

# Capacity and Delay Analysis for Social-Proximity Urban Vehicular Networks

Ning Lu<sup>†</sup>, Tom H. Luan<sup>†</sup>, Miao Wang<sup>†</sup>, Xuemin (Sherman) Shen<sup>†</sup>, Fan Bai<sup>‡</sup>

<sup>†</sup>Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, N2L 3G1

<sup>‡</sup>Electrical and Controls Integration Lab, General Motors Global R&D, Warren, MI 48105, USA

Email: {n7lu, hluan, m59wang}@uwaterloo.ca, xshen@bbcr.uwaterloo.ca, fan.bai@gm.com

**Abstract**—In this paper, the asymptotic capacity and delay performance of social-proximity urban vehicular networks with inhomogeneous vehicle density are analyzed. Specifically, we investigate the case of  $N$  vehicles in a grid-like street layout while the number of road segments increases linearly with the population of vehicles  $N$ . Each vehicle moves in a localized mobility region centered at a fixed social spot and communicates to a destination vehicle in the same mobility region via a unicast flow. With a variant of the two-hop relay scheme applied, we show that social-proximity urban networks are scalable: a constant average per-vehicle throughput can be achieved with high probability. Furthermore, although the throughput and delay of a unicast flow may degrade in a high density area, almost constant per-vehicle throughput  $\Omega(\frac{1}{\log(N)})$  and almost constant delay  $O(\log^2(N))$  (except for the polylogarithmic factor) are still achievable with high probability. By identifying the key impact factors of performance mathematically, our results might provide insight on the design and deployment of future vehicular networks.

## I. INTRODUCTION

Emerging vehicular ad hoc networks (VANETs) are vehicle-centric large-scale communication networks; by equipping vehicles with on-board wireless communication facilities, communications among vehicles in proximity can be enabled. With VANET, a variety of applications relating to the safety (e.g., collision detection and lane changing warning) and infotainment [1]–[3] (e.g., file and other valuable information sharing) can be provided to drivers and passengers on the road. This not only makes the transportation system safer and more efficient, but also revolutionizes users' in-vehicle experience.

Under the umbrella of mobile ad hoc networks, VANETs present unique characteristics in terms of mobility, density and applications. More specifically, in VANETs, vehicles have *map-restricted* and *localized* mobility with specific social features rather than moving everywhere in the entire network: a vehicle typically moves only within a bounded region which relates to the social life of the driver. For instance, a vehicle may usually move within a small area close to the driver's home or the work place. This mobility feature is also reported in [4] based on the analysis of the real-world mobility trace of taxis in the city of Warsaw. It is observed that the mobility of taxis is typically around certain social spots and the density of vehicles within the proximity area of social spots follows the empirical heavy-tailed distribution. Moreover, VANETs show high *spatial variations* of vehicle density [1] and, more importantly, are mainly involved in the

*proximity-related* applications, such as safety applications and localized file sharing. All above observations make it neither practical nor necessary to maintain a long-lasting unicast communication flow over a long-distance. With the presence of unique features aforementioned, the fundamental performance in communication delay and capacity of VANETs has not been studied in existing literature, which motivates us to study the following three issues:

- i. what is the performance in communication delay and capacity of the urban VANET with social-proximity characteristics?
- ii. is social-proximity VANET scalable as its network size continually grows in the city?
- iii. how to operate this network to attain improved network performance in terms of capacity and delay?

In this paper, we develop theoretical analysis on the throughput capacity and delay performance of the social-proximity urban VANET. Doing so, we model urban area as a scalable grid in which the number of road segments increases linearly with the number of vehicles in the network. With localized nature of vehicle's mobility, we apply restricted vehicle mobility model surrounding a fixed social spot. In addition, we target to support social-proximity applications using this network, such that unicast flows are established between vehicles in which both the source and the destination of each flow are within the proximity of location around certain social spot. With a variant of the two-hop relay scheme [5] deployed, we show that network scales with an average constant per-vehicle throughput with high probability (*w.h.p.*); even in some area with extremely high vehicle density, it is possible to achieve almost constant throughput and delay (except for a polylogarithmic factor). Finally we obtain  $delay/throughput = O(\log^3(N))$  in this social-proximity urban vehicular network, indicating that social-proximity urban vehicular networks are scalable.

The remainder of this paper is organized as follows: Section II surveys the related works. In Section III, we introduce the network models and the notations used. Section IV analyzes the asymptotic throughput capacity and average packet delay with the proposed two-hop relay scheme. Section V concludes the paper.

## II. RELATED WORK

The throughput capacity of wireless networks was initially investigated by Gupta and Kumar in [6], where it has been

shown that the per-node throughput decays at least as  $1/\sqrt{N}$  in the presence of  $N$  nodes in the network. Since then, the study of capacity scaling in different networking scenarios has received extensive attention from academia [7]–[10]. Gross-glauser and Tse [11] first studied the effect of mobility on throughput capacity and have shown that due to the mobility, the per-node throughput of the mobile ad hoc network could remain constant at the cost of enlarged delay. Motivated by this result, substantial research has been conducted, such as [5], [12]–[14], to investigate tradeoffs between the delay and throughput for mobile ad hoc networks. El Gamal *et al.* [12] studied the throughput and delay under random walk model. With a two-hop relay scheme similar to that proposed in [11], they have shown that a per-node throughput of  $\Theta(1)$ <sup>1</sup> can be achieved with a delay of  $\Theta(N \log N)$ . In [5] Neely and Modiano studied the throughput capacity and delay of a cell-partitioned ad hoc network under the i.i.d. mobility model. They found that a general delay-throughput tradeoff can be established: the ratio of delay and throughput is at least  $O(N)$  under different scheduling policies applied with or without packet redundancy.

By noticing that nodes often spend most of the time in proximity of a few preferred places within a localized area, some researchers have been motivated to study the throughput and delay under the restricted node mobility. Li *et al.* [15] investigated the impact of a restricted mobility model on throughput and delay of a cell-partitioned network. They found that smooth throughput-delay tradeoffs in mobile ad hoc networks can be obtained by controlling the mobility pattern of nodes. Garetto and Leonardi [16] showed that the delay-throughput tradeoffs can be improved by restricting the node mobility. They considered a restricted mobility that the node moves around a fixed home-point according to a Markov process, while the stationary distribution of the node location decays as a power law of exponent  $\delta$  with the distance from the home-point. They showed that it is possible to exploit node heterogeneity under a restricted mobility model to achieve  $\Theta(1/\log^2(N))$  throughput capacity and  $O(\log^4(N))$  delay by using a sophisticated bisection routing scheme. However, the capacity and delay are still unclear as the network is mainly used to support proximity-related applications. Moreover, except those in [16], nodes are typically assumed to be uniformly distributed in the network in existing literatures. However, in the urban area, the densities of vehicles in different regions are highly diversified. Therefore, it is desirable to consider the network with inhomogeneous node density. In [10] Alfano *et al.* study the capacity scaling in a network with heterogeneous node densities. However, they only consider the network with stationary nodes.

Although VANETs have received a lot of attentions, the research on the fundamental performance (asymptotic through-

put capacity and delay) of VANETs is still limited. The impact of road geometry on the capacity of the network is investigated in [17] and [18]. In [19] Wang *et al.* consider a general multi-cast capacity scaling for an arterial road system. Unlike these works that have not addressed the delay issues, it is one of the key focuses of this paper to study the delay performance.

### III. SYSTEM MODEL

#### A. Network Geometry

We consider a square urban area with a grid-like street layout, like Manhattan area or the downtown area of a city. As shown in Fig. 1, the network geometry comprises of a set of  $M$  parallel roads intersected with another set of  $M$  parallel roads. Each line segment in Fig. 1 represents a road segment with bi-directional vehicle traffics. For the simplicity of analysis, the overall area considered is normalized to one unit, i.e., each side of the grid is of unit one in length.

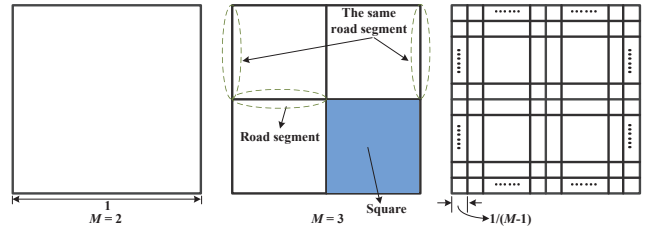


Fig. 1. A grid-like street layout.

The grid holds wrap-around conditions (i.e., the torus) to eliminate the border effects. Let  $C$  denote the number of squares in the grid. The total number of road segments (the road section between any two neighboring intersections) is therefore  $G = 2C = 2(M - 1)^2$ . We define the network density  $d = \frac{N}{G} = \frac{N}{2(M-1)^2}$ , where  $N$  is the total number of vehicles in the network. The network density  $d$  is kept constant such that the number of road segments  $G$  increases linearly with the total number of vehicles  $N$ . Note that  $d$  represents the average vehicle density in the grid. As each vehicle moves following the mobility model with social features, the spatial distribution of vehicles is inhomogeneous, as an example shown in Fig. 3(b).

*Remark:*  $M$  and  $d$  characterize the scale of the grid. A grid with a very large  $M$  and a relatively large  $d$  can model metropolitan areas like New York City; whereas a small town would have relatively small values of  $M$  and  $d$ . Therefore, from a macroscopic view, the grid-like geometry with different values of  $M$  and  $d$  can model urban scenarios of different scales.

#### B. Mobility Model

*Markovian Mobility Pattern:* The vehicle mobility is modeled by a discrete time Markovian process as follows. We consider a discrete time system where time is slotted with equal duration. The road segments are indexed from 1 to  $G$  and vehicle nodes are indexed from 1 to  $N$ . The movement of vehicles are independent to each other in the grid. The

<sup>1</sup>We use standard order notations in the paper: given nonnegative functions  $f_1(n)$  and  $f_2(n)$ ,  $f_1(n) = O(f_2(n))$  means  $f_1(n)$  is asymptotically upper bounded by  $f_2(n)$ ;  $f_1(n) = \Omega(f_2(n))$  means  $f_1(n)$  is asymptotically lower bounded by  $f_2(n)$ ; and  $f_1(n) = \Theta(f_2(n))$  means  $f_1(n)$  is asymptotically tight bounded by  $f_2(n)$ .

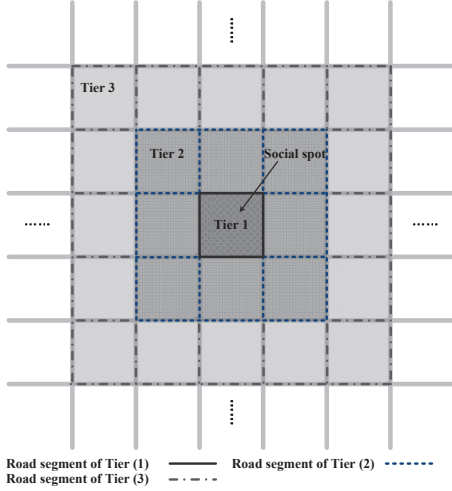


Fig. 2. Restricted mobility with different tiers for a given vehicle.

mobility of a vehicle  $k$  follows a discrete time Markovian process, denoted by  $\mathcal{C}(k)$ ,  $k \in \{1, 2, \dots, N\}$ , which is uniquely represented by a one-dimensional  $G$ -state ergodic Markov Chain.  $\mathcal{C}_t(k) = i$  if vehicle  $k$  appears on road segment  $i$ ,  $i \in \{1, 2, \dots, G\}$ , at time slot  $t$ ,  $t \in \{1, 2, \dots, T\}$ . Let  $P_{ij}(k)$  denote the transition probability that vehicle  $k$  moves from road segment  $i$  to the next road segment  $j$ ,  $j \in \{1, 2, \dots, G\}$ . Let  $\mathbf{P}(k) = \{P_{ij}(k)\}_{G \times G}$  denote the transition probability matrix of  $\mathcal{C}(k)$ ; the element  $P_{ij}(k)$  in  $\mathbf{P}(k)$  is non-zero only if  $j$  is a neighboring road segment of  $i$  in Fig. 1. Following this Markovian model, the vehicle can generate a mobility pattern in the network. The steady-state location distribution of vehicle  $k$  is  $\boldsymbol{\pi}(k) = \{\pi_i(k)\}_{1 \times G}$ , where  $\pi_i(k)$  denotes the long-term proportion of time that vehicle  $k$  stays on road segment  $i$ . In [20], it is shown that the capacity region only depends on how the node location distributes in the steady state. In [21], it is shown that the Markovian mobility model converges to its steady-state location distribution at an exponential rate. Therefore, in what follows we focus on the steady-state location distribution of the vehicles.

*Restricted Mobility Region with Social Spot:* We apply the socialized restricted mobility model to individual vehicles. Specifically, as shown in Fig. 2, the mobility region of each vehicle is composed of multiple co-centered tiers which are formed as follows. Each vehicle uniformly chooses one square out of the  $C$  squares in the grid as  $Tier(1)$  of its mobility region. It comprises of four road segments and specifies the social spot of the vehicle's mobility. The adjacent squares surrounding  $Tier(1)$  form  $Tier(2)$ , and so forth. Let  $Tier(\mathcal{A})$  denote the outermost tier of the mobility region, where  $\mathcal{A}$  is a constant scaler smaller than or equal to  $\lfloor M/2 \rfloor$ . It could be easily derived that  $Tier(\alpha)$ ,  $\alpha \in \{1, 2, \dots, \mathcal{A}\}$ , contains  $16\alpha - 12$  road segments. In this paper, the mobility of each vehicle is constrained in  $\mathcal{A}$  tiers with a total number of  $\sum_{\alpha=1}^{\mathcal{A}} 16\alpha - 12 = 4\mathcal{A}(2\mathcal{A} - 1)$  road segments.

In a randomly selected  $Tier(\alpha)$ ,  $\alpha \in \{1, 2, \dots, \mathcal{A}\}$ , the vehicle has equal steady-state location probability on each road

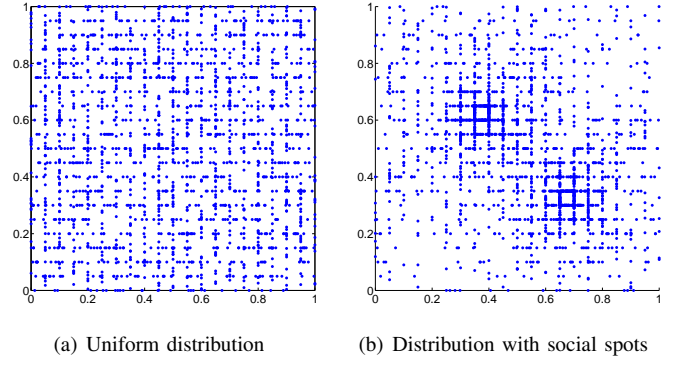


Fig. 3. Examples of homogeneous (a) and inhomogeneous (b) distributions of vehicles, in the case of  $N = 2000$ ,  $M = 21$  and  $\mathcal{A} = 10$ .

segment. Let  $\pi_\alpha$  denote the steady-state location probability of each vehicle on the road segments of  $Tier(\alpha)$ . From  $Tier(1)$  to  $Tier(\mathcal{A})$ , the steady-state location probability of vehicles is modeled to exponentially decay as a power law function with exponent  $\gamma > 0$ . As such, we have  $\pi_\alpha = \alpha^{-\gamma} \pi_1$  which indicates that a vehicle is more likely to stay in the network near its social spot. The same model has been used in [16] and its accuracy is validated in [4] through real-world measurements. As the summation of steady-state probability on road segments equals to 1, i.e.,

$$\sum_{\alpha=1}^{\mathcal{A}} (16\alpha - 12) \pi_\alpha = \sum_{\alpha=1}^{\mathcal{A}} (16\alpha - 12) \alpha^{-\gamma} \pi_1 = 1, \quad (1)$$

we have,

$$\pi_1 = \frac{1}{\sum_{\alpha=1}^{\mathcal{A}} (16\alpha - 12) \alpha^{-\gamma}}. \quad (2)$$

In summary, the mobility of each vehicle is uniquely characterized by its social spot, which is uniformly and independently selected over the  $C$  squares in the grid. The squares are indexed from 1 to  $C$ . Let  $\mathbf{H} = (H_1, H_2, \dots, H_N)$  denote the set of locations of the vehicles' social spots, with each element  $H_k$ ,  $H_k \in \{1, 2, \dots, C\}$ , denoting the location of vehicle  $k$ 's social spot. The set  $\mathbf{H}$  is fixed once the network is defined. Using the mobility model discussed above, the network presents inhomogeneous vehicle densities in the network. Fig. 3(b) shows an example of the vehicle distribution with social spots.

### C. Traffic Model

We consider that  $N$  unicast flows exist concurrently in the network. Each vehicle is exactly the source of one unicast flow and the destination of another unicast flow. We consider the case in which the source and destination vehicles of each unicast flow have the same social spot. This is motivated by the dominant proximity applications in vehicular communications. As such, the source and destination vehicles of each unicast flow are spatially close to each other. Without the loss of generality,  $N$  is considered to be even. We sort the index of vehicles such that vehicle  $k$  communicates with  $k + 1$ ,

$k \in \{1, 3, 5, \dots, N-1\}$ , and each communication pair randomly chooses a square in the grid as their social spot following the uniform distribution. The packet arrives in each unicast flow at an average rate  $\eta$ .

#### D. Communication Model

We assume that a pair of vehicles can communicate only when they are on the same road segment at the same time slot. With all vehicles using the single-channel transceiver, at each time slot, at most one transmission can occur on each road segment. The success or failure of a transmission is determined by the protocol model defined in [6] as follows. Let  $r = \frac{1}{M-1}$  denote the communication range of each vehicle which covers the entire road segment. The transmission from vehicle  $i$  to vehicle  $j$  can be successful during time slot  $t$  if and only if the following condition holds:  $d_{kj}(t) \geq (1 + \Delta)r$ , for every other vehicle  $k$  transmitting simultaneously, where  $d_{kj}(t)$  denotes the Euclidean distance between vehicle  $k$  and  $j$  at time slot  $t$ , and  $\Delta > 0$  is a guard factor.

#### E. Definitions of Throughput and Delay

Let  $L_k(T)$  be the number of packets received by the destination of flow  $k$ ,  $k \in \{1, 2, \dots, N\}$ , up to time  $T$ ; let  $\mathcal{D}_k(T)$  be the integrated delay of packets received by the destination of flow  $k$  up to time  $T$ . An asymptotic per-vehicle throughput  $\lambda(N)$  and average delay  $\mathcal{D}(N)$  are said feasible if there exist a scheduling policy and an  $N_0$  such that for any  $N > N_0$ , we have,

$$\lim_{T \rightarrow \infty} \Pr\left(\frac{L_k(T)}{T} \geq \lambda(N), \forall k\right) = 1, \quad (3)$$

$$\lim_{T \rightarrow \infty} \Pr\left(\frac{\mathcal{D}_k(T)}{L_k(T)} \leq \mathcal{D}(N), \forall k\right) = 1. \quad (4)$$

Specifically, an average per-vehicle throughput  $\tilde{\lambda}(N)$  is said feasible if there exist a scheduling policy and an  $N_0$ , such that for any  $N > N_0$ , the following holds

$$\lim_{T \rightarrow \infty} \Pr\left(\frac{\sum_{k=1}^N L_k(T)}{NT} \geq \tilde{\lambda}(N)\right) = 1. \quad (5)$$

### IV. ASYMPTOTIC CAPACITY AND DELAY ANALYSIS

In this section, we first propose a two-hop relay scheme for the packet delivery between the source and the destination. After that, we derive the asymptotic capacity and delay of the network with the two-hop relay scheme deployed.

#### A. Two-hop Relay Scheme

Let packets be delivered by using a two-hop relay scheme  $\mathcal{X}$ : a packet is either transmitted directly from the source to the destination, or relayed through one intermediate vehicle from the source to the destination. The packet transmission consists of two phases:

$\mathcal{X}$ -I: Each road segment in the grid becomes "active" in every  $1/p_{ac}$  time slots<sup>2</sup>.

<sup>2</sup>A road segment is active when vehicles on the road segment can transmit successfully without any interference of transmissions from other road segments. The value of  $p_{ac}$  is discussed later in the section.

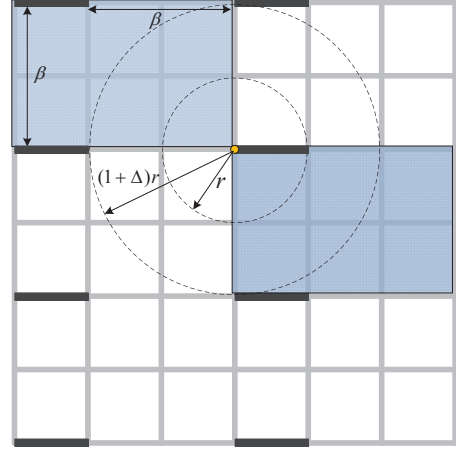


Fig. 4. An example of non-interfering transmission group of road segments.

$\mathcal{X}$ -II: For each active road segment where there are at least two vehicles,

1) if there exists at least one source-destination ( $\mathbb{S}$ - $\mathbb{D}$ ) pair on the road segment, one pair is uniformly selected. If the source has a buffering packet for transmission to the destination, it transmits the packet and evicts it from the buffer after the transmission; otherwise, the source stays idle.

2) if there is no any  $\mathbb{S}$ - $\mathbb{D}$  pair on the road segment, a vehicle, e.g.,  $v_A$ , is uniformly selected out of all vehicles on this road segment, and in the meantime another vehicle, e.g.,  $v_B$ , is independently and uniformly selected over the rest of vehicles. The following two actions are conducted equally likely:

- Let  $v_A$  be the transmitter and  $v_B$  be the receiver, and a source-to-relay transmission from  $v_A$  to  $v_B$  is scheduled. If  $v_A$  has a buffering packet to transmit,  $v_A$  transmits the packet to  $v_B$  and evicts the packet from the buffer; otherwise,  $v_A$  remains idle.
- Let  $v_B$  be the transmitter and  $v_A$  be the receiver, and a relay-to-destination transmission from  $v_B$  to  $v_A$  is scheduled. If  $v_B$  has a buffering packet destined for  $v_A$ ,  $v_B$  transmits the packet to  $v_A$  and evicts the packet from the buffer; otherwise,  $v_B$  remains idle.

In what follows, we evaluate the value of  $p_{ac}$ . As shown in Fig. 4, we partition the grid into equal-size sub-areas. Each sub-area consists of  $\beta(\beta + 1)$  squares where  $\beta$  is an integer number. The road segments highlighted in each sub-area in Fig. 4 constitute one non-interfering transmission group, such that simultaneous transmissions within one non-interfering group do not interfere with each other. Totally, there are  $2\beta(\beta + 1)$  road segments within one sub-area, and collectively  $2\beta(\beta + 1)$  non-interfering groups over the grid. With non-interfering groups transmitting iteratively, each non-interfering group becomes active every  $1/p_{ac} = 2\beta(\beta + 1)$  time slots. This indicates that the vehicles on one specific road segment obtain a transmission opportunity with probability  $p_{ac}$  at a randomly selected time slot. With the grid scale of  $M$ , the minimum distance between any two neighboring road segments of a non-interfering group is  $\frac{\beta}{M-1}$ . With the protocol model applied, we

have

$$\beta/(M-1) \geq (1+\Delta)r.$$

With  $r = 1/(M-1)$ , we have

$$\beta \geq 1 + \Delta.$$

We set  $\beta = \lceil 1 + \Delta \rceil$ , where  $\lceil x \rceil$  represents the smallest integer greater than or equal to  $x$ . By substituting it into  $1/[2\beta(\beta+1)]$ , we have

$$p_{ac} = 1/(2\lceil 1 + \Delta \rceil \lceil 2 + \Delta \rceil).$$

In what follows, we derive the throughput capacity and average packet delay with the two-hop relay scheme  $\mathcal{X}$ .

### B. Bounds of per-vehicle throughput capacity

We first derive an upper bound of per-vehicle throughput considering any possible stabilizing scheduling policies under  $\mathcal{X}$ -I. Let  $X_d(T)$  denote the number of packets delivered in the network through direct transmissions from the source to destination, and  $X_r(T)$  denote the number of packets delivered to the destination via relaying, during the interval  $[0, T]$ . Therefore, provided the arbitrary and fixed  $\epsilon > 0$ , there must exist arbitrarily large values of  $T$  such that the per-vehicle throughput  $\lambda(N)$  satisfies

$$\frac{X_d(T) + X_r(T)}{T} \geq N\lambda(N) - \epsilon. \quad (6)$$

Let  $Y(T)$  denote the total number of transmission opportunities during the interval  $[0, T]$ . From (6), we have

$$\begin{aligned} \frac{1}{T}Y(T) &\geq \frac{1}{T}X_d(T) + \frac{2}{T}X_r(T) \\ &\geq \frac{1}{T}X_d(T) + 2((N\lambda(N) - \epsilon) - \frac{1}{T}X_d(T)). \end{aligned}$$

The first inequality holds because the relayed packet reaches to the destination through at least two hops. And therefore,

$$\lambda(N) \leq \frac{\frac{1}{T}Y(T) + \frac{1}{T}X_d(T) + 2\epsilon}{2N},$$

i.e.,

$$\lambda(N) \leq \lim_{T \rightarrow \infty} \frac{\frac{1}{T}Y(T) + \frac{1}{T}X_d(T)}{2N}. \quad (7)$$

Due to the interference of transmissions, the total number of transmission opportunities is no larger than the maximum number of concurrent transmissions during  $[0, T]$ . By the law of large numbers, we have  $\lim_{T \rightarrow \infty} \frac{1}{T}Y(T) \leq Gp_{ac}$ . Similarly, we have  $\lim_{T \rightarrow \infty} \frac{1}{T}X_d(T) \leq Gp_{ac}$ , where the equality holds when there is always an  $\mathbb{S}$ - $\mathbb{D}$  transmission on each road segment of a non-interference group during each time slot. By plugging the inequalities into (7), we have

$$\lambda(N) \leq \frac{Gp_{ac} + Gp_{ac}}{2N} = \frac{p_{ac}}{d} = \frac{1}{2d\lceil 1 + \Delta \rceil \lceil 2 + \Delta \rceil}.$$

To derive a lower bound of per-vehicle throughput capacity, we start from the following lemma.

*Lemma 1:* Let  $N_i$  denote the number of vehicles whose

mobility region contains road segment  $i$ .  $N_i$  at most scales as  $O(\log(N))$  w.h.p.<sup>3</sup>.

The detailed proof is provided in Appendix.

Using the two-hop relay scheme  $\mathcal{X}$ , a destination vehicle can successfully receive a packet only when the following three events occur at the same time: i) the road segment in which the destination vehicle locates is active; ii) there exists at least one other vehicle on that road segment; and iii) the destination is selected as the receiver in either the direct source-to-destination transmission or relay-to-destination transmission. The first event occurs with probability  $p_{ac}$ . The second event occurs with a non-zero probability. From Lemma 1, the probability of occurrence of the third event is at least  $\Omega(\frac{1}{\log(N)})$  w.h.p.. Therefore, for a given  $\mathbb{S}$ - $\mathbb{D}$  pair, the throughput  $\Omega(\frac{1}{\log(N)})$  is feasible.

By summarizing the above analysis, we have the following theorem.

*Theorem 1:* For the social-proximity grid-like urban networks, with the two-hop relay scheme  $\mathcal{X}$ , the per-vehicle throughput  $\lambda(N)$  cannot be better than  $\frac{1}{2d\lceil 1 + \Delta \rceil \lceil 2 + \Delta \rceil}$  and w.h.p., scales as  $\Omega(\frac{1}{\log(N)})$ .

### C. Average per-vehicle throughput

In this part, we derive the lower bound of the average per-vehicle throughput  $\bar{\lambda}(N)$  based on the proposed two-hop relay scheme. We need the following lemmas for the proof of Theorem 2, which is stated later in the section.

*Lemma 2:* (Chebyshev's Inequality) If  $X$  is a random variable with mean  $E[X]$  and variance  $\text{Var}(X)$ , then, for any value  $k > 0$ ,

$$\Pr(|X - E[X]| \geq k) \leq \frac{\text{Var}(X)}{k^2}$$

Lemma 2 is well known and has been proved in the literature. We make use of it to prove Lemma 3.

*Lemma 3:* At least  $(1 - e^{-d})C$  squares will be a social spot of at least one  $\mathbb{S}$ - $\mathbb{D}$  pair w.h.p..

*Proof:* Let  $I$  be the number of squares that are not chosen as a *Tier*(1) (indicating the social spot) of any vehicle's mobility region. We define  $I_i$  as an indicator variable for all  $i \in \{1, 2, \dots, C\}$ ,

$$I_i = \begin{cases} 1 & \text{if } \forall n \in \{1, 2, \dots, N\}, H_n \neq i \\ 0 & \text{otherwise} \end{cases}$$

and  $\Pr(I_i = 1) = (1 - \frac{1}{C})^{\frac{N}{d}}$ . Thus, the expectation and variance of  $I_i$  are  $E[I_i] = (1 - \frac{1}{C})^{\frac{N}{d}}$  and  $\text{Var}(I_i) = (1 - \frac{1}{C})^{\frac{N}{d}} - (1 - \frac{1}{C})^N$ . Next we need to determine the variance of  $I$ . For any  $i \neq j$ ,  $j \in \{1, 2, \dots, C\}$ ,  $\text{Cov}(I_i, I_j) = E[I_i I_j] - E[I_i]E[I_j]$ , where  $\text{Cov}(I_i, I_j)$  is the covariance of variable  $I_i$  and  $I_j$ . It is easy to get that  $E[I_i I_j] = (1 - \frac{2}{C})^{\frac{N}{d}}$ . Since

<sup>3</sup>As  $N \rightarrow \infty$ , the probability approaches to 1.

$\text{Cov}(I_i, I_i) = \text{Var}(I_i)$ , we have

$$\begin{aligned}
\text{Var}(I) &= \text{Var}\left(\sum_{i=1}^C I_i\right) = \sum_{i=1}^C \sum_{j=1}^C \text{Cov}(I_i, I_j) \\
&= \sum_{i=1}^C \text{Cov}(I_i, I_i) + 2 \sum_{i=1}^C \sum_{j < i}^C \text{Cov}(I_i, I_j) \\
&= C \left[ \left(1 - \frac{1}{C}\right)^{\frac{N}{2}} - \left(1 - \frac{1}{C}\right)^N \right] \\
&\quad + C(C-1) \left[ \left(1 - \frac{2}{C}\right)^{\frac{N}{2}} - \left(1 - \frac{1}{C}\right)^N \right] \\
&\leq C \left[ \left(1 - \frac{1}{C}\right)^{\frac{N}{2}} - \left(1 - \frac{1}{C}\right)^N \right].
\end{aligned}$$

The inequality holds because  $\left(1 - \frac{2}{C}\right)^{\frac{N}{2}} - \left(1 - \frac{1}{C}\right)^N = \left(1 - \frac{2}{C}\right)^{\frac{N}{2}} - \left(1 - \frac{2}{C} + \frac{1}{C^2}\right)^{\frac{N}{2}} \leq 0$ . From Lemma 2, choosing  $k = \epsilon C$ , we have

$$\Pr(I - E[I] \geq \epsilon C) \leq \frac{C \left[ \left(1 - \frac{1}{C}\right)^{\frac{N}{2}} - \left(1 - \frac{1}{C}\right)^N \right]}{\epsilon^2 C^2}.$$

Note that  $E[I] = C \left(1 - \frac{1}{C}\right)^{\frac{N}{2}}$ . Thus,

$$\Pr\left(\frac{I}{C} \geq (\rho + \epsilon)\right) \leq \frac{\rho - \rho^2}{\epsilon^2} \cdot \frac{1}{C},$$

where  $\rho = \left(1 - \frac{1}{C}\right)^{\frac{N}{2}}$ . Since  $N = 2dC$ , as  $N \rightarrow \infty$ ,  $\rho \rightarrow e^{-d}$ . Therefore,  $\lim_{N \rightarrow \infty} \Pr(I/C \geq e^{-d}) = 0$ , i.e., the probability of  $I$  being over a constant proportion of  $C$  goes to zero as  $N \rightarrow \infty$ . Consequently, at least  $(1 - e^{-d})C$  squares will be a social spot of at least one  $\mathbb{S}\text{-}\mathbb{D}$  pair *w.h.p.* ■

Before introducing the next lemma, we define two summations of probabilities:  $p(N) = \frac{1}{G} \sum_{i=1}^G \Pr[\text{finding at least two vehicles on road segment } i \text{ during a time slot}]$  and  $q(N) = \frac{1}{G} \sum_{i=1}^G \Pr[\text{finding an } \mathbb{S}\text{-}\mathbb{D} \text{ pair on road segment } i \text{ during a time slot}]$ . Note that  $p(N)$  and  $q(N)$  are random variables because of randomness in the locations of vehicles' social spots, and all the probabilities above are associated with the steady-state location distribution of each vehicle.

**Lemma 4:**  $p(N)$  and  $q(N)$  scale as  $\Theta(1)$  *w.h.p.*. Specifically, *w.h.p.*,  $(1 - e^{-d})\pi_1^2 \leq \liminf_{N \rightarrow \infty} p(N) \leq 1$  and  $(1 - e^{-d})\pi_1^2 \leq \liminf_{N \rightarrow \infty} q(N) \leq 1$ .

*Proof:* Let  $p_i$  denote the probability of finding at least two vehicles on road segment  $i$  during a time slot. We can find that road segment  $i$  belongs to the *Tier(1)* of 2 different mobility regions (specified by different social spots), the *Tier(2)* of 10, ..., and the *Tier( $\alpha$ )* of  $8\alpha - 6$ ,  $\alpha \in \{1, 2, \dots, \mathcal{A}\}$ , as shown in Fig. 5.

Let  $N_i^\alpha$  denote the number of vehicles whose *Tier( $\alpha$ )* of the mobility region contains road segment  $i$ . Thus, the probability of finding no any vehicle on road segment  $i$  during a time slot is  $\prod_{\alpha=1}^{\mathcal{A}} (1 - \pi_\alpha)^{N_i^\alpha}$ . And the probability of finding exact one vehicle on road segment  $i$  during a time slot is  $\sum_{\alpha=1}^{\mathcal{A}} (N_i^\alpha \pi_\alpha (1 - \pi_\alpha)^{N_i^\alpha - 1} \prod_{\alpha' = 1, \neq \alpha}^{\mathcal{A}} (1 - \pi_{\alpha'})^{N_i^{\alpha'}})$ . Therefore, we have,

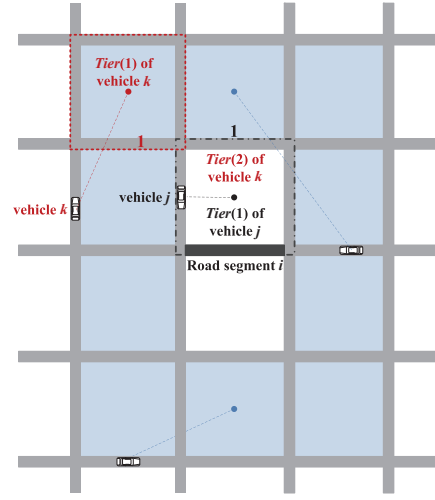


Fig. 5. An example of one given road segment contained by different vehicles' mobility region.

$$\begin{aligned}
p_i &= 1 - \prod_{\alpha=1}^{\mathcal{A}} (1 - \pi_\alpha)^{N_i^\alpha} \\
&\quad - \sum_{\alpha=1}^{\mathcal{A}} \left( N_i^\alpha \pi_\alpha (1 - \pi_\alpha)^{N_i^\alpha - 1} \prod_{\alpha' = 1, \neq \alpha}^{\mathcal{A}} (1 - \pi_{\alpha'})^{N_i^{\alpha'}} \right).
\end{aligned}$$

According to Lemma 3, the proportion of squares, each of which is the social spot of at least one  $\mathbb{S}\text{-}\mathbb{D}$  pair, is at least  $1 - e^{-d}$  *w.h.p.*. Further, we can infer that *w.h.p.*, there are at least  $2(1 - e^{-d})C$  road segments, each of which belongs to a *Tier(1)* chosen as a social spot. Let  $\mathcal{S}$  denote the set of road segments that are not contained in the mobility region of any vehicle.  $\bar{\mathcal{S}}$  is the complementary set of  $\mathcal{S}$  in  $\{1, 2, \dots, G\}$ . Note that  $N_i^\alpha$  is even and if  $N_i^\alpha = 0$ , for all  $\alpha$ ,  $p_i = 0$ . Recall that  $p(N) = \frac{1}{G} \sum_{i=1}^G p_i$ . We have,

$$\begin{aligned}
p(N) &= \frac{1}{G} \left( \sum_{i \in \mathcal{S}} p_i + \sum_{j \in \bar{\mathcal{S}}} p_j \right) \\
&= \frac{1}{G} \left( |\mathcal{S}| \cdot 0 + \sum_{j \in \bar{\mathcal{S}}} p_j \right) \geq \frac{1}{G} \sum_{j \in \bar{\mathcal{S}}} \pi_1^2,
\end{aligned}$$

since for any  $j \in \bar{\mathcal{S}}$ ,  $p_j \geq \pi_1^2$ . From Lemma 3, *w.h.p.*,

$$\begin{aligned}
\liminf_{N \rightarrow \infty} p(N) &\geq \liminf_{N \rightarrow \infty} \frac{1}{G} \pi_1^2 \cdot 2(1 - e^{-d})C \\
&= (1 - e^{-d})\pi_1^2.
\end{aligned}$$

Let  $M_i^\alpha$  denote the number of  $\mathbb{S}\text{-}\mathbb{D}$  pairs whose *Tier( $\alpha$ )* of the mobility region contains road segment  $i$ . Thus,

$$q(N) = \frac{1}{G} \sum_{i=1}^G q_i = \frac{1}{G} \sum_{i=1}^G \left( 1 - \prod_{\alpha=1}^{\mathcal{A}} (1 - \pi_\alpha^2)^{M_i^\alpha} \right).$$

Similarly, we can prove that *w.h.p.*,

$$\liminf_{N \rightarrow \infty} q(N) \geq (1 - e^{-d})\pi_1^2.$$

■

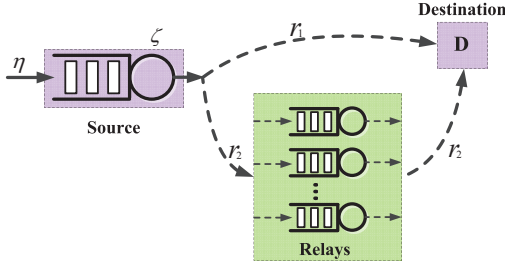


Fig. 6. A decoupling queue structure.

*Theorem 2:* For the social-proximity grid-like urban networks, with the two-hop relay scheme  $\mathcal{X}$ , the average per-vehicle throughput  $\tilde{\lambda}(N)$  scales as  $\Theta(1)$  w.h.p.. Specifically, w.h.p.,  $\liminf_{N \rightarrow \infty} \tilde{\lambda}(N) \geq \frac{1}{2d[1+\Delta][2+\Delta]}$ .

*Proof:* Based on the two-hop relay scheme  $\mathcal{X}$ , we are able to use a decoupling queue structure, similar to that in [5], to model each unicast flow, as shown in Fig. 6. Without loss of generality, we consider that the packet arrival rate  $\eta$  follows the Bernoulli process; that is to say, in each unicast flow, one packet arrives with the probability  $\eta$  at the current slot, and with the rest probability there is no packet arrival. Therefore, the source vehicle, e.g.,  $v_k$ , can be represented as a Bernoulli/Bernoulli queue with packet arrival rate  $\eta_k$  and service rate  $\zeta_k$ . The buffering packet in the source will be transmitted (served) to either its destination directly or one of the relays within the mobility region of the source. The transmission opportunity arises with probability  $\zeta_k$ .

Let  $r_1^k(N)$  denote the long term average rate at which a direct transmission to the destination is scheduled to source  $v_k$ , and  $r_2^k(N)$  denote the long term average rate at which a source-to-relay transmission is scheduled to source  $v_k$ . The transmission opportunity arises at the rate  $\zeta_k(N) = r_1^k(N) + r_2^k(N)$ . As per the definition,  $\tilde{\lambda}(N) = \frac{\sum_{k=1}^N \zeta_k(N)}{N}$ . Since the two-hop relay scheme  $\mathcal{X}$  schedules a source-to-relay transmission and a relay-to-destination transmission equally likely, the rate into the relays is equal to the rate out of the relays. During each time slot, the total number of transmission opportunities over the network is  $\sum_{k=1}^N (r_1^k(N) + 2r_2^k(N))$ . Given that the transmission opportunity arises on a road segment when it is active and at least two vehicles are on it, then we have,

$$G_{pac}p(N) = \sum_{k=1}^N (r_1^k(N) + 2r_2^k(N)). \quad (8)$$

Since the two-hop relay scheme  $\mathcal{X}$  schedules the source-to-destination transmission whenever possible and with probability  $p_{ac}q(N)$  there is a source-to-destination transmission occurs on a given road segment, then we have,

$$G_{pac}q(N) = \sum_{k=1}^N r_1^k(N). \quad (9)$$

From (8) and (9), we obtain  $\sum_{k=1}^N r_2^k(N) = \frac{G_{pac}(p(N)-q(N))}{2}$  and therefore

$$\tilde{\lambda}(N) = \frac{\sum_{k=1}^N (r_1^k(N) + r_2^k(N))}{N} = \frac{p_{ac}(p(N) + q(N))}{2d}.$$

According to Lemma 4, w.h.p., we have,

$$\liminf_{N \rightarrow \infty} \tilde{\lambda}(N) \geq \frac{p_{ac}(1 - e^{-d})\pi_1^2}{d} \quad (10)$$

which is the lower bound on the average per-vehicle throughput capacity. ■

*Remark:* The average per-vehicle throughput is analyzed as a global metric to evaluate the performance of the network with inhomogeneous vehicle density. From Theorem 2, we can infer that the constant per-vehicle throughput is feasible w.h.p. for  $N_f$   $\mathbb{S}$ - $\mathbb{D}$  pairs, where  $N_f = \Theta(N) \leq \frac{N}{2}$ . Because of vehicles' mobility with social features and the randomness of the locations of vehicle's social spots, the network shows spatial variations of vehicle density. In some hot area, i.e., the agglomeration of a large number of vehicles' social spots, the throughput of the  $\mathbb{S}$ - $\mathbb{D}$  pair in that area may drop to  $\Theta(\frac{1}{\log(N)})$ .

#### D. Average packet delay

We first analyze the average packet delay of a given unicast flow. The packet delay is accounted starting from the time slot when the packet arrives at the source until the time slot when the packet is delivered to its destination (including the queuing delay at the source or relay vehicle).

Recall that the source  $v_k$  can be represented as a Bernoulli/Bernoulli queue with arrival rate  $\eta_k$  and service rate  $\zeta_k$ . The expected number of packets buffered at the source is

$$E_k[n_s] = \frac{\eta_k(1 - \eta_k)}{\zeta_k - \eta_k}. \quad (11)$$

It is shown in [5] that packets depart from the source at the rate of  $\eta_k$  when the buffer of source is stable. For a packet sent out from the source, it is delivered to a relay vehicle, e.g.,  $v_i$ , with the probability  $\frac{r_2^k}{\zeta_k} \cdot P_{ki}$ , where  $P_{ki}$  is the contact probability between  $v_k$  and  $v_i$ . Therefore, the packet arrival rate to the relay  $v_i$  is  $\eta_{ki} = \frac{\eta_k r_2^k}{\zeta_k} P_{ki}$ . The packets depart to the destination from the relay  $v_i$  at the rate  $\zeta_{ki} = r_2^k P_{ki}$ . This is because that the source and the destination have the equal contact probability with the relay vehicles, and moreover the packet injection rate from the source to the relays equals to that from the relays to the destination, as shown in Fig. 6. With the packet arrivals and departures at the relay  $v_i$  following the Bernoulli process with mean rates  $\eta_{ki}$  and  $\zeta_{ki}$ , respectively, the average number of packets hold by  $v_i$  is

$$E_{ki}[n_r] = \frac{\eta_{ki}}{\zeta_{ki} - \eta_{ki}} = \frac{\eta_k}{\zeta_k - \eta_k}. \quad (12)$$

Note that (12) holds for every relay. From Little's law, the average packet delay of the flow from  $v_k$  is

$$\mathcal{D}_k(N) = \frac{E_k[n_s] + R_k(N)E_{ki}[n_r]}{\eta_k} = \frac{R_k(N) + 1 - \eta_k}{\zeta_k - \eta_k},$$

where  $R_k(N)$  is the total number of relay vehicles which have an overlapped mobility region with source  $v_k$ . As indicated by 13, the average packet delay is dependent of vehicle density in the proximity region of a unicast flow.

We proceed to derive the lower bound and upper bound of the average packet delay of the entire network. We neglect the queuing delay at the source vehicle in the calculation, as we are only interested in the delay caused by vehicles' mobility.

*Theorem 3:* For the social-proximity grid-like urban networks, with the two-hop relay scheme  $\mathcal{X}$ , the average packet delay  $\mathcal{D}(N)$  cannot be less than  $\frac{1}{\sum_{\alpha=1}^A (16\alpha-12)\alpha^{-2\gamma}\pi_1^2}$  and *w.h.p.*, scales as  $O(\log^2(N))$ .

*Proof:* The minimal delay of a flow is achieved when the source delivers the flow packets to its destination with the highest transmission priority. Moreover, the direct packet transmission from the source to the destination has lower average delay compared to the relay transmissions, with the condition that the contact probability between the source and one of its relay vehicles is no larger than the contact probability between the source and its destination. The source encounters the destination on a same road segment with the probability

$$P_{SD} = \sum_{\alpha=1}^A (16\alpha - 12)\alpha^{-2\gamma}\pi_1^2.$$

Therefore, the minimum packet delay is geometric distributed with mean  $1/P_{SD}$ .

Next we prove the upper bound on the average packet delay. Let  $v_A$  and  $v_B$  be a given pair of vehicles whose mobility regions are overlapped.  $v_A$  intends to transmit a packet to  $v_B$ . The transmission between  $v_A$  and  $v_B$  can be scheduled during a time slot only when the following three events occur at the same time: i)  $v_A$  and  $v_B$  are located on a same road segment during the time slot; ii) that road segment is active in the slot; and iii)  $v_A$  and  $v_B$  are both selected for a transmission from  $v_A$  to  $v_B$ . These three events occur with probability  $\phi_1$ ,  $p_{ac}$  and  $\phi_2$  respectively. Thus, the distribution of the packet delay between  $v_A$  and  $v_B$  can be treated as geometric with mean

$$\frac{1}{\phi_1\phi_2p_{ac}} = O(\log^2(N)),$$

where  $\phi_1$  is deterministic and none-zero and  $\phi_2$  is  $\Omega(\frac{1}{\log^2(N)})$  *w.h.p.* according to Lemma 1. Hence, the average packet delay  $\mathcal{D}(N)$  scales as  $O(\log^2(N))$  *w.h.p.* ■

### E. Discussion

1) *With or without packet redundancy:* With the proposed two-hop relay scheme applied, a constant average per-vehicle throughput is achievable *w.h.p.*. Note that the two-hop relay scheme does not use any packet redundancy, i.e., there is only one copy of each packet in the network. In fact, the same bounds on asymptotic throughput capacity and delay can be derived even when direct transmissions are performed between the source and destination without the assistance of any relays. This is because that the source and the destination have

non-zero contact probability and the number of concurrent transmissions scales as  $\Theta(N)$  *w.h.p.*. Therefore, we are able to use a two-hop relay scheme with packet redundancy to improve the delay performance, however, without degrading throughput in the order sense.

A better delay performance can be achieved with an improved scheduling policy. Note that the minimum delay is achieved by the direct transmission with highest schedule priority, since the source is more likely to encounter the destination than the relays (except that the relay has the same social spot as the  $\mathbb{S}$ - $\mathbb{D}$  pair). In this case, the relay may not be helpful to reduce the delay: it can delivery the packet from the source to the destination, but may incur a longer delay than the direct transmission. It is possible to conduct a source-guaranteed redundant scheduling: the source keeps a copy of each packet delivered to the relays and confirms the receipt of such packet when the source encounters the destination. If such packet has not yet been delivered to the destination, the source can transmit the packet directly. It can be envisioned that in social-proximity VANETs the throughput-sensitive applications can be supported by the two-hop relay scheme without redundancy, and the delay-sensitive applications can be better supported by the direct transmission or source-guaranteed redundant scheme. The analysis of the redundant scheduling will be considered as future work and is out the scope of this paper.

2) *Partially overlapped mobility region:* In this paper, we consider that the source and the destination have a completely overlapping mobility region, or equivalently, the source and destination have the same social spot. Therefore, all the relays contacting the source would also contact the destination. In a more general case that the source and destination are associated with different social spots, not all vehicles can be relays because vehicles may have disjoint mobility regions with the destination. In this scenario, we acknowledge that the two-hop relay scheme is no longer efficient unless a more intelligent relay selection scheme is applied.

3) *Vehicle density:* The throughput and average packet delay of a unicast flow highly depend on the vehicle density. For a large scale urban VANET, there are spatial variations in vehicle density. Vehicle density in hot social spots would be much higher than that in normal places, and accordingly each vehicle may have less opportunity to transmit packets because of potentially severe medium access contentions among vehicles. Under such circumstance, the performance in terms of throughput and delay degrades. Therefore, it may not be feasible to conduct excessive unicast flows through VANET in urban area with high vehicle density.

## V. CONCLUSION

In this paper, we have investigated the asymptotic capacity and delay performance for social-proximity urban vehicular networks. We consider a localized mobility model centering at a social spot for each vehicle. The user applications are of proximity nature, i.e., the source and the destination have the identical social spot. With the proposed two-hop relay



scheme, a constant average per-vehicle throughput can be achieved *w.h.p.*; although the throughput and average packet delay of a unicast flow highly depend on the vehicle density, almost constant per-vehicle throughput  $\Omega(\frac{1}{\log(N)})$  and almost constant delay  $O(\log^2(N))$  are feasible even in some area with high vehicle density. Our results reveal that the social-proximity vehicular network is scalable to be deployed in urban environments.

#### APPENDIX PROOF OF LEMMA 1

We first recall the Vapnik-Chervonenkis Theorem [22]. Some relevant definitions are in the following. A Range Space is a pair  $(X, \mathcal{F})$ , where  $X$  is a set and  $\mathcal{F}$  is a family of subsets of  $X$ . For any  $A \subseteq X$ , we define  $P_{\mathcal{F}}(A)$ , the projection of  $\mathcal{F}$  on  $A$ , as  $\{F \cap A : F \in \mathcal{F}\}$ . We say that  $A$  is *shattered* by  $\mathcal{F}$  if  $P_{\mathcal{F}}(A) = 2^A$ , i.e., if the projection of  $\mathcal{F}$  on  $A$  includes all possible subsets of  $A$ . The VC-dimension of  $\mathcal{F}$ , denoted by  $\text{VC-d}(\mathcal{F})$  is the cardinality of the largest set  $A$  that  $\mathcal{F}$  shatters. If arbitrarily large finite sets are shattered, the VC dimension of  $\mathcal{F}$  is infinite.

*The Vapnik-Chervonenkis Theorem:* If  $\mathcal{F}$  is a set of finite VC-dimension and  $\{Y_j\}$  is a sequence of  $N$  i.i.d. random variables with common probability distribution  $P$ , then for every  $\epsilon, \delta > 0$

$$\Pr \left\{ \sup_{F \in \mathcal{F}} \left| \frac{1}{N} \sum_{j=1}^N I(Y_j \in F) - P(F) \right| \leq \epsilon \right\} > 1 - \delta \quad (13)$$

whenever

$$N > \max \left\{ \frac{8\text{VC-d}(\mathcal{F})}{\epsilon} \log_2 \frac{16e}{\epsilon}, \frac{4}{\epsilon} \log_2 \frac{2}{\delta} \right\}. \quad (14)$$

Here  $I(Y_j \in F)$  is the indicator variable that takes value 1 if  $Y_j \in F$  and 0 otherwise.

*Proof of Lemma 1:* We use the Vapnik-Chervonenkis Theorem to prove this lemma. We denote  $F$  as the rectangular area of  $2\mathcal{A}(2\mathcal{A}-1)$  squares centered at a given road segment, as shown in Fig. 4. If a vehicle chooses a square within  $F$  as its *Tier*(1) region, its mobility region will cover the road segment at the center of  $F$ . Let  $I(Y_j \in F)$  be 1 if the *Tier*(1) region of vehicle  $j$  falls into  $F$  and 0 otherwise.  $\Pr(I(Y_j \in F) = 1) = \frac{2\mathcal{A}(2\mathcal{A}-1)}{C}$ . Let  $\mathcal{F}$  be the class of all such  $F$  rectangular areas. It is easy to show that the VC-dimension of  $\mathcal{F}$  is at most 4 [23]. Therefore, for all rectangular area  $F$ ,

$$\Pr \left\{ \sup_{F \in \mathcal{F}} \left| \frac{\sum_{j=1}^N I(Y_j \in F)}{N} - \frac{2\mathcal{A}(2\mathcal{A}-1)}{C} \right| \leq \epsilon \right\} > 1 - \delta.$$

The condition (14) holds when  $\epsilon = \delta = \frac{\Delta_\epsilon \log(N)}{N}$ , where  $\Delta_\epsilon := \max\{8\text{VC-d}(\mathcal{F}), 16e\}$ . Recall that  $C = N/2d$ . Thus, the Vapnik-Chervonenkis Theorem states that

$$\Pr \left\{ \sup_{F \in \mathcal{F}} \left\{ \sum_{j=1}^N I(Y_j \in F) \right\} \leq 4d\mathcal{A}(2\mathcal{A}-1) + \Delta_\epsilon \log(N) \right\} > 1 - \frac{\Delta_\epsilon \log(N)}{N}.$$

We conclude that *w.h.p.*, the number of vehicles whose mobility region contains a given road segment is at most  $O(\log(N))$ . ■

#### REFERENCES

- [1] F. Bai and B. Krishnamachari, "Exploiting the wisdom of the crowd: localized, distributed information-centric vanets [topics in automotive networking]," *Communications Magazine, IEEE*, vol. 48, no. 5, pp. 138–146, 2010.
- [2] T. Luan, L. Cai, J. Chen, X. Shen, and F. Bai, "VTube: Towards the Media Rich City Life with Autonomous Vehicular Content Distribution," in *SECON, 2011 Proceedings IEEE*, pp. 359–367.
- [3] R. Lu, X. Lin, T. Luan, X. Liang, and X. Shen, "Pseudonym changing at social spots: An effective strategy for location privacy in vanets," *Vehicular Technology, IEEE Transactions on*, to appear.
- [4] N. Sarafijanovic-Djukic, M. Pidrkowski, and M. Grossglauser, "Island hopping: Efficient mobility-assisted forwarding in partitioned networks," in *SECON, 2006 Proceedings IEEE*, pp. 226–235.
- [5] M. Neely and E. Modiano, "Capacity and delay tradeoffs for ad hoc mobile networks," *Information Theory, IEEE Transactions on*, vol. 51, no. 6, pp. 1917–1937, 2005.
- [6] P. Gupta and P. Kumar, "The capacity of wireless networks," *Information Theory, IEEE Transactions on*, vol. 46, no. 2, pp. 388–404, 2000.
- [7] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *Information Theory, IEEE Transactions on*, vol. 53, no. 3, pp. 1009–1018, 2007.
- [8] X. Li, "Multicast capacity of wireless ad hoc networks," *Networking, IEEE/ACM Transactions on*, vol. 17, no. 3, pp. 950–961, 2009.
- [9] P. Li, M. Pan, and Y. Fang, "The capacity of three-dimensional wireless ad hoc networks," in *INFOCOM, 2011 Proceedings IEEE*, pp. 1485–1493.
- [10] G. Alfano, M. Garetto, E. Leonardi, and V. Martina, "Capacity scaling of wireless networks with inhomogeneous node density: lower bounds," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 5, pp. 1624–1636, 2010.
- [11] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *Networking, IEEE/ACM Transactions on*, vol. 10, no. 4, pp. 477–486, 2002.
- [12] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks-part I: The fluid model," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2568–2592, 2006.
- [13] G. Sharma, R. Mazumdar, and N. Shroff, "Delay and capacity tradeoffs in mobile ad hoc networks: A global perspective," *Networking, IEEE/ACM Transactions on*, vol. 15, no. 5, pp. 981–992, 2007.
- [14] X. Lin, G. Sharma, R. Mazumdar, and N. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with brownian mobility," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2777–2784, 2006.
- [15] P. Li, Y. Fang, and J. Li, "Throughput, delay, and mobility in wireless ad hoc networks," in *INFOCOM, 2010 Proceedings IEEE*, pp. 1–9.
- [16] M. Garetto and E. Leonardi, "Restricted mobility improves delay-throughput tradeoffs in mobile ad hoc networks," *Information Theory, IEEE Transactions on*, vol. 56, no. 10, pp. 5016–5029, 2010.
- [17] H. Pishro-Nik, A. Ganz, and D. Ni, "The capacity of vehicular ad hoc networks," in *Allerton*, vol. 7, pp. 1156–1163.
- [18] M. Nekoui, A. Eslami, and H. Pishro-Nik, "Scaling laws for distance limited communications in vehicular ad hoc networks," in *Communications (ICC), 2008 IEEE International Conference on*, pp. 2253–2257.
- [19] C. Wang, C. Jiang, X. Li, S. Tang, and P. Yang, "General capacity scaling of wireless networks," in *INFOCOM, 2011 Proceedings IEEE*, pp. 712–720.
- [20] M. Neely, E. Modiano, and C. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 1, pp. 89–103, 2005.
- [21] R. Uргаonkar and M. Neely, "Network capacity region and minimum energy function for a delay-tolerant mobile ad hoc network," *Networking, IEEE/ACM Transactions on*, no. 99, pp. 1–1, 2011.
- [22] V. Vapnik and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory of Probability and its Applications*, vol. 16, p. 264, 1971.
- [23] V. Vapnik, *Statistical learning theory*. Wiley-Interscience, 1998.