Area, California.

traffic flows over the road. The input of SUMO, including the number of vehicles injected at each entry of the highway (the starting point of the vehicle) and the probability that each vehicle leaves the highway from the exits (destination of vehicles), is determined such that the expected number of vehicles passing through each road sensor location in the simulation closely matches the flow measured at that sensor on the actual highway. However, matching the traffic flow in the simulation to that of the PeMS database does not guarantee that the average speed of the vehicles in the simulation also matches the speed measured through PeMS. Therefore, the second step in generating realistic mobility model is to determine the parameters of SUMO such that the average speeds of the vehicles determined by the simulation and PeMS agree with each other. The parameters of SUMO adjusted for this purpose include the distance to the leading vehicle, the initial speed, and the acceleration and deceleration profiles.

Figs. 2 and 3 show the standard deviation bars depicted around the mean of the flow and speed of the vehicles recorded from the simulation and PeMS database, respectively. The data from 419 road sensors on highway I880-S, as shown in Fig. 1, are extracted for both high traffic density, i.e., at 18:00, and low traffic density, i.e., at 01:00. Road sensors are, on average, 1.21 km apart from each other on the road. Vehicles are randomly and uniformly assigned to the lanes based on the PeMS measurement results aggregated over all lanes. As demonstrated in the figures, once the system stabilizes at around the tenth minute, both the average flow and speed from the simulation and PeMS database are consistent with each other.

Fig. 4 shows the cumulative distribution function (cdf) of the percentage error between the values extracted from the PeMS database and the values obtained from the simulation for the flow and speed of the vehicles at different vehicle traffic densities. The values are obtained over all the sensors at different time instants after the system stabilizes. We observe that for 80% of the values of the flow and speed, the error is less than 5% at low-density traffic, and less than 12% at

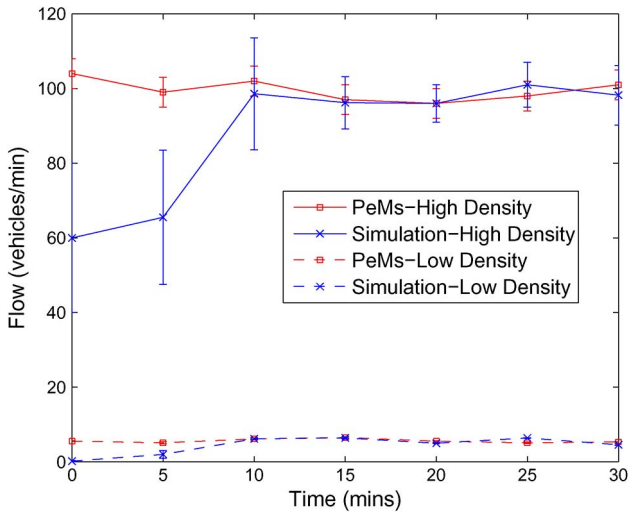

Fig. 2. Standard deviation bars depicted around the mean of the flow of the vehicles extracted from the PeMS database and obtained from the simulation at low and high vehicle traffic densities.

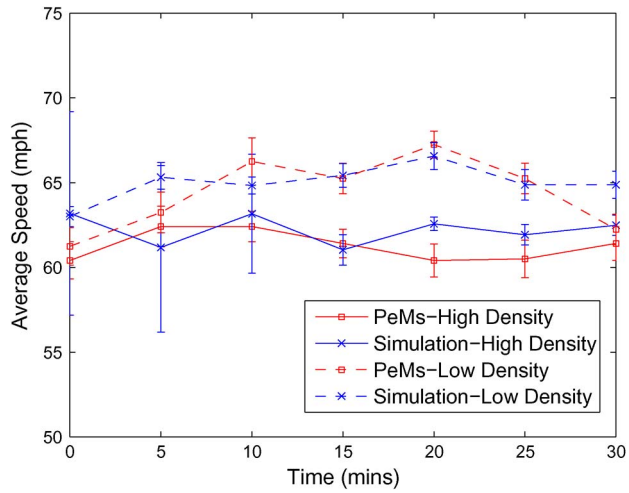

Fig. 3. Standard deviation bars depicted around the mean of the speed of the vehicles extracted from the PeMS database and obtained from the simulation at low and high vehicle traffic densities.

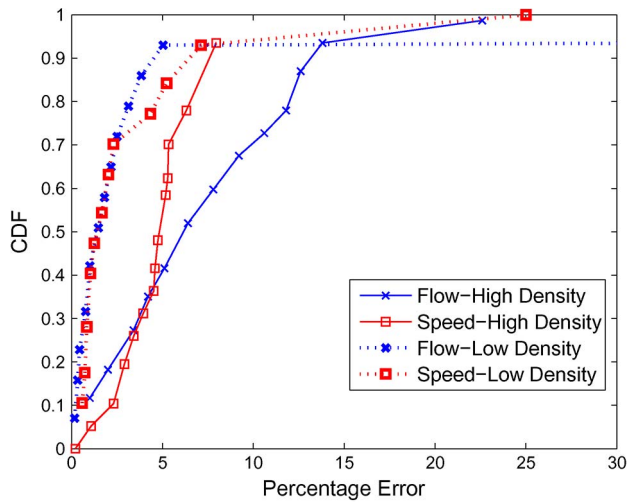

Fig. 4. CDF of the percentage error between the values extracted from the PeMS database and the values obtained from the simulation for the flow and speed of the vehicles at low and high vehicle traffic densities.

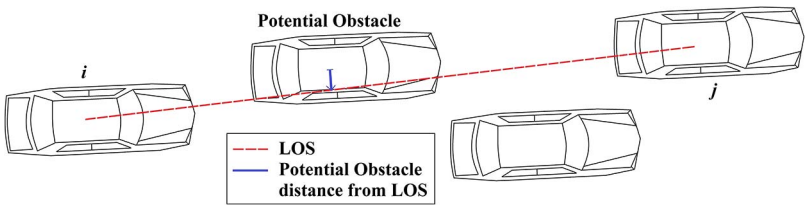

Fig. 5. Determining the vehicles potentially obstructing the LOS between vehicles i and j.

high-density traffic. This demonstrates that the mobility of the vehicles in the simulation closely matches that in the PeMS database.

IV. Vehicular Channel Models

Here, we will first describe simplistic channel models, including unit disk and lognormal shadowing models, that are commonly used in the analysis of the VANET topology characteristics. We will then describe a recently proposed and more realistic obstacle-based channel model that incorporates the effect of the moving obstacles (i.e., vehicles) on the received signal power due to their dominating influence, as shown in [36]. The obstacle-based model has not been used before in the VANET topology analysis.

A. Unit Disk Model

In the unit disk model, the vehicles can communicate with each other if they are within a threshold distance and cannot communicate otherwise. This model is widely used in the analysis of the VANET topology characteristics due to its simplicity [9], [12], [16]–[21]. However, the sharp cutoff at the threshold distance not only fails to capture the random noise that can make even nearby nodes unreachable but, in addition, does not take into account the effect of the obstacles on the received signal strength.

B. Classical Lognormal Shadowing Model

In the classical lognormal shadowing model, rather than calculating the additional attenuation due to each obstacle between the transmitter and the receiver, the probabilistic distribution of the additional attenuation is modeled with a lognormal probability density function, resulting in the following formulation for the received signal power [34], [35]:

$$P_{\mathrm{rx}}(d) = P_0 - 10n \log_{10} \frac{d}{d_0} + N \tag{1}$$

where d is the distance between the transmitter and the receiver, d_0 is the reference distance, $P_{\mathrm{rx}}(d)$ is the received signal power at distance d (in dBm), P_0 is the received signal power at the reference distance d_0 (in dBm), n is the path-loss exponent, and N is a zero-mean Gaussian random variable with variance σ^2. A vehicle can communicate with another vehicle if P_{rx} is greater than a certain threshold value [20]. Note that the lognormal shadowing model reduces to the unit disk model if $\sigma = 0$. The parameters of the lognormal model are chosen such that the mean transmission range is equal to the threshold distance in the unit disk model to have a fair comparison. The parameters n and σ of the model are chosen based on the channel measurement results reported in [31]–[35]: $n = 2.5$ and $\sigma = 5.5$ dB.

C. Obstacle-Based Channel Model

In the obstacle-based channel models, algorithms to incorporate the effect of the surrounding obstacles such as other vehicles, walls, and buildings on the received signal strength have been proposed [35], [36] rather than modeling the average additional attenuation due to these obstacles by a stochastic large-scale fading model. Usually, there are a few buildings around the highway, mostly far from the vehicles. That is why, in this study, we only consider the impact of the surrounding vehicles as obstacles. Since the additional obstacles can only further reduce the probability of the line of sight (LOS) between the transmitter and receiver vehicles, this approach gives a best case analysis for the probability of LOS, as stated in [36].

The algorithm proposed and validated in [36] is implemented for calculating the additional attenuation due to the vehicles. This algorithm consists of three main parts: First, the vehicles potentially obstructing the LOS between the transmitter vehicle i and the receiver vehicle j are determined ($getPotentialObs(i, j)$): If the distance from the center of the vehicle to the LOS line between vehicles i and j is less than half the width of the vehicle, the vehicle is considered as a potential obstacle, as shown in Fig. 5 (see line 1 of Algorithm 1).

Algorithm 1 Obstacle-Based Model: Calculation of the Additional Attenuation Between Vehicles i and j due to Surrounding Vehicles as Obstacles

```
1: [PotentialObs] = getPotentialObs(i, j)
2: if size([PotentialObs]) ≠ 0 then
3:     [ObsVeh] = getLOSobs([PotentialObs])
4:     if size([ObsVeh]) ≠ 0 then
5:         addAttenuation = calAttenuation([ObsVeh])
6:     else
7:         addAttenuation = 0
8:     end if
9: else
10:    addAttenuation = 0
11: end if
```

Second, the vehicles that obstruct the LOS between vehicles i and j are chosen from the set of the potential obstructing vehicles identified in the previous step ($getLOSobs$

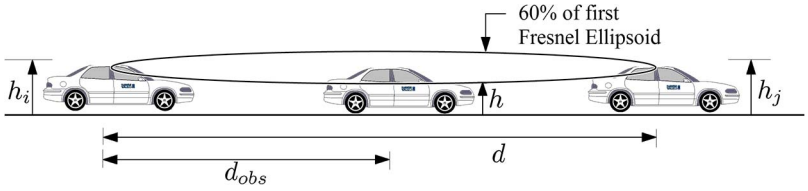

Fig. 6. Determining the vehicles that obstruct the LOS between vehicles i and j (for simplicity, vehicle antenna heights h_a are not shown).

($[PotentialObs]$)): From the electromagnetic wave propagation perspective, the LOS is not guaranteed with the existence of the visual sight line between the transmitter and the receiver. Any vehicle that obstructs the Fresnel ellipsoid might affect the transmitted signal. The effective height of the LOS line that connects vehicles i and j at the potential obstacle vehicle location when we use the first Fresnel ellipsoid is given by

$$h = (h_j - h_i)\frac{d_{\text{obs}}}{d} + h_i - 0.6r_f + h_a \tag{2}$$

where h_i and h_j are the heights of the transmitter vehicle i and the receiver vehicle j, respectively; d_{obs} is the distance between the transmitter and the obstacle; d is the distance between the transmitter and the receiver; h_a is the height of the vehicle antennas; and r_f is the radius for the first Fresnel zone ellipsoid, which is given by

$$r_f = \sqrt{\frac{\lambda d_{\text{obs}}(d - d_{\text{obs}})}{d}} \tag{3}$$

with λ denoting the wavelength. Fig. 6 shows these parameters. If the height of each potentially obstructing vehicle is known beforehand, the vehicle will obstruct the LOS between the transmitter and the receiver if h is greater than its height. Based on the assumption that the vehicle heights follow a normal distribution, as also assumed in [36], the probability of the LOS for the link between vehicles i and j is calculated as

$$\Pr(\text{LOS}|h_i, h_j) = 1 - Q\left(\frac{h - \mu_h}{\sigma_h}\right) \tag{4}$$

where μ_h and σ_h are the mean and standard deviation of the height of the obstacle vehicle (line 3 of Algorithm 1).

Third, the additional attenuation in the received signal power is calculated for the LOS obstructing vehicles determined in the previous step ($calAttenuation([ObsVehicles])$). The existing models to calculate the attenuation are empirical and vary from optimistic [41] to pessimistic approximations [42], [43]. To calculate the additional attenuation, we used the ITU-R method based on the multiple knife-edge model [44], as suggested in [36]. In this model, a complete profile is created for all the LOS obstructing vehicles, and the signal attenuation is calculated based on the vehicle height, distance from the transmitting vehicle, wavelength of electromagnetic waves, and position of the vehicles (see line 5 of Algorithm 1).

V. Performance Metrics

The performance metrics are used in the comparison of different signal propagation models. For the formal definition of these metrics, we represent the vehicular network topology at time t by a graph $G(t) = (V(t), E(t))$, where $V(t)$ is the set of vehicles and $E(t) \subset V(t) \times V(t)$ are the directed edges representing the wireless communication links between the vehicles.

A. Neighbor Distance Distribution

Neighbor distance distribution is defined as the distribution of the distance of the neighbors of the vehicles in the network. Let us denote the set of neighbors at distance p away from vehicle i at time t by $N_i(t,p)$ such that $N_i(t,p) = \{j|(i,j) \in E(t), d_{ij} = p\}$, where d_{ij} is the distance between vehicles i and j. The neighbor distance distribution as a function of the distance p denoted by $f(p)$ is then formulated as $f(p) = (1/T)\sum_{t=1}^{T}(1/|V(t)|)\sum_{i=1}^{|V(t)|}|N_i(t,p)|$, where $|V(t)|$ is the total number of the vehicles in the network at time t, and T is the total simulation time. Neighbor distance distribution measures the distribution of the communicating vehicles over space.

B. Node Degree

The node degree of a vehicle is defined as the number of the neighboring vehicles it can communicate with. Let us denote the set of neighbors of vehicle i at time t by $N_i(t)$ such that $N_i(t) = \{j|(i,j) \in E(t)\}$. Note that $N_i(t) = \int_0^\infty N_i(t,p)dp$. The degree of vehicle i is then formulated as $deg_i(t) = |N_i(t)|$. Node degree measures the density of the network from the physical connectivity point of view.

C. Link Duration

Link duration is defined as the time span between the instants at which the communication link between two vehicles is established and lost. Let us denote the times when the link from vehicle i to vehicle j is established and broken by t_0 and t_f, respectively, such that $(i,j) \notin E(t_0 - \epsilon)$, $(i,j) \notin E(t_f + \epsilon)$ for arbitrarily small $\epsilon > 0$ and $(i,j) \in E(\tau)$, $\forall \tau \in [t_0, t_f]$. Then, the duration of the link from vehicle i to vehicle j, which is denoted by l_{ij}, is formulated as $l_{ij} = t_f - t_0$. Link duration measures the stability of a connection over time.

D. Number of Clusters

Number of clusters is defined as the number of coexistent nonconnected groups of vehicles at a given instant. We define cluster as a connected group of vehicles within which there exists a path between any pair of nodes. Formally, let us denote the existence of a path from vehicle i to vehicle j at time t

by the binary variable $p_{ij}(t)$ such that $p_{ij}(t)$ takes value 1 if $(i,j) \in E(t)$ or there exists k for which $(i,k) \in E(t)$ and $p_{kj}(t) = 1$, and value 0 otherwise. Let us also define the cluster in which vehicle i is located at time t as $C_i(t) = i \cup \{j | p_{ij}(t) = 1, p_{ji}(t) = 1\}$. The set of unique clusters in the network at time t is formulated as

$$C(t) = \{C_i(t) | C_i(t) \cap C_k(t) = \emptyset,\ \forall k \in [1, i-1],\ \forall i \in V(t)\}. \tag{5}$$

The number of clusters denoted by $c(t)$ is then equal to $|C(t)|$. Number of clusters measures the degree of fragmentation in the network in terms of the number of mutually isolated groups of vehicles.

E. Size of the Largest Cluster

Size of the largest cluster is defined as the number of vehicles in the largest cluster of the vehicular network. Formally, the size of the largest cluster, which is denoted by $c_{\max}(t)$, is formulated as $c_{\max}(t) = \max_{i \in V(t)} |C_i(t)|$.

F. Clustering Coefficient

Clustering coefficient is defined as the ratio of the number of the links within a cluster to the maximum number of the links that could exist within a cluster. Let us denote the set of the links within cluster $C_i(t)$ by $E_{C_i}(t) = \{(i,j) | (i,j) \in E(t); i,j \in C_i(t)\}$. The clustering coefficient of the same cluster denoted by $k_{C_i}(t)$ is then formulated as

$$k_{C_i}(t) = \frac{|E_{C_i}(t)|}{|C_i(t)|\,(|C_i(t)| - 1)}. \tag{6}$$

Clustering coefficient measures the degree of connectivity of the vehicles within a cluster. Note that the clustering coefficient has a maximum value 1 if the cluster is a clique.

G. Closeness Centrality

Closeness centrality of a vehicle is defined as the inverse of the sum of the number of hops to all other nodes that are reachable from that vehicle. Formally, the closeness centrality of vehicle i is defined for every i such that there exists at least one reachable vehicle from vehicle i, i.e., $\{j | p_{ij}(t) = 1, j \neq i\} \neq \emptyset$, which is denoted by $CC_i(t)$ and formulated as

$$CC_i(t) = \frac{1}{\sum_{\{j | p_{ij}(t) = 1\}} h_{ij}} \tag{7}$$

where h_{ij} is the number of hops in the shortest path from vehicle i to vehicle j. Closeness centrality measures how long it will take information to spread from a given vehicle to other reachable vehicles in the network.

VI. Matched Lognormal Shadowing Model

The classical lognormal shadowing model is based on specifying the probabilistic distribution of the additional attenuation due to the vehicles instead of calculating the attenuation due to each vehicle acting as an obstacle separately. Since the decision for the existence of the communication link between every pair of vehicles requires specifying the value of a random variable only, independent of the number of the vehicles, the runtime of the simulations using this model is reasonable at $O(|V(t)|^2)$ at time t. However, the parameters of this model, including the path-loss exponent and the standard deviation of the Gaussian distribution, are fixed independent of the density of the surrounding vehicles leading to unrealistic simulations.

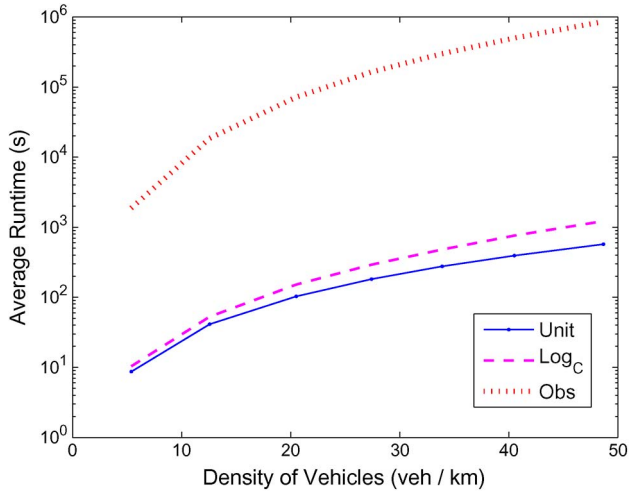

Fig. 7. Average runtime of the analysis of the performance metrics based on the 1800-s simulation of the scenario where the vehicles are generated using the real data at different traffic densities and communicating with a mean transmission range of 500 m under different channel models.

The obstacle-based channel model, on the other hand, incorporates the effect of each vehicle on the received signal strength separately. Since the decision for the existence of the communication link between every pair of vehicles requires determining the effect of every other vehicle on the received signal strength, the accurate representation of the channel comes at the cost of high runtime complexity at $O(|V(t)|^3)$ at time t. This prevents the integration of this model into the network simulators.

Fig. 7 shows the average runtime of the analysis of the performance metrics based on the 1800-s simulation of the scenario where the vehicles are generated using the real data at different traffic densities and communicating with a mean transmission range of 500 m under unit disk, classical lognormal, and obstacle-based channel models, which are denoted by $Unit$, Log_C, and Obs, respectively. We observe that the average runtime of the obstacle-based model is about 100 times more than that of the unit disk and classical lognormal models increasing at a much higher rate as a function of the vehicle traffic density.

Here, we propose a matching mechanism to tune the parameters of the lognormal model such that the performance metrics summarizing the link characteristics over space agree with those of the obstacle-based model (see Section VI-A). We also introduce a correlation model to take into account the evolution of the link characteristics over time and propose a mechanism to tune the parameters of this model to match the performance metrics, summarizing the time characteristics of the

links to those of the obstacle-based model (see Section VI-B). The resulting matched lognormal model provides performance close to that of the obstacle-based model at much lower computational cost and implementation complexity, allowing its integration into the network simulators. The reason for choosing a lognormal channel model in the matching mechanism is that a lognormal channel model is commonly used, as supported by the channel measurement results reported in [31]–[35], and allows adjusting the probabilistic behavior of the communication links over both space and time.

Algorithm 2 Matching Parameters of the Lognormal Shadowing Model

Input: $nValues$, $\sigma Values$
Output: n_m, σ_m
1: $NDObs = cdfND(ObsModel)$;
2: $Error_{\min} = \infty$;
3: **for all** $nValues$ **do**
4: **for all** $\sigma Values$ **do**
5: $NDLog = cdfND(LogModel, n, \sigma)$;
6: $Error = Diff(NDObs, NDLog)$;
7: **if** $Error < Error_{\min}$ **then**
8: $Error_{\min} = Error$;
9: $n_m = n$;
10: $\sigma_m = \sigma$;
11: **end if**
12: **end for**
13: **end for**

A. Matching Parameters of the Lognormal Model

The parameters of the lognormal model that need to be matched for the spatial evolution of the link characteristics include the path-loss exponent and the standard deviation of the Gaussian random variable.

Our matching algorithm to tune these path-loss exponent and standard deviation parameters is given in Algorithm 2. The inputs of the algorithm are the set of possible values for the path-loss exponent and standard deviation denoted by $nValues$ and $\sigma Values$, respectively. $nValues$ and $\sigma Values$ are chosen in the range of [1.5–5] and [5–20], respectively, based on the channel measurement results reported in [31]–[35]. The outputs of the algorithm are the values for the path-loss exponent and standard deviation that provide the best match to the obstacle-based model, which are denoted by n_m and σ_m, respectively.

The algorithm starts by extracting the cdf of the node degree metric for the obstacle-based model by using function $cdfND$ with the parameter $ObsModel$ representing the obstacle-based model and storing the resulting cdf in the variable $NDObs$ and initializing the minimum error to infinity by using variable $Error_{\min}$ (lines 1 and 2). The algorithm then computes the cdf of the node degree metric of the lognormal model, which is represented by $LogModel$, and storing the resulting cdf in the variable $NDLog$ for every possible value of the path-loss exponent and standard deviation, stored in the variables n and σ, respectively, in each iteration (lines 3–5). The difference between the cdf of the obstacle-based and lognormal models is then calculated by using Kolmogorov–Smirnov statistic, which is defined as $sup_x|NDObs(x) - NDLog(x)|$ in the function called $Diff$ (line 6). The values of the path-loss exponent and standard deviation that provide the minimum difference are then selected to provide the best match to the obstacle-based model (lines 7–11). If the value of $Error_{\min}$ is not low enough at the end of the algorithm, the matching process can be repeated with an extended set of values for $nValues$ and $\sigma Values$ at the input. However, the proposed initial ranges for the path-loss exponent and standard deviation parameters always provided a low $Error_{\min}$ value in the extensive simulations given in Section VII.

The reason for choosing the cdf of the node degree in the matching algorithm is that node degree is a measure of the spatial distribution of the nodes in the network. The aforementioned process is repeated for the neighbor distance distribution metric to validate the values n_m and σ_m. To obtain common values for n_m and σ_m from the matching algorithm of both the node degree and neighbor distance distribution metrics, we have determined all the values of n and σ that achieve $Error$ within the 5% of the $Error_{\min}$ of both metrics. Then, out of all the n and σ values that overlap, the ones corresponding to the minimum error of the node degree metric are assigned to n_m and σ_m, respectively. If there are no overlapping n and σ values, then the percentage of getting closer to $Error_{\min}$ can be increased. However, such overlapping values were always found in the extensive simulations given in Section VII. The matching of the remaining performance metrics are justified in Section VII. Similarly, if it is important for an application to match another metric that is a measure of the spatial distribution of the nodes more closely than node degree, it can be used in the matching algorithm. However, the matching of the remaining performance metrics should again be justified in that case.

B. Matching Time Correlation

The instances of the Gaussian variable used in the classical lognormal model are independently calculated at each time step of the simulation resulting in zero correlation of the link characteristics over time. The obstacle-based model, on the other hand, provides the time correlation of the link characteristics implicitly due to the slow changes in the relative locations of the obstacles between the transmitter and receiver vehicles. We therefore extend the classical lognormal model to include a correlation model that takes into account the temporal evolution of the link characteristics.

We used the Gudmunson model with an exponential correlation function for this work [45]. This model has been previously used for spatially correlated processes [22], [46]. However, as described in [47], this model can be also used for time correlation. The model describes the correlation of the shadowing process at time difference Δt by

$$Corr(\Delta t) = \sigma^2 \cdot \exp(-\alpha \Delta t)$$

where σ is the standard deviation of the Gaussian variable at each time instant, and α is the correlation factor.

Algorithm 3 Matching Correlation Factor for the Lognormal Shadowing Model

Input: n_m, σ_m, $\alpha Values$
Output: α_m
1: $LDObs = cdfLD(ObsModel)$;
2: $Error_{\min} = \infty$;
3: **for all** $\alpha Values$ **do**
4: $LDLog = cdfLD(LogModel, n_m, \sigma_m, \alpha)$
5: $Error = Diff(LDObs, LDLog)$
6: **if** $Error < Error_{\min}$ **then**
7: $Error_{\min} = Error$;
8: $\alpha_m = \alpha$;
9: **end if**
10: **end for**

The matching algorithm to tune the value of the correlation factor is given in Algorithm 3. The inputs of the algorithm are the values of the path-loss exponent and standard deviation providing the best match with the spatial link characteristics of the obstacle-based model, i.e., n_m and σ_m, respectively, and the set of possible values for the correlation factor, which is denoted by $\alpha Values$. $\alpha Values$ are chosen in the range of [0–2] based on the correlation coefficient values used in [22] and [45]–[47]. The output of the algorithm is the value of the correlation factor that provides the best match with the temporal link characteristics of the obstacle-based model, which is denoted by α_m.

The algorithm starts by determining the cdf of the link duration metric for the obstacle-based model by using function $cdfLD$ and storing the resulting cdf in the variable $LDObs$ and initializing the minimum error to infinity (see lines 1 and 2). The reason for choosing the cdf of the link duration is that link duration is a measure of the stability of the links over time. The algorithm continues by computing the cdf of the link duration metric of the lognormal model with the matched parameters n_m and σ_m, and every possible value of the correlation factor stored in the variable α, and storing the resulting cdf in the variable $LDLog$ in each iteration (see lines 3 and 4). The difference between the cdf of the link duration of the lognormal and obstacle-based models is then calculated by using Kolmogorov–Smirnov statistics in the function $Diff$ (see line 5). The value of the correlation factor giving the minimum difference is selected as the best match value σ_m (see lines 6–9). If the value of $Error_{\min}$ is not low enough at the end of the algorithm, the matching process can be repeated with an extended set of values for $\alpha Values$ at the input. However, the proposed initial range for the correlation factor always provided a low $Error_{\min}$ value in the extensive simulations given in Section VII.

VII. Simulation Results

The goal of the simulations is to compare the effect of different channel models, including the unit disk, classical lognormal fading, obstacle-based, and matched lognormal channel models on the topology characteristics of VANETs located on a large-scale highway by comparing the node degree, neighbor distance distribution, link duration, closeness centrality, number of clusters, size of the largest cluster, and clustering coefficient metrics of the resulting communication graphs, as explained in Section V. Although we provide the cdf of the absolute values of all the metrics, we focus on the closeness of different channel models in terms of these metrics. A certain discrepancy between the obstacle-based model and one of the unit disk, classical lognormal fading, and matched lognormal channel models in some of the metrics may be tolerable for some vehicle applications.

In all the figures here, the unit disk, classical lognormal, matched lognormal, and obstacle-based channel models are denoted by $Unit$, Log_C, Log_M, and Obs, respectively, and R refers to the mean transmission range of the vehicles, which is achieved by adjusting the transmission power in all the channel models.

The topology of the VANET is obtained by using the accurate microscopic mobility modeling of SUMO while determining its input and parameters based on the PeMS database, as explained and validated in detail in Section III. The flow and speed data of 419 road sensors on highway I880-S, as shown in Fig. 1, at both high traffic density, i.e., 121 vehicles/km at 18:00, and low traffic density, i.e., 11 vehicles/km at 01:00, are used for this purpose. The performance metrics are extracted after the system stabilizes around the 10th minute, as shown in Figs. 2 and 3. The vehicle mobility output of SUMO is then input to MATLAB, where the channel models are implemented and the performance metrics are derived and plotted. The data sets can be found in [48].

Fig. 8(a) and (b) shows the neighbor distance distribution for low- and high-density networks, respectively. Note that the distance in the neighbor distance distribution is quantized into bins of 25 m in this figure. The matched lognormal model follows the obstacle-based model closely. The difference between the obstacle-based model and the commonly used unit disk and classical lognormal fading models, on the other hand, increases as the vehicle density increases and the transmission range becomes greater than 100 m. To elaborate on the effect of this different behavior on the performance metrics, we therefore plot the rest of the graphs at transmission ranges of 100 and 500 m for both low and high vehicle traffic densities.

Fig. 9(a) and (b) shows the cdf of the node degree metric for different channel models and transmission ranges in low- and high-density networks, respectively. The cdf of the node degree shows the total number of the vehicles each vehicle can communicate with, which is summed over all the distances but considering every vehicle separately over time. Neighbor distance distribution, on the other hand, differentiates the number of vehicles each vehicle can communicate with over distance but averages these values over time and all the vehicles in the network. Therefore, node degree and neighbor distance distribution together determine the neighborhood of each vehicle across distance and time. All channel models generate the node degree distribution very close to each other at a low transmission range for both low- and high-density networks. As

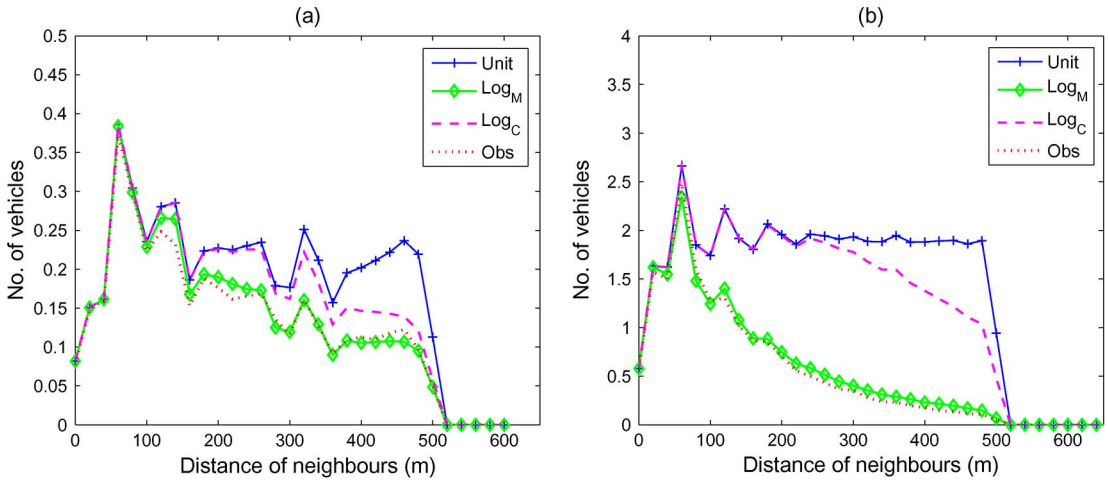

Fig. 8. Neighbor distance distribution for (a) low-density and (b) high-density networks at the 500-m transmission range.

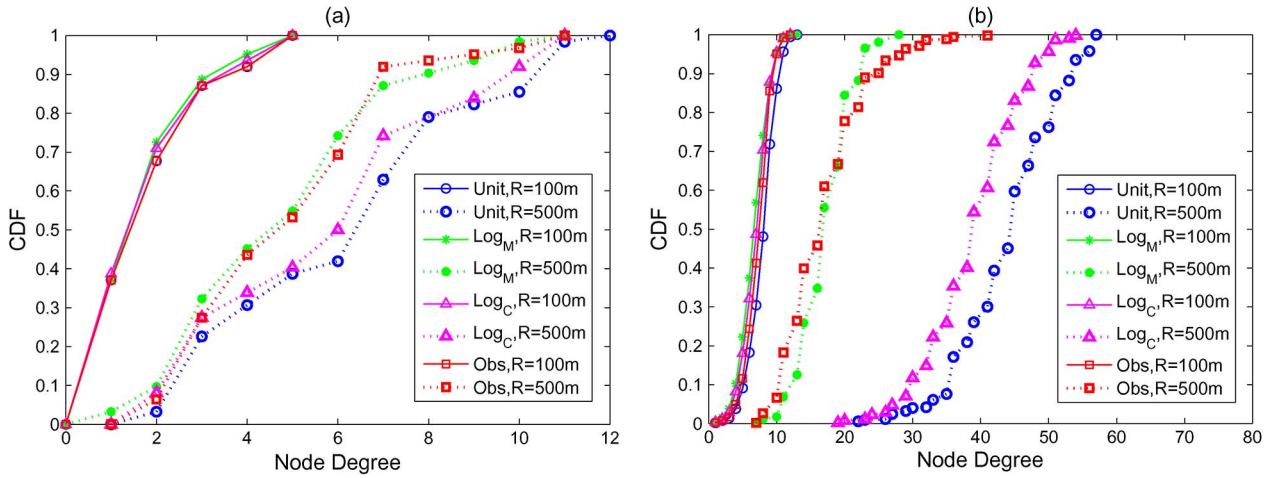

Fig. 9. CDF of the node degree metric for different channel models and transmission ranges in (a) low-density and (b) high-density networks.

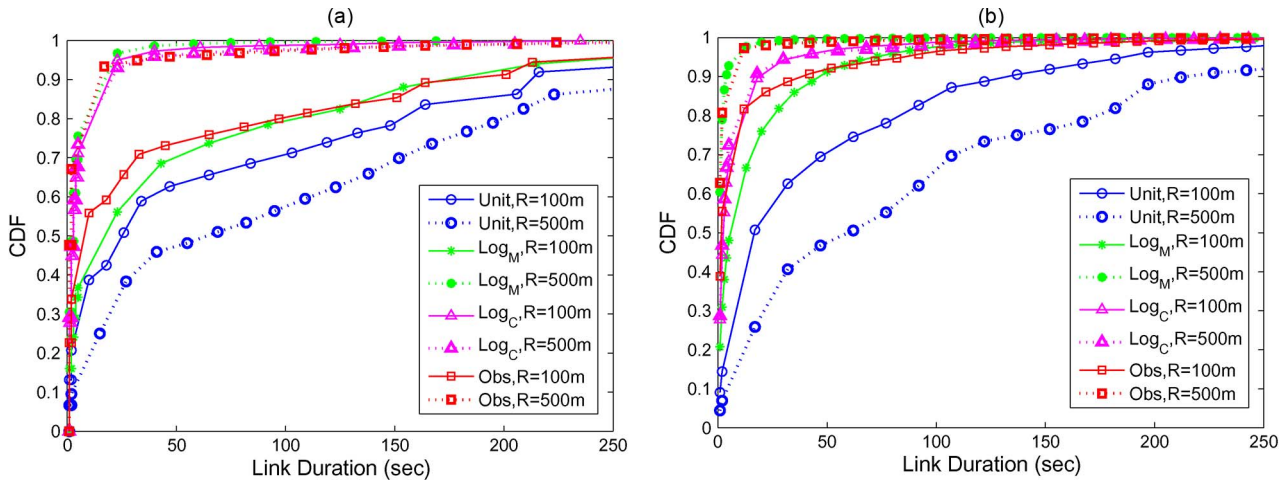

Fig. 10. CDF of the link duration metric for different channel models and transmission ranges in (a) low-density and (b) high-density networks.

the transmission range and the density of the network increase, the discrepancy between the obstacle-based model and the unit disk and classical lognormal fading models increases, as expected from the difference observed in the neighbor distance distribution. The matched lognormal model, on the other hand, still agrees with the obstacle-based model for all scenarios.

Fig. 10(a) and (b) shows the CDF of the link duration metric for different channel models and transmission ranges in low- and

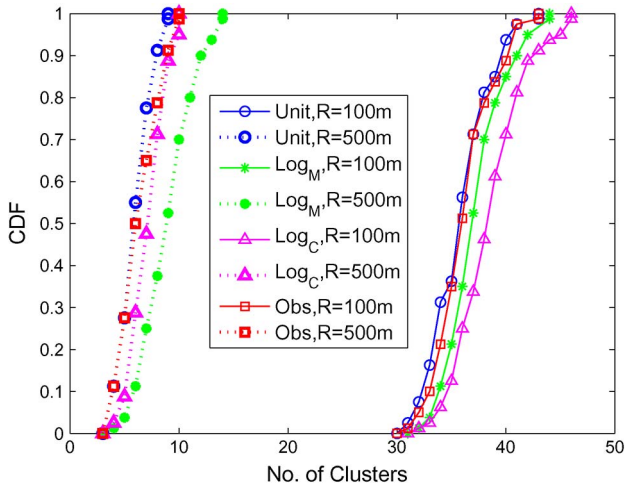

Fig. 11. CDF of the number of clusters metric for different channel models and transmission ranges in low-density networks.

high-density networks, respectively. The link duration of the matched lognormal model again is very close to that of the obstacle-based model for all scenarios. In addition, the link duration for the obstacle-based model is smaller than that of the unit disk model and larger than that of the lognormal model. The main reason is that the nodes can always communicate with each other within a threshold distance for the unit disk model, creating high correlation of the connectivity behavior over time and, thus, much higher link duration. On the other hand, the connections between the vehicles are probabilistically determined for the lognormal model, where the probability is independently chosen in each step, creating low correlation of the connectivity behavior and, thus, much lower link duration. The link duration for the obstacle-based model is closer to the unit disk model at a low transmission range and closer to the lognormal model at a high transmission range, which means that the correlation of the connectivity behavior decreases as the transmission range increases in the obstacle-based model. Moreover, we observe that the behavior of the link duration as a function of the transmission range does not agree for the unit disk and obstacle-based models. As the transmission range increases from 100 to 500 m, the link duration decreases under the obstacle-based model, whereas it increases under the unit disk model. The main reason for this behavior is that with the increase in the transmission range, the vehicles stay as neighbors for a longer time under the unit disk model, whereas the number of vehicles and their probability of acting as obstacles increase at a larger transmission range under the obstacle-based model. Furthermore, we notice that the link duration under the obstacle-based model decreases as the vehicle traffic density increases due to the increasing probability of the vehicles acting as obstacles at higher density.

Fig. 11 shows the cdf of the number of clusters metric for different channel models and transmission ranges in low-density networks. Since the number of clusters is very low for high-density networks when the transmission range is between 100 and 500 m, we did not include a separate graph for the high-density network. The distributions of the number of clusters based on the unit disk and obstacle-based models are very close to each other. The main reason for this similarity even at different transmission ranges is that the vehicle that acts as an obstacle between two vehicles also acts at the same time as a bridge between them, resulting in an indirect connection through the obstructing vehicle. The vehicles are directly connected when unit disk model is used, whereas they are connected through the obstacles in the obstacle-based model, resulting in the same number of vehicles within clusters. The number of clusters of the matched lognormal model is still close to that of the obstacle-based model but slightly worse than that of the unit disk and classical lognormal models. This small difference mainly results from not including the spatial correlation of the Gaussian variables in the lognormal model. The obstacle-based model implicitly includes the correlation among the links closer to each other in space due to the inclusion of the obstacles, whereas the random variables generated for each link in the matched lognormal model are independently chosen over space.

Fig. 12(a) and (b) shows the cdf of the size of the largest cluster for different channel models and transmission ranges in low- and high-density networks, respectively. The size of the largest cluster for unit disk and matched lognormal models is very close to that of the obstacle-based model, whereas the largest cluster size is very different for classical lognormal and obstacle-based models. This demonstrates the unsuitability of the classical lognormal model formulated with the same parameters at different network densities and the benefit of the additional matching process included in the matched lognormal model.

Fig. 13(a) and (b) show the cdf of the clustering coefficient metric for different channel models and transmission ranges in low- and high-density networks, respectively. Although the unit disk and obstacle-based models agree with each other in the number of clusters and size of the largest cluster metrics, we observe that their performance is very different when the clustering coefficient is considered. The reason is that the clustering coefficient provides the degree of connectivity of the vehicles within a cluster, which differentiates between the direct connectivity and the connection through the obstacles, unlike the number of clusters and size of the largest cluster metrics. On the other hand, the matched lognormal model provides clustering coefficient values much closer to those of the obstacle-based model than the unit disk and classical lognormal models in most scenarios. The slight difference between the matched lognormal and obstacle-based models again mainly comes from not including the spatial correlation of the Gaussian variables in the lognormal model.

Fig. 14(a) and (b) shows the cdf of the closeness centrality metric for different channel models and transmission ranges in low- and high-density networks, respectively. Similarly, the matched lognormal model provides very close performance to that of the obstacle-based model, unlike the classical lognormal model.

VIII. Validation of Results

Section VII demonstrates that the calibration of the path-loss exponent, standard deviation, and correlation factor parameters

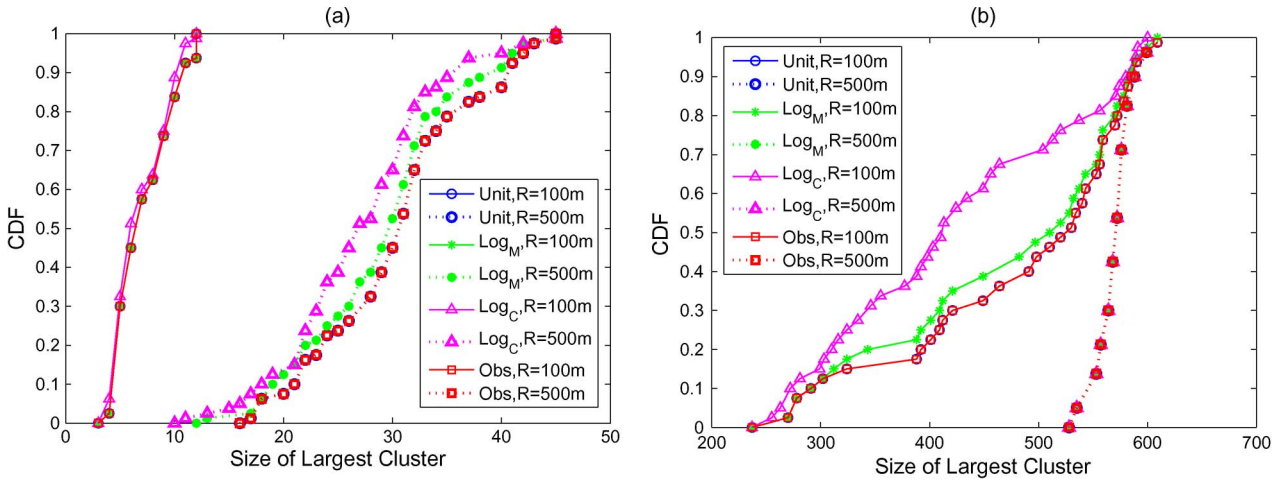

Fig. 12. CDF of the size of the largest cluster metric for different channel models and transmission ranges in (a) low-density and (b) high-density networks.

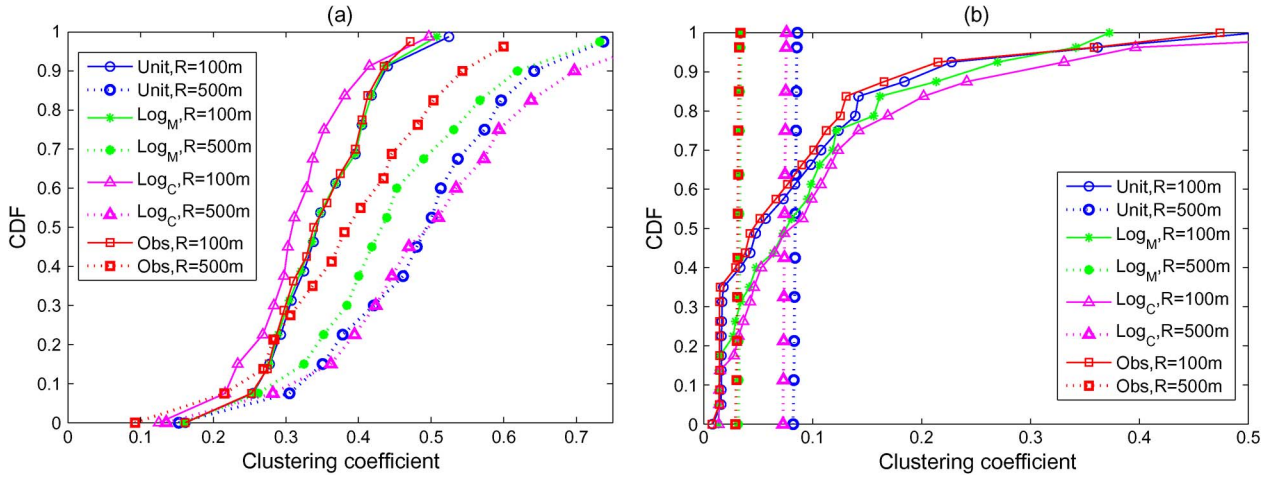

Fig. 13. CDF of the clustering coefficient metric for different channel models and transmission ranges in (a) low-density and (b) high-density networks.

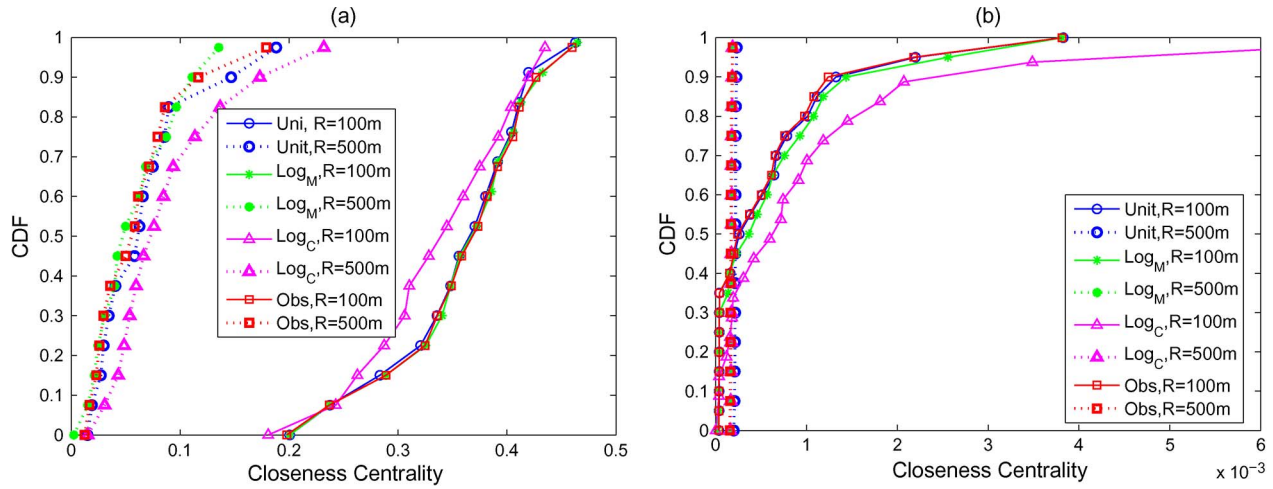

Fig. 14. CDF of the closeness centrality metric for different channel models and transmission ranges in (a) low-density and (b) high-density networks.

in the matched lognormal model achieves very close performance to the obstacle-based model in terms of several metrics at different traffic densities. This section aims to validate the dependence of the values of these parameters on the vehicle traffic density only, across different highways. This allows the users of the matched lognormal model to choose these parameters, considering the traffic density of their highway only, independent of the highway they are using for their simulation.

To validate the proposed matched lognormal model, we have extended the simulations for various vehicle traffic densities and three additional highway roads I5-S, I80-E, and I580-W in California and checked the agreement of the resulting matched

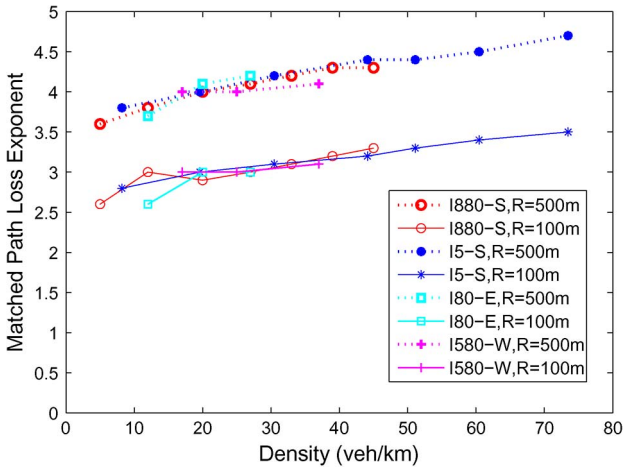

Fig. 15. Matched path-loss exponent for I880-S, I5-S, I80-E, and I580-W highway roads at different vehicle traffic densities and transmission ranges.

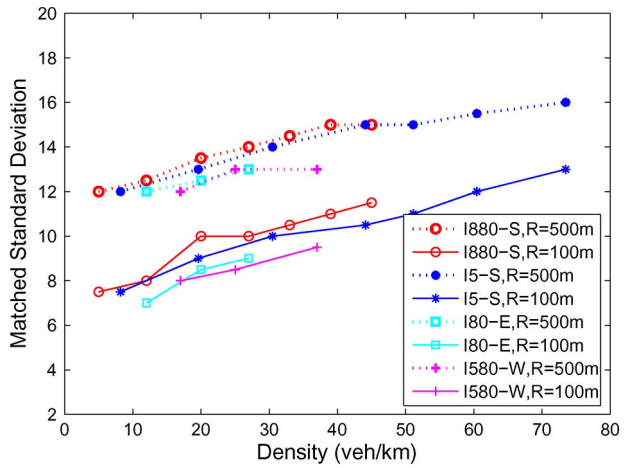

Fig. 16. Matched standard deviation for I880-S, I5-S, I80-E, and I580-W highway roads at different vehicle traffic densities and transmission ranges.

parameters, including the path-loss exponent, standard deviation, and correlation factor for the matched lognormal model on these highway roads. The realistic mobility over the additional roads is generated by determining the input and parameters of the microscopic mobility simulator SUMO based on the flow and speed information provided by the PeMS database, as explained in detail in Section III. The vehicle mobility output of SUMO is then input to MATLAB, where the values of the path-loss exponent, standard deviation, and correlation factor that provide the best match to the obstacle-based model are determined. These roads are chosen such that they are different in vehicle traffic density, the number of lanes, and intersections.

Figs. 15–17 show the matched path-loss exponent, standard deviation, and correlation factor values, respectively, for the I880-S, I5-S, I80-E, and I580-W highway roads at different vehicle traffic densities and transmission ranges. We observe that the matched values of these parameters are consistent across different highways at different vehicle traffic densities and different transmission ranges.

We have further validated the dependence of the matched parameters on the traffic density by checking the agreement between the cdf of the performance metrics obtained by matching

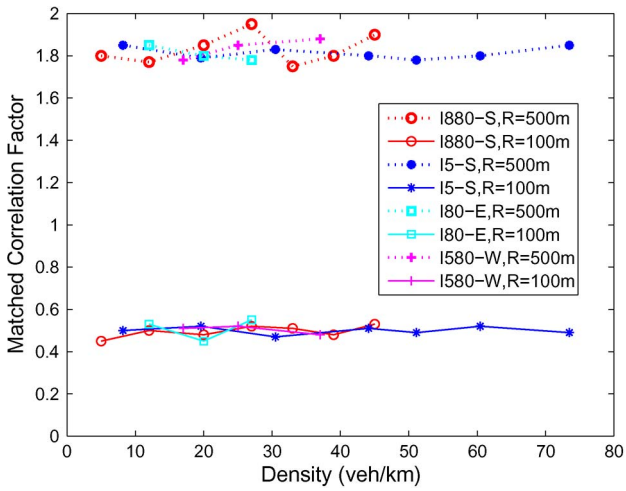

Fig. 17. Matched correlation factor for I880-S, I5-S, I80-E, and I580-W highway roads at different vehicle traffic densities and transmission ranges.

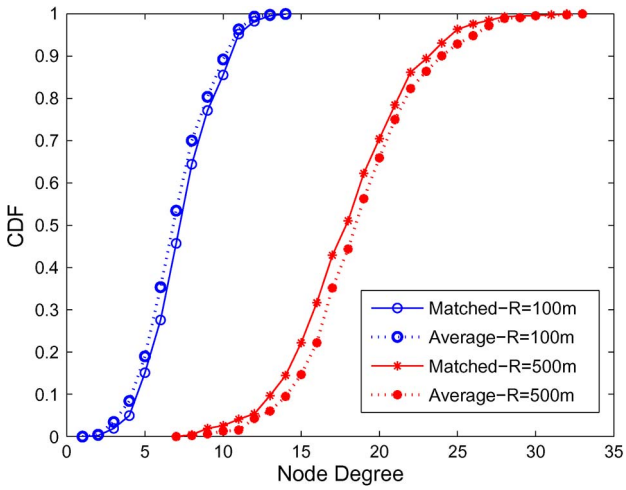

Fig. 18. CDF of the node degree obtained by matching to the obstacle-based model and using the matched parameters averaged over several roads on I580-W at high vehicle traffic density.

to the obstacle-based model and using the matched parameters averaged over several roads. As an example, Figs. 18 and 19 show the cdf of the node degree and link duration on I580-W at high vehicle traffic density, respectively. We observe good agreement of the cdf's.

IX. Conclusion

We analyze the spatial and temporal evolution of the VANET topology characteristics by using both realistic large-scale mobility traces and realistic channel models. The realistic large-scale mobility traces are obtained by using accurate microscopic mobility modeling of SUMO, determining its input and parameters based on the vehicle flow and speed data extracted from the PeMS database. The realistic channel model is obtained by implementing a recently proposed obstacle-based channel model that takes all the vehicles around the transmitter and the receiver into account in determining the received signal strength. The performance of the obstacle-based model is compared to the most commonly used and more simplistic

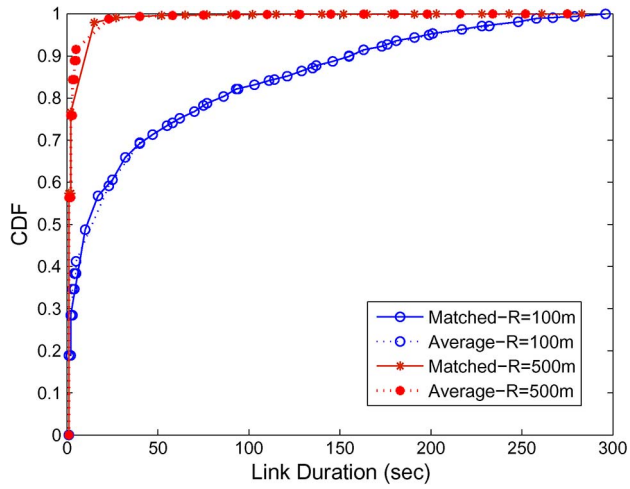

Fig. 19. CDF of the link duration obtained by matching to the obstacle-based model and using the matched parameters averaged over several roads on I580-W at high vehicle traffic density.

channel models, including unit disk and lognormal shadowing models.

The extensive investigation of the system metrics regarding the link characteristics over both time and space, including node degree, neighbor distance distribution, number of clusters, and link duration, reveals that tuning the parameters appropriately and introducing time correlation for the Gaussian random variable in the lognormal model provides a good match with the more sophisticated and computationally expensive obstacle-based model. We validate the dependence of the values of these parameters on the vehicle traffic density only based on the real data of four different highways in California.

As future work, we plan to extend the proposed realistic simulation model and matching mechanism for urban traffic scenarios.

REFERENCES

[1] N. Akhtar, O. Ozkasap, and S. Ergen, "VANET topology characteristics under realistic mobility and channel models," in *Proc. IEEE WCNC*, Apr. 2013, pp. 1774–1779.

[2] R. Chen, W.-L. Jin, and A. Regan, "Broadcasting safety information in vehicular networks: Issues and approaches," *IEEE Netw.*, vol. 24, no. 1, pp. 20–25, Jan./Feb. 2010.

[3] W. Chen and S. Cai, "Ad hoc peer-to-peer network architecture for vehicle safety communications," *IEEE Commun. Mag.*, vol. 43, no. 4, pp. 100–107, Apr. 2005.

[4] M. Torrent-Moreno, J. Mittag, P. Santi, and H. Hartenstein, "Vehicle-to-vehicle communication: Fair transmit power control for safety-critical information," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3684–3703, Sep. 2009.

[5] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET routing on city roads using real-time vehicular traffic information," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3609–3626, Sep. 2009.

[6] Y. Toor, P. Muhlethaler, and A. Laouiti, "Vehicle ad hoc networks: Applications and related technical issues," *IEEE Commun. Surveys Tuts.*, vol. 10, no. 3, pp. 74–88, Third Quart. 2008.

[7] S. A. H. Tabatabaei, M. Fleury, N. N. Qadri, and M. Ghanbari, "Improving propagation modeling in urban environments for vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 705–716, Sep. 2011.

[8] D. Dhoutaut, A. Régis, and F. Spies, "Impact of radio propagation models in vehicular ad hoc networks simulations," in *Proc. 3rd Int. Workshop VANET*, Sep. 2006, pp. 40–49.

[9] S. Uppoor and M. Fiore, "Large-scale urban vehicular mobility for networking research," in *Proc. IEEE VNC*, Nov. 2011, pp. 62–69.

[10] Tapascologne Project. [Online]. Available: http://sourceforge.net/apps/mediawiki/sumo/index.php?title=TAPASCologne

[11] D. Naboulsi and M. Fiore, "On the instantaneous topology of a large-scale urban vehicular network: The cologne case," in *Proc. ACM Int. Symp. MobiHoc*, Jul. 2013, pp. 167–176.

[12] G. Pallis, D. Katsaros, M. D. Dikaiados, N. Loulloudes ,and L.Tassiulas, "On the structure and evolution of vehicular networks," *Proc. MASCOTS*, Sep. 2009, pp. 1–10.

[13] V. Naumov, R. Baumann, and T. Gross, "An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces," in *Proc. 7th ACM Int. Symp. MobiHoc*, May 2006, pp. 108–119.

[14] M. Boban, W. Viriyasitavat, and O. Tonguz, "Modeling vehicle-to-vehicle line of sight channels and its impact on application-layer performance," in *Proc. 10th ACM Int. Workshop VANET*, 2013, pp. 91–94.

[15] Drive-in (Distributed Routing and Infotainment through Vehicular Inter-Networking). [Online]. Available: http://drive-in.cmuportugal.org

[16] W. Viriyasitavat, F. Bai, and O. K. Tonguz, "Dynamics of network connectivity in urban vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 3, pp. 515–533, Mar. 2011.

[17] M. Fiore and J. Harri, "The networking shape of vehicular mobility," in *Proc. ACM Int. Symp. MobiHoc*, May 2008, pp. 261–272.

[18] R. Meireles, M. Ferreira, and J. Barros, "Vehicular connectivity models: From single-hop links to large-scale behavior," in *Proc. IEEE VTC Fall*, Sep. 2009, pp. 1–5.

[19] R. Protzmann, B. Schunemann, and I. Radusch, "The influences of communication models on the simulated effectiveness of V2X applications," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 149–155, Nov. 2011.

[20] S. C. Ng, W. Zhang, Y. Zhang, Y. Yang, and G. Mao, "Analysis of access and connectivity probabilities in vehicular relay networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 140–150, Jan. 2011.

[21] X. Jin, W. Su, and Y. Wei, "Quantitative analysis of the VANET connectivity: Theory and application," in *Proc. IEEE VTC Spring*, May 2011, pp. 1–5.

[22] J. Gozalvez, M. Sepulcre, and R. Bauza, "Impact of the radio channel modelling on the performance of VANET communication protocols," *Telecommun. Syst.*, vol. 50, no. 3, pp. 149–167, Jul. 2012.

[23] Performance Measurement System (PeMS). [Online]. Available:http://pems.dot.ca.gov/

[24] SUMO—Simulation of Urban Mobility. [Online]. Available: http://sumo.sourceforge.net

[25] J. Harri, F. Filali, and C. Bonnet, "Mobility models for vehicular ad hoc networks: A survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 4, pp. 19–41, Fourth Quart. 2009.

[26] Vissim. [Online]. Available: http://vision-traffic.ptvgroup.com/en-uk/products/ptv-vissim/

[27] H. Conceicao, L. Damas, M. Ferreira, and J. Barros, "Large scale simulation of V2V environments," in *Proc. ACM Symp. Appl. Comput.*, Mar. 2008, pp. 28–33.

[28] B. Raney, A. Voellmy, N. Cetin, M. Vrtic, and K. Nagel, "Towards a microscopic traffic simulation of all of Switzerland," in *Proc. Int. Conf. Comput. Sci.*, Apr. 2002, pp. 371–380.

[29] J. Maurer, T. Fugen, M. Porebska, T. Zwick, and W. Wiesbeck, "A ray-optical channel model for mobile to mobile communications," presented at the 4th MCM COST, Wroclaw, Poland, Feb. 2008.

[30] C. F. Mecklenbrauker *et al.*, "Vehicular channel characterization and its implications for wireless system design and performance," *Proc. IEEE*, vol. 99, no. 7, pp. 1189–1212, Jul. 2011.

[31] J. Karedal, N. Czink, A. Paier, F. Tufvesson, and A. F. Molisch, "Path loss modeling for vehicle-to-vehicle communications," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 323–328, Jan. 2011.

[32] L. Cheng, B. E. Henty, F. Bai, and D. D. Stancil, "Highway and rural propagation channel modeling for vehicle-to-vehicle communications at 5.9 GHz," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, Jul. 2008, pp. 1–4.

[33] G. Setiawan, S. Iskandar, S. Kanhere, and K. chan Lan, "The effect of radio models on vehicular network simulations," in *Proc. 14th World Congr. ITS*, Oct. 2007, pp. 1–10.

[34] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 GHz dedicated short range communication (DSRC) frequency band," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1501–1516, Oct. 2007.

[35] O. Renaudin, V.-M. Kolmonen, P. Vainikainen, and C. Oestges, "Non-stationary narrowband MIMO inter-vehicle channel characterization in the 5-GHz band," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 2007–2015, May 2010.

[36] M. Boban, T. T. V. Vinhoza, M. Ferreira, J. Barros, and O. K. Tonguz, "Impact of vehicles as obstacles in vehicular ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 15–28, Jan. 2011.
[37] J. Karedal *et al.*, "A geometry-based stochastic MIMO model for vehicle-to-vehicle communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3646–3657, Jul. 2009.
[38] J. Maurer, T. Fugen, and W. Wiesbeck, "Narrow-band measurement and analysis of the inter-vehicle transmission channel at 5.2 GHz," in *Proc. IEEE VTC Spring*, May 2002, pp. 1274–1278.
[39] I. Sen and D. Matolak, "Vehicle-to-vehicle channel models for the 5-GHz band," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 235–245, Jun. 2008.
[40] D. Krajzewicz, M. Bonert, and P. Wagner, "The open source traffic simulation package SUMO," in *Proc. RoboCup*, Jun. 2006, pp. 371–380.
[41] J. Epstein and D. Peterson, "An experimental study of wave propagation at 850 Mc," *Proc. IRE*, vol. 41, no. 5, pp. 595–611, May 1953.
[42] J. Deygout, "Multiple knife-edge diffraction of microwaves," *IEEE Trans. Antennas Propag.*, vol. AP-14, no. 4, pp. 480–489, Jul. 1966.
[43] C. Giovaneli, "An analysis of simplified solutions for multiple knife-edge diffraction," *IEEE Trans. Antennas Propag.*, vol. AP-32, no. 3, pp. 297–301, Mar. 1984.
[44] "Propagation by Diffraction," Geneva, Switzerland, p. 526, Feb. 2007.
[45] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.
[46] M. Sepulcre and J. Gozalvez, "Adaptive wireless vehicular communication techniques under correlated radio channels," in *Proc. IEEE VTC Spring*, Apr. 2009, pp. 1–5.
[47] S. Szyszkowicz, H. Yanikomeroglu, and J. Thompson, "On the feasibility of wireless shadowing correlation models," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4222–4236, Nov. 2010.
[48] Data sets of vehicle mobility and communication channel models for realistic and efficient VANET highway simulation. [Online]. Available: http://cs-people.bu.edu/nabeel/VANETs_Mobility_Dataset/

Nabeel Akhtar (M'13) received the B.Sc. degree in computer science from Lahore University of Management Sciences, Lahore, Pakistan, in 2011 and the M.Sc. degree in computer science and engineering from Koc University, Istanbul, Turkey, in 2013, under the supervision of Prof. O. Ozkasap and Prof. S. C. Ergen. He is currently working toward the Ph.D. degree with Boston University, Boston, MA, USA, working under Prof. I. Matta.

His research interests include vehicular networks, vehicular mobility modeling, distributed algorithms, and future Internet architectures.

Sinem Coleri Ergen (M'13) received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2000 and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, CA, USA, in 2002 and 2005, respectively.

She was a Research Scientist with the Wireless Sensor Networks Berkeley Lab under the sponsorship of Pirelli and Telecom Italia from 2006 to 2009. Since September 2009, she has been an Assistant Professor with the Department of Electrical and Electronics Engineering, Koc University, Istanbul, Turkey. Her research interests are in wireless communications and networking with applications in sensor networks and transportation systems.

Dr. Ergen received the Science Academy Young Scientist Award (BAGEP) in 2014; the Turk Telekom Collaborative Research Award in 2011 and 2012; the Marie Curie Reintegration Grant in 2010; the Regents Fellowship from the University of California, Berkeley, in 2000; and the Bilkent University Full Scholarship in 1995.

Oznur Ozkasap (M'11) received the M.S. and Ph.D. degrees in computer engineering from Ege University, Izmir, Turkey, in 1994 and 2000, respectively.

From 1997 to 1999, she was a Graduate Research Assistant with the Department of Computer Science, Cornell University, Ithaca, NY, USA, where she completed her Ph.D. dissertation. She is currently an Associate Professor with the Department of Computer Engineering, Koc University, Istanbul, Turkey, which she joined in 2000. Her research interests include distributed systems, multicast protocols, peer-to-peer systems, bioinspired distributed algorithms, mobile ad hoc networks, energy efficiency, cloud computing, and computer networks.

Prof. Ozkasap served as an Area Editor of the *Computer Networks* journal, Elsevier Science, and as a Management Committee Member of the European COST Action IC0804: Energy efficiency in large-scale distributed systems.