A Probabilistic Model for Message Propagation in Two-Dimensional Vehicular Ad-Hoc Networks

Yanyan Zhuang
Dept of Computer Science
University of Victoria
Victoria, BC, Canada
yyzhuang@uvic.ca

Jianping Pan
Dept of Computer Science
University of Victoria
Victoria, BC, Canada
pan@uvic.ca

Lin Cai
Dept of Elec & Comp Engr
University of Victoria
Victoria, BC, Canada
cai@uvic.ca

ABSTRACT

Vehicular ad-hoc networks (VANET) promise to enhance the road safety and travel comfort significantly in both highway and city scenarios. Message propagation, either for emergency or pleasure purposes, constitutes a major category of VANET applications, and is particularly challenging in infrastructure-less vehicle-to-vehicle communication scenarios. In this paper, we study the connectivity property of message propagation in two-dimensional VANET. We first derive the exact expression for the average size of the connected components in the one-dimensional case, i.e., messages propagating along a main street, and give a close approximation to the size distribution. We further derive the connectivity of message propagation in the two-dimensional ladder case, i.e., covering the main and two side streets, and formulate the problem for the two-dimensional lattice case to cover all the blocks in a district. Extensive simulation has been conducted to verify the analytical model and provide further insights in message propagation with and without geographic constraints, respectively. The simulation results show the efficacy of the model and the tradeoff between these two message forwarding strategies and provide guidelines for future network planning and protocol development.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless communication*

General Terms

Performance

Keywords

VANET, message propagation, network connectivity

1. INTRODUCTION

Vehicular ad-hoc networks (VANET), a major component of the future intelligent transportation systems (ITS) [1],

promise to enhance the road safety and travel comfort significantly in both highway and city scenarios. With the allocation of frequency bands for the dedicated short-range communications (DSRC) [2] around the world, vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) wireless communications have attracted a lot of attention from academia, industry and standards organizations for protocol design, performance evaluation and system prototyping.

Message propagation, either for emergency or pleasure purposes, constitutes a major category of VANET applications. For example, emergency electronic break light messages can alert the drivers of following vehicles, even when their view has been obstructed by others. Also, vehicles in congested areas can notify the drivers of incoming vehicles to detour well in advance. Further, parking lots, hotels and restaurants can advertise their availability to vehicles several blocks away, reducing the extra time and fuel wasted when the drivers are looking for empty spots.

However, for message propagation in infrastructure-less V2V communication scenarios, it is very challenging to ensure that the message is delivered reliably and in time. Many research efforts have focused on MAC and routing protocol design and performance evaluation, but relatively few have considered the fundamental limits of message propagation, and how to explore both vehicle and traffic characteristics in highway or city scenarios to improve message delivery.

In this paper, we study the connectivity property of message propagation in two-dimensional VANET, i.e., how likely a message is propagated a certain distance away from the source. We first derive the exact expression for the average size of the connected components (vehicles connected through wireless communications) in the one-dimensional (1-d) case, i.e., messages propagating along a main street, and give a close approximation to the size distribution. The results show that due to the dimensional constraint on message propagation, there is always a non-negligible probability that message delivery cannot be guaranteed. To improve the connectivity and coverage, two-dimensional (2-d) message propagation is necessary. Thus we further derive the connectivity of message propagation for the 2-d ladder case and formulate the problem for the 2-d lattice case, for covering main-side streets and city blocks, respectively.

The contributions of this paper are threefold. First, although the limitation on the connectivity in the 1-d case has been shown in the literature, we have derived the exact expression for the average size of the connected components, using a much simpler and more accurate approach validated by simulation results. Second, we have derived the exact ex-

pression for the connectivity probability for the 2-d ladder case, for which we have not found similar results the literature yet. 2-d ladder is of practical importance in VANET, and the results can be applied to the often-seen main-side street scenarios (e.g., congestion detour to the side streets from a main street), or the scenario where messages are aggregated and re-propagated for each adjacent street. Simulation results have shown the accuracy of this model. Third, we have also formulated the problem for the 2-d lattice case, which generalizes the 1-d and 2-d ladder case. Although no closed-form expression is obtained for the 2-d lattice case due to the intrinsic hardness of the 2-d percolation process, we demonstrate the tradeoff between two message forwarding schemes with and without geographic constraints, respectively, and their performance, providing guidelines for future network planning and protocol development.

The rest of the paper is organized as follows. In Section 2, we briefly review the latest results in traffic modeling, percolation theory and message propagation. We then present our probabilistic models in Section 3, for 1-d (main street), 2-d ladder (main-side streets) and 2-d lattice (city blocks) cases, and compute the connectivity probability accordingly. Simulation results are presented in Section 4, to demonstrate the tradeoff between two message forwarding schemes for these three cases, followed by further discussion and concluding remarks in Section 5 and 6, respectively.

2. BACKGROUND AND RELATED WORK

With its high social-economy impact and unique features such as high-mobility vehicles and their highly predictable trajectory due to road constraints, VANET has offered many research opportunities and challenges. Here, we focus on the traffic modeling and message propagation results that are most relevant to our work.

2.1 Spatio-Temporal Vehicle Traffic Models

Early papers [3,4] have found that inter-vehicle distances follow an exponential distribution. Through recent statistical analysis of empirical data collected from real world scenarios, [5–8] again found that an exponential model is a good fit for highway vehicle traffic in terms of inter-vehicle distance and inter-contact time distribution. Given a single parameter, the vehicle density λ (number of vehicles per meter), this model is able to describe the characteristics and variation of highway traffic. In urban scenarios, vehicular traffic is much more complicated due to road grids, traffic lights and stop signs, but the exponential distribution is still widely used in the literature. For example, [9] assumed exponential distribution in a lattice-shaped road network, and studied network connectivity at any time instants, which is equivalent to our study at a snapshot of the network.

In this paper we assume that the inter-vehicle distances follow an independent and identical distribution, and use exponential distribution as an example. Section 3 shows that our approach can be applied to other distributions, so long as the the connectivity probability of two adjacent intersections, p, can be derived mathematically or obtained empirically through measurement.

2.2 Percolation Theory and Applications

In classic percolation theory [10], a stochastic percolating process is modeled as the process of liquid seeping through a porous object, which is usually modeled as a d-dimensional square lattice. Declare each edge in the lattice *open* with probability p, and closed otherwise, the liquid penetration is related to the existence of an *infinite connected cluster* of open edges: all open clusters have finite size when $p < p_c$, but there exists an infinite open cluster when $p \geq p_c$, where p_c is the critical probability. In contrast to liquid penetration, where edges can be open to all directions, there is another percolation process called *directed percolation* [11], i.e., edges only open to certain directions. According to the prediction on directed percolation [12], if $p < p_c$, the node connectivity probability will steadily decrease and eventually reach 0 w.r.t. the distance to the source; if $p \geq p_c$, the node connectivity probability will decrease gradually and eventually scatter around an asymptotic value. These percolation processes are similar to the message propagation in two-dimensional VANET: whether there exists a critical threshold that the entire network is almost surely connected, is certainly also of our interest.

However, our main focus is to determine the probability that a message is delivered to certain blocks away from the source. Although similar, percolation theory focuses more on the coverage: an infinite connected cluster does not guarantee messages are always reliably delivered; there may be a lot of isolated nodes that are not connected to the infinite cluster. In Section 4, the sub-critical and super-critical behaviors from the simulation look similar to percolation, but we also derive and formulate the accurate connectivity probability at each intersection from the message source in the 1-d and 2-d cases in Section 3, which is more microscopic and important when compared with identifying whether a network percolates or not, for message propagation.

2.3 Message Propagation and Connectivity

Existing studies have shown that the network connectivity in the 1-d case, e.g., on a highway, is always limited, whereas for 2-d cases, e.g., city blocks, network connectivity can be guaranteed if the density among nearby nodes is above a certain threshold. In [14], nodes were uniformly distributed along the interval $[0, z]$, forming a one-dimensional multi-hop radio network. Using Laplace Transforms, [14] derived the probability of network connectivity as a function of transmission range. In this paper, however, we assume that the inter-vehicle distance follows a certain statistical distribution. At the same vehicle density, the results in [14] are more optimistic due to the less bursty traffic.

By using an equivalent $GI|D|\infty$ queuing model, [15] derived the connectivity distribution in one-dimensional networks. Node positions were equivalent to customer arrivals, and the transmission range was modeled as a constant service time. The results were obtained by Laplace-Stieltjes Transforms (LST) and simplifications were used to obtain a closed-form expression. In Section 3, we discuss and compare the model in [15] with our model. Derivation and simulation show that our approach is simpler than [5], and the results are more accurate than [15]. [16] studied emergency message propagation with time constraints, and derived lower bounds on the probability that a car at a certain distance can receive the messages on time. Their lower bounds depends on channel reliability and message dissemination strategy. However, the inter-vehicle distance distribution is based on a set of n equally spaced vehicles. In [17], the fraction of the vehicles that belong to the largest connected component in a two-dimensional graph was analyzed.
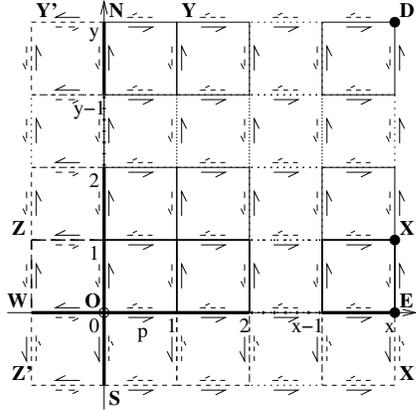
Figure 1: Two-Dimensional City Grid.

These papers, however, did not consider how likely the message can propagate to a certain location in the network.

It is proved in [9] that Poisson positioning allows the existence of a finite critical vehicle density, and with non-Poisson positioning, there still exists a critical vehicle density, but at a larger value. It is proved in [18] that the critical density exists for a hexagon model, and the connectivity probability increases sharply within a short interval around this density. These papers either applied percolation theory in a scenario other than the traditional physics and biology systems, or proved the existence of a critical threshold; whereas the approach in this paper can be applied to the cases with general vehicle distributions, and our work focuses more on the connectivity property of message propagation (how likely a message can be delivered to a location) rather than coverage (how many vehicles can receive the message).

3. PROBABILISTIC MODEL

In this paper, we consider a two-dimensional city grid as shown in Figure 1. Suppose W–E is the main street stretching from west to east (similarly S–N from south to north), and a message is generated by a vehicle at O. The question we want to answer is: how likely the message will be propagated to a certain location by the vehicles on the street? E.g., intersection E for the 1-d case, X for the 2-d ladder case, and D for the 2-d lattice case, respectively. For modeling, we use the geographically-constrained (*geo-constrained*) message forwarding strategy, where vehicles only forward messages to their neighbors who are further away from the message source, following the solid arrows in Figure 1. Please refer to Section 4 and 5 for evaluation and discussion on other forwarding strategies. Here we only consider the wireless communication and relay between vehicles, and do not consider vehicle mobility, which is at a much lower speed when compared with the speed of electromagnetic waves. The "carry-and-forward" effect for delay-tolerant message propagation in mobile VANET is our future work.

3.1 One-Dimensional Message Propagation

For the 1-d case, we focus on the main street W–E, from O to E as an example. If the message can reach E, it means O and E are in a "connected" cluster, in which any vehicle can reach each other by wireless transmission and relay. *Cluster size*, the distance between the first and last vehicles in the same cluster, is of great importance in message propagation, so in the following we characterize the cluster size first.

3.1.1 Moments of Cluster Size

The one-dimensional cluster size has been studied in [15]. Using a $GI|D|\infty$ queuing model, the expected cluster size $E[B]$ is approximated by

$$E[B] \approx R + \frac{\int_0^R x f_X(x)\mathrm{d}x}{1 - F_X(R)}, \qquad (1)$$

where R is the transmission range, and $F_X(x) = \int_{-\infty}^{x} f_X(x)\mathrm{d}x$ is the cumulative distribution function (CDF) of the inter-vehicle distance. Assuming the vehicle arrival follows a Poisson process, i.e., $f_X(x) = \lambda e^{-\lambda x}$ for $x > 0$, thus (1) gives

$$E[B] \approx R + \frac{\int_0^R \lambda x e^{-\lambda x}\mathrm{d}x}{e^{-\lambda R}} = \frac{1 - e^{-\lambda R}}{\lambda e^{-\lambda R}}. \qquad (2)$$

As the simulation results show below, the above results are not accurate in low-density networks. Thus, we use a simple approach to derive the moments of cluster size without any approximation as follows.

Denote the random variable (RV) for cluster size by C. Because the inter-vehicle distances are assumed i.i.d. RVs[1], we have the following recursion for the expectation of C:

$$\begin{aligned} E[C] &= E[C|X_1 < R] \times \Pr\{X_1 < R\} \\ &= (E[X_1|X_1 < R] + E[C]) \times \Pr\{X_1 < R\}, \ (3) \end{aligned}$$

where X_1 is the distance between the first two vehicles in the cluster. For the second step in (3), given the i.i.d. distribution of the inter-vehicle distance, the average cluster size starting from any vehicle (which is treated as the cluster head) is the same. Thus, $E[C|X_1 < R] = E[X_1|X_1 < R] + E[C']$, where $E[X_1|X_1 < R]$ is the average distance between the first and the second vehicles in the cluster, and $E[C']$ (equal to $E[C]$ asymptotically) is the average cluster size if we choose the second vehicle as the cluster head.

This recursive approach works for any inter-vehicle distance distributions and is simpler than [5]. In the following, we use exponential distribution as an example: $\Pr\{X_1 < R\} = 1 - e^{-\lambda R}$. Given the conditional expectation $\overline{X_1'} = E[X_1|X_1 < R] = \int_0^R \lambda x e^{-\lambda x}/(1 - e^{-\lambda R})\mathrm{d}x$, we have

$$E[C] = \frac{1 - e^{-\lambda R}}{e^{-\lambda R}} \times \overline{X_1'} = \frac{1 - e^{-\lambda R}(\lambda R + 1)}{\lambda e^{-\lambda R}}. \qquad (4)$$

This simple approach gives the *exact* average cluster size without any simplification. Figure 2 compares $E[B]$ and $E[C]$ with simulation results. λ is set to be 0.015 (number of vehicles per meter), and the x-axis and y-axis are the radio transmission range R and the expected cluster size in log scale, respectively. In the figure, when $\lambda \times R$ is sufficiently large, results from both $E[B]$ and $E[C]$ match the simulation well; but for small R and λ, our approach is much more accurate. Similar results are obtained for other traffic densities λ and are omitted here for brevity.

Higher-order moments of C can be derived in a similar way. For instance, the second-order moment of C is

$$\begin{aligned} E[C^2] &= \Pr\{X_1 < R\} \times E[(C + X_1)^2|X_1 < R] \\ &= \frac{1 - e^{-\lambda R}}{e^{-\lambda R}} \times \left(E[C^2] + 2E[C]\overline{X_1'} + \overline{X_1'^2} \right), (5) \end{aligned}$$

---

[1]As the width of the street is negligible when compared with the street length, all vehicles in a two-way street are assumed along a line, no matter which lane they actually occupy.
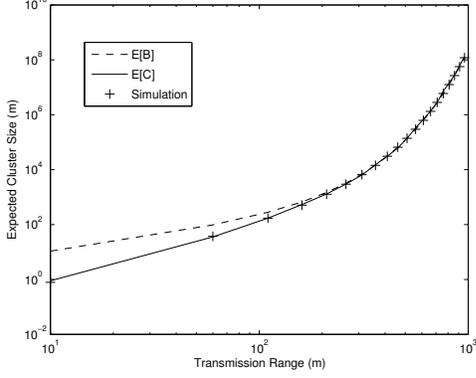
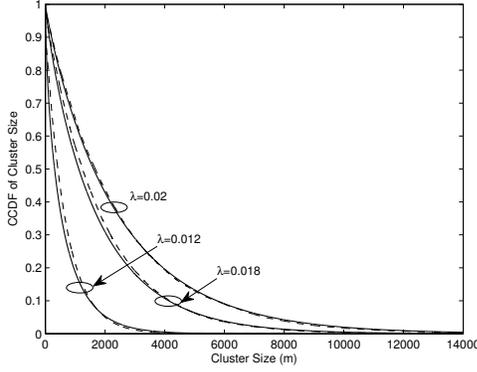**Figure 2: Comparison on Expected Cluster Size.**

Figure 3: Cluster Size Distribution Approximation (Solid line: Gamma Approx., Dashed: Simulation).

where $\overline{X_1'^2} = E[X_1^2 | X_1 < R] = \int_0^R \lambda x^2 e^{-\lambda x} \mathrm{d}x$. If we derive all the moments, they can characterize a unique distribution. In the next subsection, we study the density function of C, which is very useful for deriving the connectivity probability.

3.1.2 Cluster Size Distribution

Let X_i's $(i = 1, 2, ...)$ be the RVs of the inter-vehicle distance between the i-th and $(i + 1)$-th vehicle, given that they are in the same cluster. We have

$$f_{X_i}(x) = f(x | 0 \leq x \leq R) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda R}}, \quad \text{for } 0 \leq x \leq R. \quad (6)$$

Suppose that there are k vehicles in a cluster, the Laplace Transform of the cluster size distribution is

$$f_{C|k}^*(s) = f_{\sum_{i=1}^k X_i}^*(s) = \left[\frac{\lambda}{1 - e^{-\lambda R}} \times \frac{1 - e^{-(s+\lambda)R}}{s + \lambda} \right]^k. \quad (7)$$

If we can get $f_{C|k}$ by taking the inverse-Laplace Transform on (7), $f_C(x) = \sum_{k=1}^{\infty} f_{C|k} \Pr\{k\}$, where $\Pr\{k\} = (1 - e^{-\lambda R})^{k-1} e^{-\lambda R}$ is the probability that there are k vehicles in a cluster. Unfortunately, (7) does not give any closed-form result by the inverse-Laplace Transform. The problem is complicated since C is the sum of k *truncated* exponential RVs, given that they are smaller than the cutoff value R, and k itself follows a Geometric distribution.

According to the last subsection, however, we can get the statistical moments of cluster size. Using these moments, we use a known *Gamma* distribution for approximation, since the Gamma distribution has been widely used to model the

sum of exponentially distributed random variables. In Figure 3, vehicle transmission range is $200\ m$, the dashed curves are from simulation and the solid curves are from the corresponding Gamma approximation:

$$f_C(x) = x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)}, \quad \text{for } x > 0, \quad (8)$$

where $k = (E[C^2]/E[C]^2 - 1)^{-1}$ and $\theta = E[C]/k$ to ensure that the first and second-order moments of the Gamma approximated random variables are the same as $E[C]$ and $E[C^2]$. As shown in the figure, the Gamma distribution is a good fit for the cluster size distribution.

Once we have the cluster size distribution or approximation, the probability that a message can reach a location at distance d away from the source, is the probability that the cluster size is larger than d, i.e., $\int_d^{\infty} f_C(x) \mathrm{d}x$. From $f_C(x)$, it is clear that no matter how large the traffic density is, there is always a non-negligible probability that the gap between two adjacent vehicles is beyond the radio transmission range R. Once a vehicle is disconnected from the vehicle in front of it, the rest of the 1-d network is also disconnected. It has been proved that in a 1-d network, the network is almost surely disconnected [13], and the probability that a further-away location can be reached decays very quickly, suggested by the Gamma approximation of C. Thus, we should explore the message propagation beyond one dimension, i.e., not just along the main street.

On the other hand, based on the above approximated density function of C, we are able to get the probability that two nearby intersections are connected. This probability is useful for studying the two-dimensional connectivity.

3.2 Two-Dimensional Message Propagation

The probability that two given intersections are connected is very important for message propagation, and it has not been purposely investigated in the existing percolation theory. We denote p as the probability that two adjacent intersections of a street segment in a 2-d grid are connected. In the following, we first derive p; then, we study the connectivity property in the 2-d ladder and lattice case.

3.2.1 Bond Probability

p, as defined above, is equivalent to the "bond probability" in percolation theory. If wireless transmissions to other perpendicular streets are heavily shadowed, which might be the case in cities with high-rise buildings along the street, p can be simplified as the probability that the one-dimensional cluster size is larger than the distance between two intersections, so it can be calculated directly using the cluster size distribution in (8). For more realistic situations when wireless transmissions can reach perpendicular streets, p can be derived as follows.

As shown in Figure 4, V_e, V_s, V_w and V_n represent the vehicles (if existing) closest to the right intersection on the east, south, west, and north streets, respectively. Their distance to the right intersection are d_e, d_s, d_w and d_n. Vehicle V_o is on the street between the two intersections and is closest to the left intersection. V_o is connected to the source node, and its distance to the left intersection is d_o. The distribution of d_o is a truncated exponential function $\lambda e^{-\lambda t}/(1 - e^{-\lambda R})$, for $0 \leq t \leq R$. We consider the following two disjoint cases that at least one of the vehicles V_e, V_s or V_n is connected to V_o, i.e., message propagation can
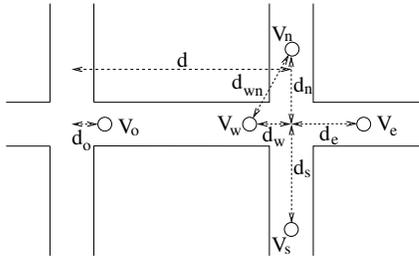
Figure 4: Bond Probability Illustration.

continue in a new street segment, regardless of its direction, therefore the two intersections are connected.

Case One: The cluster originating from V_o has a size larger than $d - d_o$, where d is the distance between two adjacent intersections. By considering V_o's location and cluster size distribution, the probability for this case is

$$p_1 = \int_0^R \int_{d-t}^\infty f_C(x) \mathrm{d}x \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda R}} \mathrm{d}t.$$

In this case, V_e is connected to V_o. V_s is connected, either because it is within the transmission range of V_w or V_e, or it is within the transmission range of V_n, and V_n is within the transmission range of V_w or V_e.

Case Two: In this case, the cluster size originating from V_o is smaller than $d - d_o$; the last vehicle connected to V_o is V_w, and $d_e + d_w > R$. In addition, one of V_n and V_s is within the transmission range of V_w. In this case, V_e is outside the direct transmission range of V_w, but at least one vehicle on the south or north street can receive the message and continue the propagation.

Denote the cluster size originating from V_o as x and let $d_0 = t$, to be connected to V_w, the minimum of d_s and d_n must be no larger than $\sqrt{(\eta R)^2 - (d - x - t)^2}$, where $\eta \in (0, 1)$ is the shadowing effect for signal transmissions to other perpendicular streets. Thus, given x and t, the conditional probability that at least one vehicle on the north or south street is within the transmission range of V_w is $(1 - e^{-2\lambda\sqrt{(\eta R)^2 - (d-x-t)^2}})$. The probability of this case is

$$p_2 = \int_0^R \int_{d-t-\eta R}^{d-t} (1 - e^{-2\lambda\sqrt{(\eta R)^2 - (d-x-t)^2}}) f_C(x) \mathrm{d}x \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda R}} \mathrm{d}t.$$

Considering the above two disjoint cases, p is given by

$$p = p_1 + p_2. \tag{9}$$

As shown in Figure 5, the analytical results given in (9) match well with the simulation. Here, the distance between two adjacent intersections is $d = 500\ m$ and $\eta = 1$. The simulation is done by scattering vehicles randomly in a 2-d square lattice, according to the exponential inter-vehicle distance distribution. The detailed description of simulation settings is given in Section 4.

3.2.2 Ladder Connectivity

In this subsection we look at the 2-d ladder connectivity, e.g., two side streets Z–X and Z'–X' along the main street W–E in Figure 1. Events such as congestion detour messages are of interest to the main street and its side streets. We derive the connectivity probability between any two intersections in this ladder topology, with the constraint that the message can only be propagated further away from the source, i.e., geo-constrained forwarding, as an example.
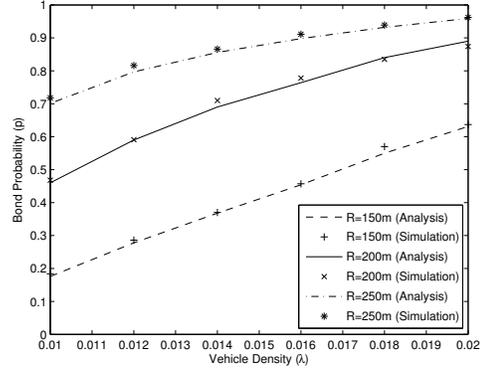

Figure 5: Bond Probability Validation.

In a ladder topology, as shown in Figure 6, the source is at $(0, 0)$ and the destination is at (x, y) where $y = \pm 1$ (we use $y = 1$ for illustration). Looking at the first row in Figure 6, where $(1, 1)$ is the destination, there are two paths to reach $(1, 1)$: path A_1 where $(0, 0)$ and $(1, 1)$ are connected by the intermediate node $(0, 1)$, and path B_1 where the intermediate node is $(1, 0)$. Given that each edge is connected with p, by the principle of inclusion-exclusion (PIE), we have the connectivity probability for $(1, 1)$

$$
\begin{aligned}
P(1, 1) &= P(A_1 + B_1) = P(A_1) + P(B_1) - P(A_1 B_1) \\
&= P(A_1) + P(B_1) - P(B_1|A_1)P(A_1) \\
&= p^2 + p^2 - p^2 * p^2 = 2p^2 - p^4. \tag{10}
\end{aligned}
$$

When $x \geq 1$, recursion is needed to derive the probability. For example, in the second row of Figure 6, where $(2, 1)$ is the destination, it has paths A_2 and B_2. B_2 is dependent on A_2. As shown in the $B|A$ column, event $(B_2|A_2)$ can be degenerated to a horizontal segment plus a triangle (i.e., $(1, 1)$ and $(2, 1)$ merge to a single point), given the segments in A_2 are already connected. That is, $P(B_2|A_2) = p * (p + p^2 - p^3)$, since the two paths (from $(1, 0)$ to $(2, 1)$) in the degenerated triangle are independent (illustrated in the second sub-row of the second row). On the other hand, $P(B_2)$ is simply $p * P(1, 1)$, where $P(1, 1)$ is given in (10). Using this recursive method, the connectivity probability for $(2, 1)$ can be derived as

$$
\begin{aligned}
P(2, 1) &= p^3 + p * P(1, 1) - P(B_2|A_2)P(A_2) \\
&= 3p^3 - 2p^5 - p^6 + p^7. \tag{11}
\end{aligned}
$$

Similar recursion can be done when $x = 3$. In the third row of Figure 6, we notice that $P(A_3)$ is simply p^4, $P(B_3)$ is related to $P(2, 1)$, and the triangle in $B_3|A_3$ can be decomposed into two sub-cases, one is simply p and the other is related to $B_2|A_2$. Applying the same recursion for squares and degenerated triangles iteratively for all x's, we get the following for $P(x, 1)$:

$$P(x, 1) = p[p^x + P(x - 1, 1) - p^x \theta(x)], \quad x \geq 1 \tag{12}$$

with $P(0, 1) = p$ and

$$\theta(x) = p[p + \theta(x - 1) - p\theta(x - 1)], \quad x \geq 1 \tag{13}$$

with $\theta(0) = 0$, where x can be any integer. The derivation is symmetric for the ladder connectivity in the other three quadrants for the main-side street scenarios.
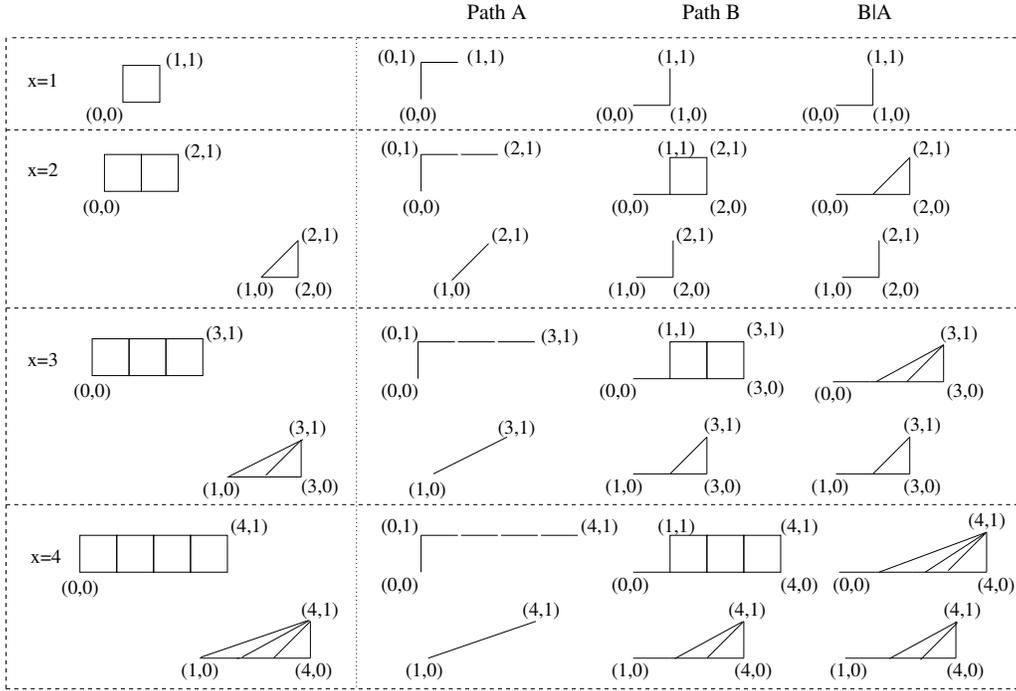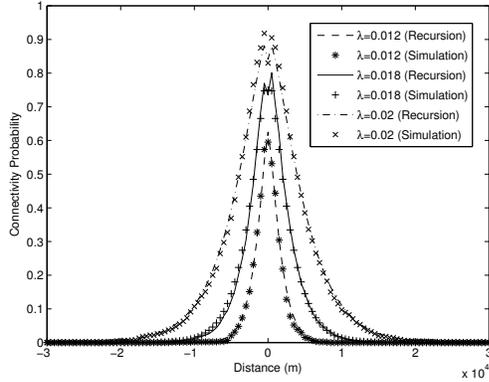
Figure 6: Ladder Connectivity Illustration.

Figure 7: Ladder Connectivity Validation.

Figure 7 shows the results from recursion (12) and (13), and compares them with simulation results. The vehicle transmission range is 200 m and the distance between two adjacent intersections is 500 m. The x-axis is the west-east coordinate of side street Z–X, which is 500 m away from the parallel main street. The message source O is on the main street with x-coordinate equal to 0. As shown in the figure, the recursive analysis and the simulation match well with each other, and we have not found similar results in the literature yet. The results show that the connectivity probability decays fast w.r.t. the distance, for all traffic densities. However, when compared with the results from the 1-d case, the connectivity probability at the same x-coordinate offset with the same traffic density is much higher. For example, with $\lambda = 0.02$, the connectivity probability at $x = 4\ km$ offset for the 2-d ladder case is above 0.5, while the connectivity for the 1-d case is only around 0.2, as shown in Figure 3, since there are more paths available from the source in the 2-d ladder case, while the 1-d case only has one.

3.2.3 Lattice Connectivity

For the general 2-d lattice case, e.g., the origin-destination pair O–D in Figure 1, we have the following numerical formulation to derive the end-to-end connectivity probability. First, we enumerate all the possible paths from $(0,0)$ to (x, y), then by the principle of inclusion-exclusion (PIE), $P(x, y)$ can be obtained by calculating the probabilities of different combinations of paths and crosschecking their overlapping street segments. As an example, we assume the destination point is $(2, 1)$, and then all the possible paths starting from $(0, 0)$ are

$$
\begin{aligned}
A: & \quad (0,0) \rightarrow (1,0) \rightarrow (2,0) \rightarrow (2,1) \\
B: & \quad (0,0) \rightarrow (1,0) \rightarrow (1,1) \rightarrow (2,1) \\
C: & \quad (0,0) \rightarrow (0,1) \rightarrow (1,1) \rightarrow (2,1)
\end{aligned}
$$

By PIE,

$$
\begin{aligned}
P(2,1) &= P(A + B + C) \\
&= P(A) + P(B) + P(C) - P(AB) \\
&\quad - P(AC) - P(BC) + P(ABC) \\
&= p^3 + p^3 + p^3 - p^5 - p^5 - p^6 + p^7,
\end{aligned}
$$

which has the same result as that in Section 3.2.2. $P(AB) = p^5$ and $P(BC) = p^6$ since paths A and B together have 5 non-overlapping street segments, while paths B and C have 6 non-overlapping ones.

Ideally, any $P(x, y)$ can be computed by this enumeration-combination method. Unfortunately, this approach suffers from the combinatorial explosion problem and does not scale, and the computation becomes intractable when $x + y$ becomes larger. For instance, when $x = 5$, $y = 3$, the number of different paths is $\binom{x+y}{y} = 56$, and the number of different combinations of these 56 paths can be as many as $\binom{56}{28} = 7.6487 \times 10^{15}$, each of which has $|x| + |y| = 8$ street segments. If we store these street segments in a bit map, which requires

38 bits per path since there are $(x+1)y + x(y+1) = 38$ unique street segments, then the memory required will be $38 \times 7.6487 \times 10^{15}/8$ bytes $\approx 3.63 \times 10^7$ GB—this is only for a 5×3 lattice. The problem of deriving connectivity probabilities with arbitrary x and y is a central problem of directed percolation [11] in Physics and Stochastic Processes, which still remains unsolved after many years of efforts.

4. PERFORMANCE EVALUATION

In the last section, we have derived the connectivity probability for the 1-d and the 2-d ladder case, and found the 2-d lattice case is not yet analytically solvable. In this section, we go further and evaluate the static connectivity [19] of these cases by simulation. In vehicular networks, there are different ways for vehicles to forward the messages they have received. Geo-constrained forwarding (GF) that we studied in Section 3 is commonly used to avoid message redundancy, link contention, hidden terminal and broadcast storm problems. Unconstrained forwarding (UF), on the other hand, is an extension of the former in the sense that vehicles can forward messages to neighbors in all directions, so messages may potentially go backwards, as shown by both the solid and dashed arrows in Figure 1. Unconstrained forwarding may improve network connectivity, but may also increase the overhead in terms of broadcast cost, i.e., the number of transmissions of each message. It is interesting to quantify and compare the performance and cost of these two message forwarding strategies, with or without geographic constraints: unconstrained forwarding is equivalent to percolation process, whereas geo-constrained forwarding corresponds to *directed* percolation in [11].

In both message forwarding strategies, whenever a vehicle overhears the same message from the neighbors in the opposite direction, it will not rebroadcast that message again, in order to reduce overhead and possible collisions; otherwise, the vehicle rebroadcasts the message periodically up to a retransmission limit. We use simulation to compare these two forwarding strategies, with and without geographic constraints, respectively, and reveal the tradeoff between them. All simulation is done in Matlab. By using an exponential random variable generator and a square lattice map, vehicle locations are constrained to the vertical and horizontal streets. Street are 500 m apart, all vehicles have the same transmission range R, varying from 150 to 300 m with $\eta = 1$, and all the results are averaged over 1,000 simulation runs.

4.1 Geo-Constrained Forwarding

In geo-constrained forwarding, similar to *directed* percolation [11], the probability of each street segment to be open is determined by p in (9). We first plot the network connectivity that measures the percentage of the vehicles connected to the source node in a 3×3 km^2 grid map. Figure 8 shows the change of network connectivity, from 0 to 1, with different vehicle densities. Such transitions are similar to the results in [18]. Combining Figures 5 and 8, we note that the critical threshold p_c (above which the network connectivity is close to 1) is much larger than 0.5 (the critical threshold for the *undirected* bond percolation in a grid [9]).

In Figure 9, different curves in each sub-figure show the connectivity probability for the vehicles on different parallel streets in the lattice network, and the x-axis represents the coordinates along each street. The coordinate of the source is $(0,0)$, and the transmission range R is 200 m. Different
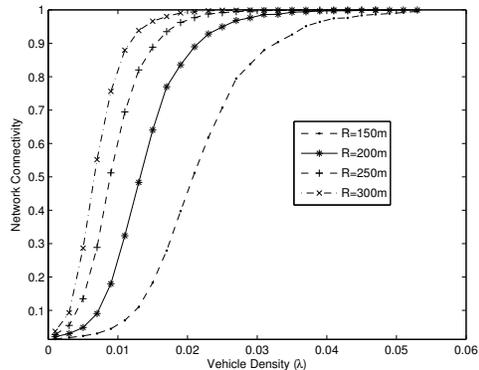

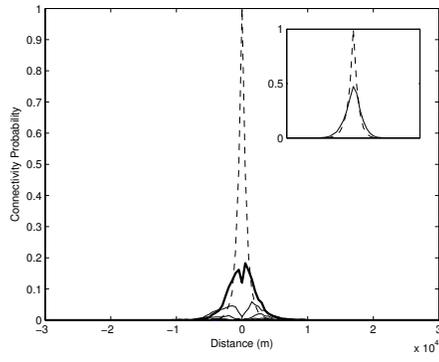**Figure 8: Network Connectivity with GF.**

sub-figures show the results with different vehicle densities λ, varying from 0.01 to 0.015 vehicles per meter. The values of λ are carefully chosen according to Figure 8, so the bond probability p is below, close, or above the critical threshold for the directed bond percolation.

Typically, the further away a street is from the message source, the lower the connectivity probability is for the vehicles on that street, thus the message propagation within a few blocks away from the source (e.g., the bold curves in this figure) is of major concern. In Figure 9 we plot the connectivity probability up to 30 km, in order to give a big picture of the entire network, e.g., how the connectivity probability decays or converges. Figures 9(a)–(d) show the simulation results for geo-constrained forwarding, of which the results are: around the critical point (with λ from 0.012 to 0.015), even with a small change in traffic density, the connectivity probabilities change significantly; once the critical point is reached, the peak of the connectivity probability for each street converges to an asymptotic value, as shown in Figure 9(d), but the connectivity decays eventually.

Another interesting observation is that, except the main street, the connectivity probability on all the other streets is not monotonic w.r.t. x. The dashed, Λ-shaped curves in Figures 9(a)–(d) are for the main street; the bold curves are for the side streets that are 3 blocks away; the solid curves in the zoom-in window show the results for the 2-d ladder case. All curves, except the ones for the main street, have an M-shape, which means being closer to the message source does not necessarily lead to a higher connectivity probability. This is because, the closer an intersection is to the source, the smaller number of paths exist between this intersection and the source, although each path is shorter. If an intersection is on either of the streets where the source is located, for example along W–E in Figure 1, vehicles can only forward data in one direction: as a result, the connectivity for them is the same as that in the 1-d case.
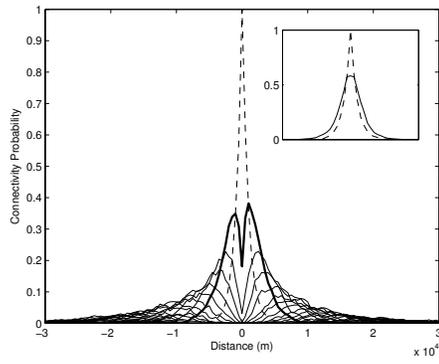
4.2 Unconstrained Forwarding

Observing the cluster size distribution in the 1-d case and the limited connectivity probability in a 2-d lattice with geo-constrained forwarding, we find that clusters are likely to be small and a large number of vehicles are unable to receive the message. By allowing the message to go through other directions, as indicated by the dashed arrows in Figure 1, clusters can be extended in these directions, i.e., including more neighbors for message delivery.
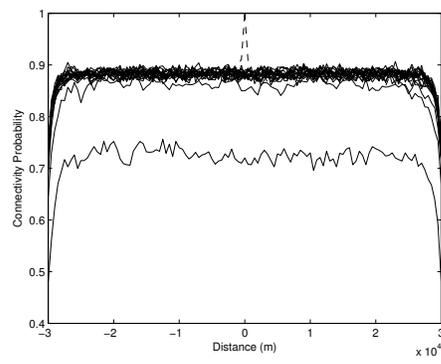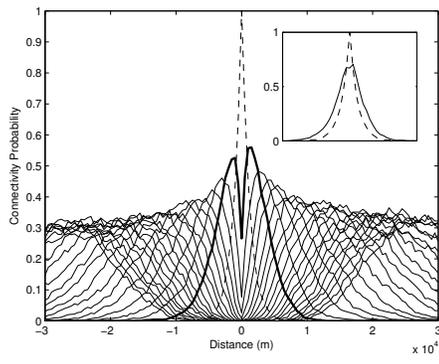
(a) GF, $\lambda = 0.01$
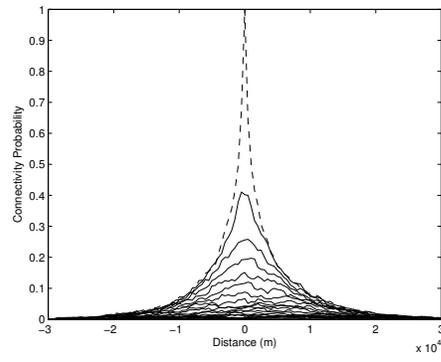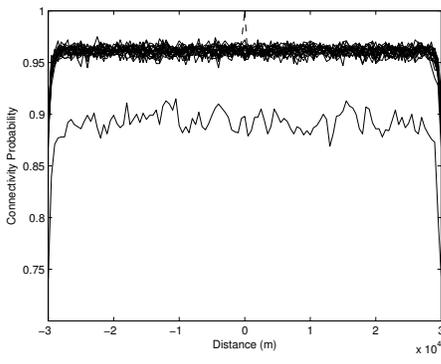
(e) UF, $\lambda = 0.01$

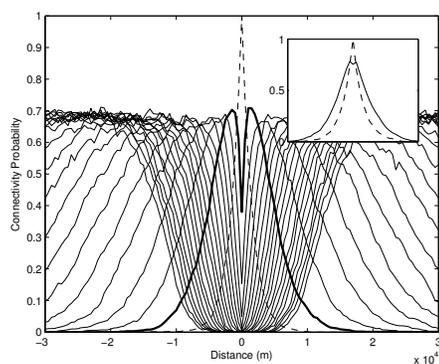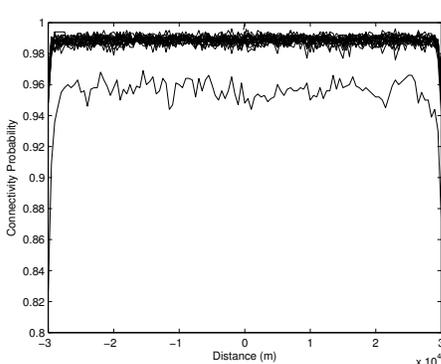(b) GF, $\lambda = 0.012$

(f) UF, $\lambda = 0.012$

(c) GF, $\lambda = 0.0135$

(g) UF, $\lambda = 0.0135$

(d) GF, $\lambda = 0.015$

(h) UF, $\lambda = 0.015$

Figure 9: Connectivity Probability: Geo-Constrained Forwarding (GF) vs. Unconstrained Forwarding (UF).
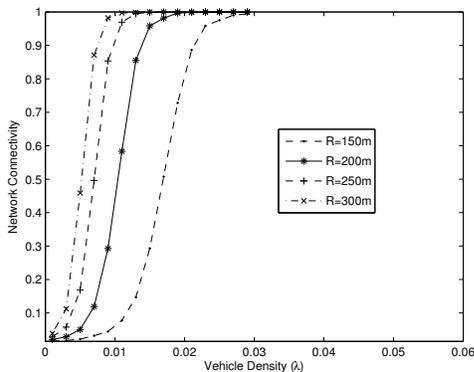
Figure 10: Network Connectivity with UF.

As shown in Figure 10, the transitional behavior in unconstrained forwarding is similar to that in geo-constrained forwarding, except that the critical threshold is much smaller, and the transitions is also sharper. p can be found in Figure 5, e.g., when $R = 200\ m$ and λ is slightly larger than 0.01, the corresponding p is exactly 0.5. From Figure 9, we can see another difference between these two forwarding schemes. At super-critical stage, the connectivity probability in geo-constrained forwarding converges to a value that is much smaller than 1, whereas in unconstrained forwarding, the connectivity probability is almost 1 as p approaches the critical value, given the same vehicle density. The Λ-shaped curves in Figures 9(e)–(h) still correspond to the main street, where the street farthest away from the source has a much lower connectivity probability due to edge effect.

In contrast to geo-constrained forwarding, the connectivity probability here is either monotonically deceasing with distance or scattering around an asymptotic value. Taking a closer observation, the results in Figure 9(e) with $\lambda = 0.01$ for unconstrained forwarding is similar to Figure 9(b) with $\lambda = 0.012$ for geo-constrained forwarding, because the latter has relatively limited choices of path selection. As shown in Figures 9(e) and (f), λ is only increased by 0.002, i.e., adding one vehicle every 500 m, the connectivity probability of vehicles increases drastically from below 20% to higher than 85%. When λ is further increased to 0.0135, as shown in Figure 9(g), the connectivity probability of most vehicles exceeds 95%. After that, further increasing λ does not change the connectivity probability significantly, as in Figure 9(h). This tells us that when p is above the critical threshold 0.5, there is no more significant increase in network connectivity. The above results can help the application to decide whether it is necessary to increase p, e.g., by using a larger transmission power, or use unconstrained forwarding to improve connectivity.

4.3 Broadcast Cost

The high connectivity achieved by unconstrained forwarding is at the expense of a higher broadcast cost than geo-constrained forwarding. Figures 11(a) and (b) show the average number of transmissions for each message with these two forwarding schemes, using the same parameter settings for comparison. Here, we use a simplified setting by assuming that there is no transmission error for each broadcast, and the vehicle within the transmission range and farthest away from the transmitter will rebroadcast. Geo-constrained forwarding apparently has much lower broadcast cost, although its network connectivity is limited. Therefore it reduces the amount of network resources required, such as communication bandwidth and buffer space.

The decision whether to use geo-constrained or unconstrained forwarding, thus also depends on the application. If messages need to be propagated as far as possible, e.g., hotel, dining, or parking lot information, then using unconstrained forwarding (probably at a lower message frequency) can let the information traverse the network without a high demand on vehicle density. Road hazard or traffic congestion, on the other hand, only has high impact on nearby streets, therefore the messages need to be propagated within a few hops as soon as possible; in this case geo-constrained forwarding is preferable due to its low broadcast cost and potentially fewer collisions.

5. FURTHER DISCUSSION

In this paper, we have derived the connectivity probability for the 1-d and 2-d ladder cases, based on traffic density λ, radio transmission range R, and street segment length d, which can be encapsulated into the connectivity probability of two adjacent intersections, i.e., p. To derive p, we have to make assumptions on λ (e.g., homogeneous exponential distribution). However, no matter how p is derived mathematically or obtained empirically, our approach to deriving the connectivity probability for the 2-d ladder case still applies. Also, our approach to deriving the connectivity probability in the 1-d case is much simpler and more accurate. We used geo-constrained forwarding for derivation, but unconstrained forwarding can be considered as well.

Note that, with a simple extension, ladder connectivity in (12) and (13) can be used to derive the connectivity probability where the bond probability of horizontal street segments is different from that of vertical street segments. In addition, this approach can be used to deal with dynamic traffic shifting: when the vehicles in the main street are no longer connected (e.g., due to detour), the two side streets can form new ladders themselves, and the recursive method can be applied to recalculate the connectivity probability. Other practical factors, such as packet loss due to wireless channel impairments or collisions and how the physical, link, and network layers react to them, can be further considered in deriving the bond probability, so our work can be extended to include these realistic wireless channel characteristics and the implementation details of the MAC and routing protocols. Our results also can be used to study the tradeoff between connectivity and throughput: if each node use a lower data rate for transmission, then the transmission range can be larger, but each message transmission will occupy the channel longer, and more wireless resources will be consumed. Incorporating throughput, collision and scheduling analysis is critical in modeling vehicular networks, and it is open for further research. Nevertheless, the physical connectivity in our study is a necessary condition for the connectivity at the network and higher layers.

So far our work has been limited to a snapshot of the network only, and has not considered vehicle mobility, or vehicles traveling at different speed in different road segments, i.e., the snapshot may become time variant. One possible extension to the existing work is to use mobility trace in the existing simulation, and study the characteristics of network connectivity under certain mobility models. Also, "carry-
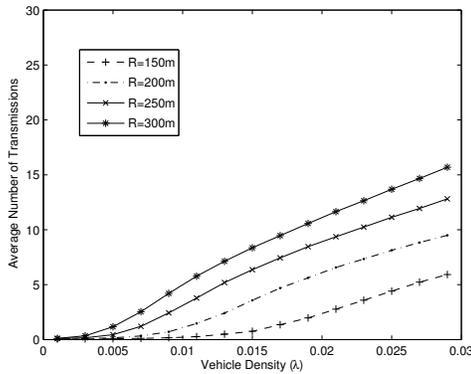
(a) Geo-Constrained Forwarding
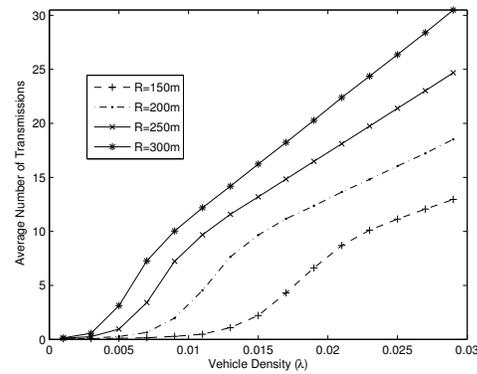
(b) Unconstrained Forwarding

Figure 11: Broadcast Overhead: Average Number of Transmissions.

and-forward" messages between vehicles traveling in opposite directions can be applied for delay-tolerant applications. Different from unconstrained forwarding, carry-and-forward utilizes vehicle mobility, instead of forwarding directions, to opportunistically extend the cluster size. In two-dimensional networks, this problem is much more difficult since vehicles may turn to the other directions at intersections.

6. CONCLUSIONS

In this paper, we have considered the fundamental limits of message propagation in VANET, and explored both vehicle and message characteristics in one and two-dimensional cases for highway and city scenarios. Starting from a one-dimensional street, we have derived and analyzed the network connectivity properties in the 2-d ladder and lattice cases. As obtaining the analytical result for the lattice topology is very difficult (a closed-form or recursive expression, if found, would constitute a major milestone in Stochastic Processes.), we have used extensive simulation to gain insights from the tradeoff between geo-constrained and unconstrained forwarding, for future network planning and protocol development. Pursuing further results in an arbitrary two-dimensional lattice network, and solving the problems posed in Section 5, thus constitute our ongoing work.

7. REFERENCES

[1] J. Blum, A. Eskandarian, and L. Hoffman, "Challenges of inter-vehicle ad hoc networks," *IEEE Trans on Intelligent Transportation Syst*, 5(4):347–351, 2004.

[2] "Dedicated Short Range Communications (DSRC) Home," [Online]. `http://www.leearmstrong.com/DSRC/DSRCHomeset.htm`

[3] H. Saito, "Performance analysis of combined vehicular communication", *IEICE Trans on Communications*, E89B(5):1486–1494, 2006.

[4] M. Schönhof, A. Kesting, M. Treiber and D. Helbing, "Coupled vehicle and information flows: message transport on a dynamic vehicle network", *Physica A–Stat Mech and ITS Apps*, 363(1):73–81, 2006.

[5] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar and O. K. Tonguz, "Routing in Sparse Vehicular Ad Hoc Wireless Networks", *IEEE J on Selected Areas in Communications*, 25(8):1538–1556, 2007.

[6] S. Ukkusuri and L. Du, "Geometric connectivity of vehicular ad hoc networks: Analytical characterization", *Transportation Research Part C: Emerging Technologies*, 16(5):615–634, 2008.

[7] F. Bai and B. Krishnamachari, "Spatio-temporal variations of vehicle traffic in VANETs: facts and implications," in *Proc. ACM VANET'09*, 2009.

[8] H. Zhu, L. Fu, G. Xue, etc, "Impact of traffic influxes: understanding exponential inter-contact time in VANETs", in *Proc. IEEE INFOCOM'10*, 2010.

[9] S. Shioda, J. Harada, Y. Watanabe, etc, "Fundamental characteristics of connectivity in vehicular ad hoc networks", *Proc. IEEE PIMRC*, 2008.

[10] G. Grimmett, "Percolation", 2nd Ed, *Springer*, 1999.

[11] L. C. Chen and F. Y. Wu, "Directed percolation in two dimensions: An exact solution", in *Differential Geometry and Physics, Nankai Tracts in Math.*, Vol. 10, pp. 160-168, 2006.

[12] C. Kurrer and K. Schulten, "Dependence of percolation thresholds on lattice connectivity", *Phys. Rev.*, E 48, pp. 614-617, 1993.

[13] O. Dousse, "Asymptotic properties of wireless multi-hop networks", *Ph.D. Thesis, EPFL*, 2005.

[14] M. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *IEEE Communications Letters*, 6(10), pp. 437–439, 2002.

[15] D. Miorandi and E. Altman, "Connectivity in one dimensional ad hoc networks: a queuing theoretic approach," *ACM/Springer Wireless Networks*, 2006.

[16] G. Resta, P. Santi and J. Simon, "Analysis of multi-hop emergency message propagation in vehicular ad hoc networks," in *Proc. ACM MobiHoc'07*, 2007.

[17] M. Kafsi, P. Papadimitratos, O. Dousse, T. Alpcan, and J.-P. Hubaux, "VANET Connectivity Analysis," in *Proc. IEEE Autonet'08*, 2008.

[18] H. Cai, X. Jia and M. Sha, "Critical sensor density for partial connectivity in large area wireless sensor networks", in *Proc. IEEE INFOCOM'10*, 2010.

[19] W. Viriyasitavat, O.K. Tonguz and F. Bai, "Network Connectivity of VANETs in Urban Areas," in *Proc. IEEE SECON'09*, 2009.