

Distributed Routing for Vehicular Ad Hoc Networks: Throughput-Delay Tradeoff

Ali Abedi Majid Ghaderi Carey Williamson

Department of Computer Science, University of Calgary
{abedia, mghaderi, carey}@cpsc.ucalgary.ca

Abstract—In this paper, we address the problem of *low-latency routing* in a vehicular highway network. To cover long highways while minimizing the number of required roadside access points, we utilize vehicle-to-vehicle communication to propagate data in the network. Vehicular networks are highly dynamic, and hence routing algorithms that require global network state information or centralized coordination are not suitable for such networks. Instead, we develop a novel *distributed routing algorithm* that requires *minimal coordination* among vehicles, while achieving a highly efficient throughput-delay tradeoff. Specifically, we show that the proposed algorithm achieves a throughput that is within a factor of $1/e$ of the throughput of an algorithm that centrally coordinates vehicle transmissions in a highly dense network, and yet its end-to-end delay is approximately half of that of a widely studied ALOHA-based randomized routing algorithm. We evaluate our algorithm analytically and through simulations and compare its throughput-delay performance against the ALOHA-based randomized routing.

Index Terms—Vehicular ad hoc networks, Routing, Throughput, Delay.

I. INTRODUCTION

In vehicular ad hoc networks (VANETs), vehicles equipped with short-range radios communicate with each other and possibly with a roadside infrastructure to enable a range of applications from Internet access and driver assistance to transportation safety and emergency response. Although cellular-based systems are widely used in vehicles, mainly to report accidents in a timely manner (*e.g.*, OnStar [1]), emerging vehicular ad hoc networks enable inter-vehicle communication to offer a variety of services to end users at a lower cost.

To support emerging vehicular networks, *Dedicated Short Range Communication* (DSRC) [2] is specifically designed for communication between vehicles and roadside access points. In the US, the FCC has allocated 75 MHz of the spectrum in the 5.9 GHz band to DSRC. The MAC and physical layer specifications of DSRC are defined in IEEE 802.11p [3], which amends IEEE 802.11 for communication in outdoor high-speed vehicular environments, where the physical layer properties change rapidly and communication opportunities are transient and brief.

Although a VANET can be considered as a traditional mobile ad hoc network, there are several key differences. For example, the mobility pattern in vehicular networks is more structured and predictable, since vehicles are restricted to streets or highways. Also, the average speed of nodes in a VANET is typically higher compared to a general mobile

ad hoc network. Finally, power consumption is no longer a primary concern in a VANET, since the embedded wireless devices rely on (relatively unlimited) vehicle battery power.

Many applications, especially those related to safety, and soft real-time applications such as video streaming, require a low-latency network to operate properly. Most data dissemination schemes for wireless ad hoc networks cannot be used directly in vehicular networks due to the high mobility rate in VANETs, which causes frequent communication breakdowns and excessive delays.

Routing is the backbone of any data dissemination scheme in vehicular networks. There is a large body of research on routing in traditional ad hoc networks that can be potentially used in vehicular networks. In general, routing algorithms in ad hoc networks can be divided into two categories, namely static and dynamic routing. Static routing algorithms [4], [5] establish a path in the network *a priori* so that every node along the path knows the next hop toward the destination. In dynamic routing algorithms, on the other hand, there is no *a priori* established path. Instead, every node chooses the next hop toward the destination *opportunistically* [6]–[8]. Opportunistic routing takes advantage of the broadcast nature of the wireless channel to dynamically choose the best next hop¹ toward the destination. At each step of opportunistic routing, the current next hop broadcasts a packet to its neighborhood. Among all the nodes in the neighborhood that have correctly received the packet, one is selected by means of local cooperation to serve as the next hop. We refer to the set of potential next hops as *partners* and the newly selected next hop as a *relay*. We argue that due to the highly dynamic nature of VANETs, hop-by-hop opportunistic routing algorithms are more appropriate than end-to-end static routing algorithms.

An important step in opportunistic routing is the relay selection process, *i.e.*, the *relaying strategy*. Two opposing strategies have been proposed in the literature. The first strategy is *uncoordinated randomized relaying* (*RANDOM*), in which every partner individually decides whether to be a relay or not (*i.e.*, to forward or not) similar to the traditional ALOHA protocol. Hence, *no coordination* among partners is required. One such strategy, called *asynchronous transmission*, was recently proposed and analyzed by Johnson *et al.* [9] for

¹The best next hop is chosen with respect to some routing metric such as geographic distance to the destination.

data dissemination in vehicular networks. It can be shown that RANDOM achieves throughput that is within a factor of $1/e$ of that for a strategy that perfectly schedules all packets in the network; however, RANDOM suffers from long delays [9]. The second strategy is *coordinated geographic relaying (GEOGRAPHIC)*, in which the partner that is closest to the destination is selected as the relay. Obviously, this strategy requires knowledge of the position of every partner, and perfect coordination amongst them. A family of geographic routing algorithms exists in the literature [7], [8]. In general, GEOGRAPHIC minimizes the path length (*i.e.*, the distance packets travel to reach the destination), but suffers from low throughput and high end-to-end delay, since coordination amongst partners requires multiple rounds of message exchange to select a relay consensually. The relay selection is particularly challenging in VANETs due to rapid changes in network connectivity, unreliable nature of wireless channels, and the fact that some partners may not be within transmission range of each other.

In this paper, we develop a novel threshold-based relaying strategy called *uncoordinated optimized relaying (OPTIMIZED)*. OPTIMIZED achieves path length similar to GEOGRAPHIC and throughput similar to RANDOM, while requiring no coordination among partners. The basic idea is that, in a highway network, those partners (*i.e.*, potential relays) that are farther away from the transmitter should have a higher chance to become a relay and forward data. With the RANDOM strategy, relays are uniformly selected among partners, while in OPTIMIZED the probability is related to the distance from the transmitter. The relay selection still happens in a distributed and uncoordinated manner, which is desirable for vehicular environments.

Our contributions can be summarized as follows:

- 1) We develop a novel relaying strategy that is both throughput and delay efficient, while requiring no coordination for relay selection.
- 2) We build a mathematical model based on the theory of non-homogeneous Poisson processes to analyze the performance of the proposed relaying strategy.
- 3) We conduct simulations to investigate the performance of the proposed strategy, comparing it with a well-known randomized strategy.

The rest of this paper is organized as follows. In Section II, we briefly survey some well-known ad-hoc routing paradigms that might be used in vehicular ad-hoc networks. In Section III, we describe the system model considered in this paper along with our assumptions. Section IV is dedicated to our proposed relaying strategy and its analysis. Simulation results are presented in Section V, where we study the performance of the proposed strategy and validate the accuracy of our analysis. Finally, our conclusions as well as future research directions are discussed in Section VI.

II. RELATED WORK

To establish communication between vehicles and access points, traditional multi-hop routing techniques such as

AODV [10] could be used. However, their route discovery mechanisms impose significant overhead in highly dynamic vehicular networks. In particular, the discovered route between a source and destination breaks frequently because of the high mobility of nodes, and the ensuing flooding of new route discovery packets consumes time and network bandwidth.

In contrast, the opportunistic routing paradigm [11] exploits the dynamics of mobile ad hoc networks by choosing the next hop opportunistically after the transmission of the packet, rather than following a pre-established path as in AODV. After a packet is relayed, one node among the nodes that received the packet is selected as the forwarder. In opportunistic routing, however, nodes require some information about their 1-hop neighbors in order to decide on the next hop, so they frequently send out beacon messages announcing their current location, speed, and direction. With high mobility rates in a VANET, such information quickly becomes obsolete, necessitating more frequent beacon messages, and increasing the coordination overhead of the routing protocol [11], [12]. For instance, the ExOR opportunistic routing scheme [13] transmits batches of packets across each hop, rather than single packets, to amortize the coordination overhead. However, these schemes still cannot guarantee packet delivery, since topology changes between coordination phases are not detected. Moreover, in some opportunistic routing schemes, a node is selected as the next hop only if it can relay the packet to other potential forwarders. Although the node considers potential forwarders in its neighborhood (*i.e.*, based on recently received beacon messages), occasionally all of these nodes miss the relayed packet, leading to additional retransmission coordination overhead.

A closely-related problem that has received considerable attention in mobile ad-hoc networks is the *broadcast storm* problem [14], in which multiple forwarders try to relay a packet at almost the same time. This issue is especially important in vehicular ad-hoc networks when the node density is high. Several routing schemes have been proposed to manage the broadcast storm problem. One solution is to select the node farthest from the source in order to make the multi-hop forwarding more efficient [15]. To this end, different mechanisms have been proposed to bias the forwarding probability based on a node distance from the transmitter.

One such group of broadcast schemes adopts a probabilistic mechanism to select the next hop. Wisitpongphan *et al.* [14] propose three probabilistic schemes for rebroadcasting the received message. The weighted p -persistent broadcasting scheme employs a rebroadcasting probability function which is one at the transmission range and linearly decreases to reach zero at the transmitter position. A similar approach has been proposed in [16] to dynamically adjust backoff timers; in particular, the contention window is set inversely proportional to the distance from the source.

The idea of assigning a higher priority to the farthest node from the transmitter has also been utilized in time-slotted systems. The slotted 1-persistent and p -persistent broadcasting schemes proposed in [14] try to bias the rebroadcasting priority

towards more distant nodes. In these schemes, when a node receives a packet, it delays rebroadcasting so that farther nodes have a higher priority in forwarding the packet. A similar rebroadcasting scheme has been proposed in [17], which prioritizes the nodes based on their distance from the transmitter by delaying retransmissions.

Other schemes select the next hop deterministically using a next-hop selection phase. For instance, the Urban Multi-hop Broadcast (UMB) protocol [18] addresses the broadcast storm problem in vehicular networks using a coordination phase to choose the next hop. To reduce hop-count, this technique selects the vehicles farthest from the source as relay nodes, using the black-burst method. The main drawback of this scheme is the requirement for repeaters at intersections, making deployment challenging. The authors later addressed this issue in [19] by introducing a fully Ad-hoc Multi-hop Broadcast (AMB).

All of the above mentioned works employ *heuristic* forwarding policies to reduce delay and improve throughput. Our main contribution in this work is to mathematically characterize the *optimal forwarding strategy* for this family of opportunistic routing algorithms. Specifically, we prove that the optimal forwarding strategy is a *threshold-based step function* of the distance from the transmitter. We then develop a model based on the theory of Poisson Processes to find the optimal threshold that characterizes this step function. Our model is appreciably general and takes into consideration several important network parameters such as the vehicle density on the road and probabilistic packet reception. Using the insights obtained from our analysis, we then develop an optimal routing scheme (OPTIMIZED) that achieves throughput and delay comparable to fully coordinated opportunistic routing schemes without requiring any node coordination.

III. SYSTEM MODEL

In this section, we describe the system model and discuss the assumptions considered in this paper.

A. Network Model

Our network model consists of multiple roadside wireless access points (APs) installed along the highway, every few kilometers. The APs are connected to the Internet via some backbone network infrastructure (wired or wireless), as depicted in Figure 1. For modeling purposes, we assume that vehicles arrive at the highway according to a Poisson process, whose rate determines the density of vehicles on the highway. Once on the highway, the *freeway mobility model* [20] determines the speed of all vehicles based on a set of rules to emulate the movement of vehicles.

Since it is not possible to populate the entire highway with APs, multi-hop transmission is used between nodes to propagate packets to the appropriate roadside AP. A vehicle can communicate directly with a roadside AP, or with another nearby vehicle on the highway, as shown in Figure 1.

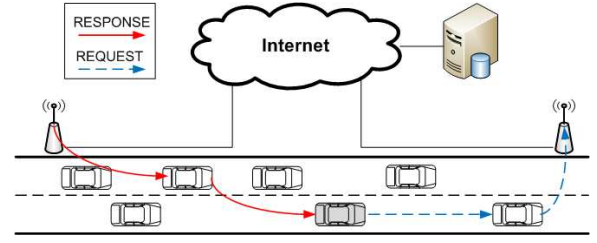


Figure 1. Network structure and query/response mechanism.

B. Application Model

A variety of applications are envisioned for VANETs [21]. In general, these applications are designed based on two paradigms, namely *push-based* and *pull-based* models, or a combination of the two models referred to as the *hybrid* model. In the push-based model, popular data such as traffic information is broadcast to vehicles on the road in a specific region. In contrast, in the pull-based model, a *query-response* model is employed, where a user queries for a specific information item on-demand. For instance, user applications such as web browsing and entertainment services have recently received considerable attention in VANETs [21], [22]. These applications require *low-latency* pull-based systems to operate effectively. For concreteness, in this paper, we focus on pull-based applications and cast our routing algorithm in the context of the underlying query/response mechanism, as shown in Figure 1.

C. Routing Model

If a vehicle is within the transmission range of an access point, then direct communication takes place between the vehicle and the access point. Otherwise, an opportunistic multi-hop routing is used to deliver packets from roadside APs to vehicles and vice versa. At each step of the multi-hop routing, *i.e.*, a *hop*, a vehicle (or an AP) broadcasts a packet to its neighborhood. Among the nodes that have successfully received the packet (*i.e.*, partners), one is selected to forward the packet next (*i.e.*, a relay). The mechanism that selects the relay is called the *relaying strategy*, which is the subject of Section IV.

D. Packet Reception Model

We consider two commonly used wireless channel models [23], namely the *Unit Disk Graph model* (UDG) and the *Log-Normal Shadowing model* (LNS). UDG is a simple model that assumes all receivers within the transmission range of the transmitter receive all packets successfully, implying a perfect circle as the reception area. LNS, on the other hand, is more realistic and has been widely used to model packet reception probability in wireless networks.

Let $f(x)$ denote the probability of receiving a packet correctly at distance x from the transmitter. LNS defines this function as follows [23]:

$$f(x) = \begin{cases} 1 - \frac{1}{2} \left(\frac{x}{R}\right)^{2\beta}, & 0 < x \leq R \\ \frac{1}{2} \left(\frac{2R-x}{R}\right)^{2\beta}, & R < x \leq 2R \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where β is the power attenuation factor, which typically varies between 2 and 6 depending on the propagation environment. For the highway environment, we set $\beta = 2$. The parameter R denotes the transmission range and is set to 1000 meters as the new IEEE 802.11p standard supports wireless communication up to one kilometer [9].

E. Mobility Model

The freeway mobility model [20] is adopted to emulate vehicle movements on a highway. This model can simulate a bidirectional multi-lane highway that gets the highway path from an input map. The advantage of this model is its consideration of the temporal, spatial, and geographical dependencies.

The key characteristics of our mobility model are:

- Vehicular traffic is unidirectional.
- Each vehicle is restricted to one lane.
- Vehicles enter the highway at the speed V_{min} .
- Vehicular speeds are adjusted dynamically with time using $V_i(t + \Delta t) = V_i(t) + a \text{rand}() \Delta t$, where $V_i(t)$ is the speed of vehicle i at time t and a is the acceleration factor. Function $\text{rand}()$ returns a uniformly distributed random number in the range $[-1, 1]$. The resulting speed is constrained to fall within the range (V_{min}, V_{max}) .
- If the distance between any two vehicles is less than a specified safety threshold distance, then the speed of the approaching vehicle is decreased to avoid a collision between the two vehicles. In our simulation, the safety distance is set to $2V_{max}$ meters, based on the two-second rule that helps drivers maintain a safe following distance.

F. Location Information

Throughout this work, we assume that location information is available to all vehicles in the network. Location information can be obtained from Global Positioning System (GPS) devices, which are embedded in many vehicles. In the absence of GPS, approximate location information can be estimated, for example, from the signal quality received at the destination. Although the estimation methods introduce error, we will show in Subsection V-E that the proposed strategy can tolerate estimation errors, with negligible performance degradation.

Several different distance estimation techniques have been proposed for ad hoc networks [24]. A widely used technique is based on *Received Signal Strength Indicator (RSSI)*, which provides passive estimation with no extra communication overhead. This technique uses the signal attenuation models that relate signal propagation distance to signal strength received at the destination. To estimate distance in a highway environment based on RSSI, we can apply Frii's equation, which yields the relation $d = G\sqrt{P_t/P_r}$, where G is a constant that captures the effect of transmitting and receiving antennas for a given carrier frequency, P_t is the transmission power, and P_r is the received power at the receiver. We note that G can be estimated as a system calibration parameter, P_t is known as it is set by the access point, and P_r can be obtained from the RSSI measurements. Real RSSI measurements in a vehicular

network reported in [25] show a predictable relation between the measured RSSI and distance as the highway environment is mostly free of obstacles and other interference effects.

IV. UNCOORDINATED OPTIMIZED RELAYING

Consider the network state after a node² has broadcast a packet to its neighborhood. A relaying strategy specifies which node among the nodes that have received the packet (*i.e.*, partners) will relay the packet next.

A simple strategy is to choose a node randomly and uniformly from the set of partners [9]. This approach is reminiscent of the ALOHA random channel access protocol. As in the ALOHA protocol, three outcomes are possible: (1) if zero nodes decide to forward the packet, delivery fails; (2) if one node decides to forward the packet, delivery succeeds; and (3) if multiple nodes forward the packet at the same time, a wireless collision occurs, and packet delivery fails.

If there are N partners, in a distributed implementation of this simple strategy, each node will relay the packet with probability $p^* = 1/N$. This strategy asymptotically achieves a throughput that is within a factor of $1/e$ of a centralized strategy that perfectly schedules all packets in the network [9]. Although simple and robust to different network topologies and mobility patterns, this scheme is *oblivious* to node *locations* and hence is not optimal in terms of path length and end-to-end delay.

Ideally, a relaying strategy should take into consideration the location of potential relays, and select the relay that is *closest to the destination*. This is similar to what GEOGRAPHIC does by coordinating transmissions of the partners. Our goal is to design a relaying strategy that considers node positions while selecting a relay, and yet does not require any global or local coordination.

A. Optimal Relaying Strategy

Consider a node located at distance x from the current relay node. Let $p(x)$ ($0 \leq p(x) \leq 1$) denote the relaying strategy, *i.e.*, $p(x)$ specifies the probability that the node located at distance x relays the packet. Our goal is to find a relaying strategy $p(x)$, so that:

- 1) the average *distance* between two adjacent relay nodes is maximized, and,
- 2) the probability of having only *one* node relaying the packet is maximized.

Consequently, we need to understand the (geographic and statistical) distribution of the nodes receiving a forwarded packet correctly, based on the packet reception model used.

A vehicle at distance x from the current transmitter receives the packet successfully with probability $f(x)$, where $f(x)$ is the packet reception probability function as presented in Subsection III-D. Conditioned on having received the packet successfully, the vehicle at distance x will relay the packet with probability $p(x)$, where $p(x)$ is the relaying strategy.

²Throughout the paper, we use the term "node" to refer to a vehicle or a roadside access point.

Therefore, the probability that a vehicle at distance x relays the packet is given by $f(x)p(x)$.

For the sake of analysis, we assume that vehicles are distributed on the highway according to a Poisson process with rate λ . We will later, in Subsection V-D, investigate the impact of this assumption on the performance of our strategy using simulations. Using the theory of Poisson processes [26], the vehicles that will relay the packet form a non-homogeneous Poisson process with rate $\lambda(x) = \lambda f(x)p(x)$. In the following proposition, we outline a simple proof of this result for completeness.

Proposition 1. *Suppose that Π is a Poisson process over a line with rate λ . An Event of Π at location x is counted independently with probability $p(x)$. Then the new process of counted events denoted by Π_p forms a non-homogeneous Poisson process with rate $\lambda(x) = \lambda p(x)$.*

Proof: Define $N(a, b)$ ($a, b \geq 0$) as the number of counted events in the interval $(a, b]$. The counting process Π_p is a non-homogeneous Poisson process if it satisfies four conditions [27]:

- 1) $N(0, 0) = 0$
- 2) Number of events in non-overlapping intervals are independent
- 3) $\lim_{h \rightarrow 0} \frac{\mathbb{P}\{N(x+h, x)=1\}}{h} = \lambda(x)$
- 4) $\lim_{h \rightarrow 0} \frac{\mathbb{P}\{N(x+h, x)>1\}}{h} = 0$

Conditions 1, 2 and 4 are satisfied since the counted events are a subset of all events in Π , and all these conditions are true for a homogeneous Poisson process [27]. To show that condition 3 follows as well, define $M(a, b)$ ($a, b \geq 0$) as the number of Π 's events that occur in the interval $(a, b]$, in contrast with $N(a, b)$ that gives the number of *counted* events in the same interval. We obtain that

$$\begin{aligned} \mathbb{P}\{N(x+h, x) = 1\} &= \mathbb{P}\{M(x+h, x) = 1 \wedge N(x+h, x) = 1\} \\ &\quad + \mathbb{P}\{M(x+h, x) \geq 2 \wedge N(x+h, x) = 1\}, \\ &= \mathbb{P}\{M(x+h, x) = 1 \wedge N(x+h, x) = 1\}, \\ &= \lambda h p(x). \end{aligned}$$

□

Vehicles on the highway form a Poisson point process with rate λ . A point is counted if it can correctly receive a packet and forward it. Since the counting process is independent for each point, according to Proposition 1, the distribution of vehicles relaying the packet form a non-homogeneous Poisson process with rate $\lambda(x) = \lambda f(x)p(x)$.

Using the theory of non-homogeneous Poisson processes, the probability of occurrence of n events in a non-homogeneous Poisson process is given by the following theorem [26].

Theorem 2. *The probability of occurrence of n events in the interval $(a, b]$ for a non-homogeneous Poisson process with*

rate $\lambda(x)$ is given by:

$$\mathbb{P}\{N(a, b) = n\} = \frac{(\int_a^b \lambda(x) dx)^n e^{-\int_a^b \lambda(x) dx}}{n!}. \quad (2)$$

A packet is successfully forwarded if exactly one vehicle relays it. In order to avoid concurrent relaying that causes a collision, or no relaying at all, we set the number of events (*i.e.*, number of vehicles that relay the packet) in Theorem 2 to 1. We therefore obtain that

$$\mathbb{P}\{N(0, \infty) = 1\} = (\int_0^\infty \lambda(x) dx) e^{-\int_0^\infty \lambda(x) dx}. \quad (3)$$

Our first requirement is to maximize the probability of observing exactly one relay. A function of the form $\alpha e^{-\alpha}$ achieves its maximum at $\alpha = 1$, and the maximum is equal to $1/e$. Therefore, to maximize $\mathbb{P}\{N(0, \infty) = 1\}$, we have

$$\int_0^\infty \lambda(x) dx = \int_0^\infty \lambda f(x)p(x) dx = 1, \quad (4)$$

or, equivalently,

$$\int_0^\infty f(x)p(x) dx = 1/\lambda. \quad (5)$$

Our second requirement is to maximize the average hop length. Thus, the following expression, which is the average distance between two adjacent relays, should be maximized:

$$\int_0^\infty x \lambda f(x)p(x) dx. \quad (6)$$

To summarize, we want to find a $p(x)$ that maximizes expression (6) while satisfying the constraint (4). Since $0 \leq p(x) \leq 1$, $\int_0^\infty f(x)p(x) dx$ is the area under a curve that is less than or equal to $f(x)$ at each point. We have to find $p(x)$ so that the area under the curve $x f(x)p(x)$ is maximized. To this end, we show that the optimal relaying strategy $p(x)$ that satisfies these conditions is a *threshold-based* policy, and then compute the optimal threshold for UDG and LNS packet reception models.

Theorem 3. *The optimal relaying strategy $p(x)$ is of the form*

$$p(x) = \begin{cases} 0, & \text{if } x < T \\ 1, & \text{if } x \geq T \end{cases} \quad (7)$$

where T ($T > 0$) is a distance threshold.

Proof: We begin by tiling the area under the curve³ $f(x)p(x)$ into small rectangles of width Δx and height δ . Since $0 \leq p(x) \leq 1$, the optimal relaying strategy $p(x)$ essentially chooses a subset of the tiles under the curve $f(x)$ so that expression (6) is maximized subject to constraint (4).

Any rectangle with height δ under the curve $f(x)p(x)$ is mapped to a rectangle with height $x\delta$ under the curve $x f(x)p(x)$ because the two curves differ only by a factor of x . Next, we iteratively select rectangles under the curve $f(x)p(x)$ that give rectangles under the curve $x f(x)p(x)$ with the largest area. This process is continued until the total selected area

³At this stage of the proof, we do not know the exact shape of this curve as $p(x)$ is unknown.

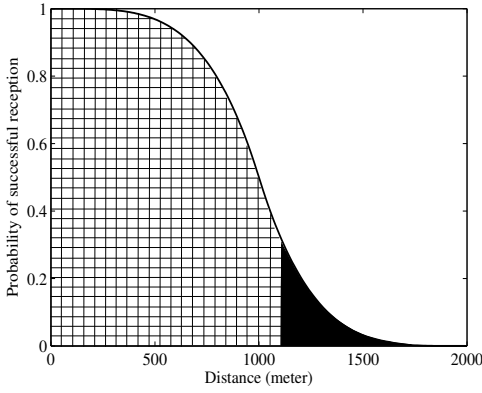


Figure 2. To maximize the average distance between relays, rightmost tiles are chosen. The total area of the chosen tiles is equal to $1/\lambda$.

under the curve $f(x)p(x)$ equals $1/\lambda$ (see Eqn. (5)). To select a rectangle, all of the rectangles below it must be already selected in order to form a continuous area under a curve. In each iteration, the algorithm selects the rightmost rectangle because it has the largest x producing the largest area for a fixed Δx . Consequently, columns are selected from right to left until the total selected area becomes equal to $1/\lambda$.

Figure 2 shows an example of how the algorithm selects the rectangles and the final result. The black region is equal to the area under the curve $f(x)p(x)$. Clearly, $p(x)$ is a unit step function shifted by a constant T as in Eqn. (7). This geometric proof holds for any packet reception function $f(x)$ that is monotonically decreasing with x (a realistic assumption). \square

B. Computing Optimal Threshold

Every vehicle should know its distance x from the current transmitter. If vehicles are equipped with GPS then x is readily available. Otherwise, x can be estimated using the received signal strength, as discussed earlier in Subsection III-F. In order to fully characterize the optimal relaying strategy $p(x)$, we need to specify T in Eqn. (7), which depends on vehicle density and packet reception model. In this subsection, we compute T for the two packet reception models described in Subsection III-D, namely UDG and LNS models. It is straightforward to extend our analysis to other packet reception models.

1) *Unit Disk Graph Model:* In UDG, $f(x) = 0$ for $x > R$ indicating that $T \leq R$. By substituting $f(x)$ and $p(x)$ in Eqn. (4), we obtain that

$$\int_0^\infty \lambda f_{\text{UDG}}(x)p(x) dx = \int_T^R \lambda dx = 1. \quad (8)$$

Consequently, it is obtained that

$$T = R - \frac{1}{\lambda}. \quad (9)$$

Obviously, as $\lambda \rightarrow \infty$ we have $T \rightarrow R$, which is the maximum possible hop length. Therefore, for UDG model, as the vehicle density on the highway increases, our relaying strategy actually achieves the largest hop length resulting in *minimum end-to-end delay*.

2) *Log-Normal Shadowing Model:* In LNS, $f(x)$ has a critical point at R . Therefore, there are two possible regions for T as follows.

(i) If $T \leq R$:

$$\begin{aligned} \int_0^\infty \lambda f_{\text{LNS}}(x)p(x) dx &= \int_T^R \lambda \left(1 - \frac{1}{2} \left(\frac{x}{R}\right)^{2\beta}\right) dx \\ &\quad + \int_R^{2R} \lambda \frac{1}{2} \left(\frac{2R-x}{R}\right)^{2\beta} dx \\ &= 1. \end{aligned} \quad (10)$$

After some algebraic manipulation, we obtain that

$$(R - T) - \frac{T^{2\beta+1}}{2R^{2\beta}(2\beta + 1)} = \frac{1}{\lambda}. \quad (11)$$

(ii) If $T > R$:

$$\int_T^{2R} \frac{1}{2} \left(\frac{2R-x}{R}\right)^{2\beta} dx = \frac{1}{\lambda}, \quad (12)$$

which, results in the following expression

$$T = 2R - \left(\frac{2R^{2\beta}(2\beta + 1)}{\lambda}\right)^{\frac{1}{2\beta+1}}. \quad (13)$$

To find T , Eqns. (11) and (13) are solved numerically. However, only one of these equations results in a valid choice for T .

C. Recovery From Unsuccessful Relaying

Although our relaying strategy is optimized to achieve the highest successful relaying probability, there is still a non-zero chance (equal to $1 - \frac{1}{e}$) that there would be no relay or more than one relay after the threshold T . In the case of no relaying or collision, we need a recovery mechanism to retransmit the packet.

Different mechanisms can be used to handle the recovery process. One mechanism is to dynamically adjust the threshold based on the reason for forwarding failure. If the failure was due to the collision of concurrent transmissions, the threshold will increase to narrow down the relaying region. On the other hand, if the channel was idle during the last time slot, then the threshold will decrease to expand the relaying region and accommodate more vehicles. This mechanism resembles the dynamic adjustment of *contention window* in IEEE 802.11 standard.

Another alternative, which we adopt in this work, is to introduce back-off regions to relay the packet when the first region fails. These adjacent regions are separated by the thresholds that specify the length of each region as illustrated in Figure 3. Regions are ordered according to their thresholds, a region with a larger threshold has higher priority for relaying. For a region i , the optimal threshold T_i is computed so that the probability of successful relaying within the region is maximized. That is, the width of the region i , *i.e.*, $T_i - T_{i+1}$, is chosen so that the probability of having only one relay in the region is maximum. Upon receiving a packet, region R_1 (*i.e.*, the furthest region) tries to relay the packet. Vehicles in

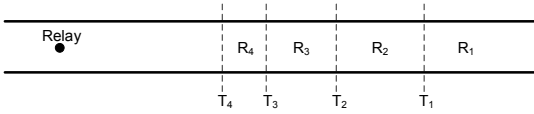


Figure 3. The priority of regions for relaying a packet decreases from right to left.

other regions listen to the channel to determine if R_1 could successfully relay the packet. If region i fails to relay the packet successfully, region $i+1$ will take its turn. This process continues until the packet is relayed successfully.

The same technique we used to compute the first threshold in the previous subsection can be used to compute other thresholds. For instance, to compute the second threshold (T_2), we have

$$\int_{T_2}^{T_1} f(x) dx = \frac{1}{\lambda}, \quad (14)$$

where $f(x)$ is the packet reception function. This equation can be solved for T_2 . Once T_2 is computed, other thresholds can be computed iteratively.

V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of OPTIMIZED through simulation, and compare it against RANDOM proposed in [9]. The reviewed rebroadcasting schemes such as weighted p -persistent [14] do not consider vehicle density in their rebroadcast probability function. Since vehicle density changes during the day, we consider RANDOM for comparison. Although simple, RANDOM can adapt rebroadcasting probability to vehicle density. We also investigate the impact of the Poisson assumption for vehicle distribution, and distance estimation inaccuracies on the performance of OPTIMIZED. The performance measures of interest in this section are the end-to-end delay and throughput.

A. Simulator Overview

The simulator was built using the OMNeT++ discrete-event simulator [28]. A one-lane unidirectional highway was simulated using the freeway mobility model described in Section III-E. The highway was 10 Km long, with a single roadside AP at the beginning of the segment. The Log-Normal Shadowing (LNS) packet reception model was used to obtain the packet reception probability at a given location.

During the initialization phase of the simulation, the simulation parameters are set to their initial values, as shown in Table I. Each time slot is 0.1 second. The simulation has a 1000-second warmup period, during which vehicles enter the highway and move according to the freeway mobility model. Following the warmup period, performance data for the relaying strategies are recorded for an additional 500 seconds of simulation run. During this phase, new vehicles continue to arrive to the highway according to a Poisson process. The inter-arrival times are exponentially distributed with mean μ . The exponential distribution has been found to be a good model for vehicles entering a highway [29]. Our simulator, however, is able to use any general distribution for vehicle inter-arrival times.

Table I
SIMULATION PARAMETER SETTINGS.

Simulation Settings		
Section	Parameter	Value
General	Vehicle arrival rate (μ)	0.2 – 1.0 vehicle/sec
	Packet generation rate	0.02 packet/time-slot
	Simulation Length	1500 sec
Freeway mobility model	(V_{min}, V_{max})	(30, 60), (80, 100) Km/hr
	Max acceleration	1.0 Km/sec
	Safety distance	$2V_{max}$
Packet reception model	R	1000 m
	β	2

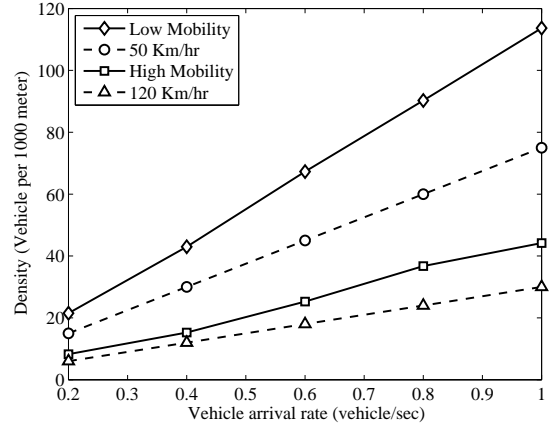


Figure 4. Measured vehicle density for different vehicle arrival rates.

We study one-hop and multi-hop performance of OPTIMIZED and RANDOM. Simulation and analytical results are presented for different vehicle densities, which are controlled by changing the rate at which vehicles enter the highway. Figure 4 shows the relationship between vehicle density on the highway and vehicle arrival rate. In the figure, the *Low Mobility* and *High Mobility* refer to (V_{min}, V_{max}) of (30, 60) and (80, 100) Km/hr, respectively.

Next, we present simulation results for the single-hop and multi-hop network settings. Each simulation is repeated 5 times and the average is presented with 95% confidence intervals.

B. Single-Hop Performance

We consider single-hop throughput to study the performance of the relaying strategies, and validate the analytical results presented in Section IV. More precisely, we want to see if the calculated optimal threshold achieves the maximum throughput when a realistic mobility model is used. For the ease of exposition, we ignore recovery/retransmissions and focus on the throughput achieved with a single threshold. We define the single-hop *throughput* as the *fraction of packets that can be relayed successfully*.

1) *Effect of Relaying Threshold on Throughput*: Figure 5 shows the effect of the threshold T on the throughput. The throughput observed is highly dependent on the distance

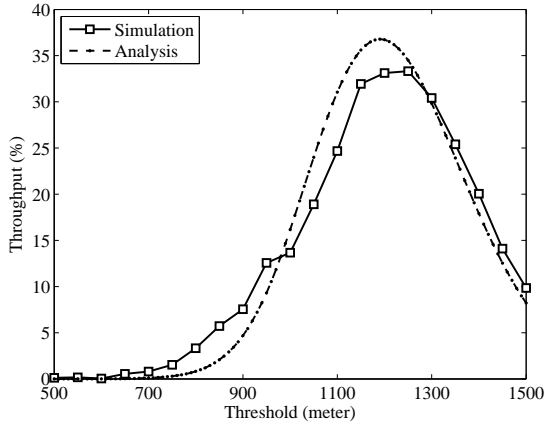


Figure 5. Effect of threshold T on the single-hop throughput ($\mu = 0.8$).

threshold T . When the threshold value is small, the relaying region is large, and many nodes attempt to relay packets simultaneously. The ensuing collisions dramatically reduce the throughput. At some larger threshold value, the throughput reaches its maximum value, and then falls sharply after this point because the probability of having at least one relaying node in this narrow region is very low.

In Section IV-A, we computed the probability of successfully relaying a packet in OPTIMIZED. Figure 5 shows strong agreement between the analytical results (dashed line) obtained in Subsection IV-B and the simulation results (solid line) for the achieved throughput. In particular, the shape of the distribution is similar, and the optimal threshold values obtained from both approaches are very close to each other. We have validated our analytical results against simulations in additional scenarios, which show similar behavior to Figure 5, and hence the results are not presented here.

2) Effect of Vehicle Arrival Rate on Relaying Threshold:

Figure 6 shows the effect of vehicle arrival rate on the choice of thresholds. In particular, we focus on thresholds T_1 and T_2 corresponding to Region 1 and Region 2, respectively. As the vehicle arrival rate increases, the threshold values increase slightly, making the regions smaller. This result makes sense intuitively. Since we want to have exactly one relay in each region on average, the length of all regions shrink if vehicle density increases (due to increased vehicle arrival rate). More precisely, this agrees with the analytical model since the optimal threshold for the first region comes from the equation $\int_T^\infty \lambda f(x) dx = 1$, where $f(x)$ gives the probability of receiving a packet at distance x from the transmitter and λ is the rate of the Poisson process that determines the distribution of vehicles on the highway. Therefore, if vehicle density (λ) increases, then T increases to satisfy the equation. A similar justification applies for other regions.

3) *Effect of Vehicle Arrival Rate on Throughput:* The simulation results for the throughput are presented in Figure 7. We observe that as the vehicle density increases (*i.e.*, vehicle arrival rate increases) the achieved throughput slowly approaches its theoretical limit of $1/e$. Moreover, in low densities, the throughput of OPTIMIZED is within 8% of that for RANDOM.

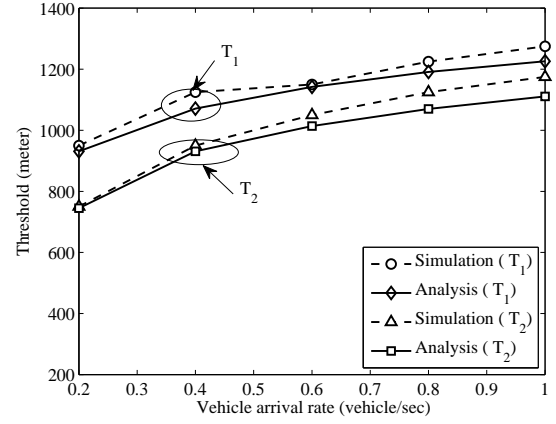


Figure 6. Effect of vehicle arrival rate on relaying thresholds.

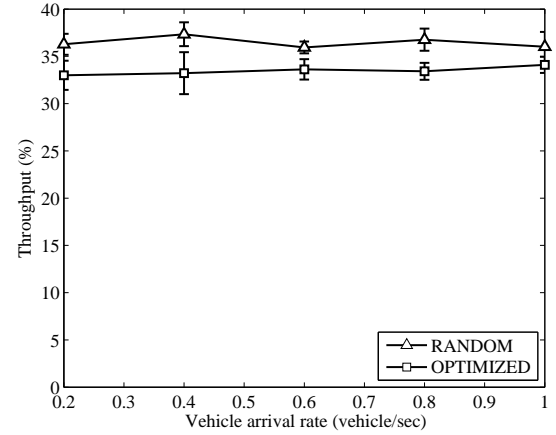


Figure 7. Effect of vehicle arrival rate on single-hop throughput.

C. Multi-Hop Performance

In the multi-hop relaying scenario, a new packet is generated every 50 time slots (each time slot is 0.1 second.) and injected into the network. The roadside AP broadcasts the packet and vehicles try to forward it based on the relaying strategy in use. This procedure continues until the destination vehicle (7 Km away from the AP) is reached.

Figures 8 and 9 present the simulation results for low and high mobility scenarios, respectively (speed range for low and high mobility are (30, 60) and (80, 100) Km/hr, respectively). We observe that OPTIMIZED needs about half as many hops to reach the destination, and its end-to-end delay is about 50% less than that for RANDOM. On average, the next hop in RANDOM is $\frac{R}{2}$ away, while the usage of thresholds in our scheme gives preference to those partners far from the previous hop, resulting in half as many hops as for RANDOM.

A slight decline in the number of hops is observed when the vehicle density increases. Recalling Figure 7, an increase in vehicle density results in a slight increase of all thresholds. This fact increases the average distance between two adjacent relays in the multi-hop path, and as a result, the number of hops required to reach the destination decreases slightly when the vehicle density increases. Moreover, the throughput of single-hop relaying approaches the $1/e$ limit when density

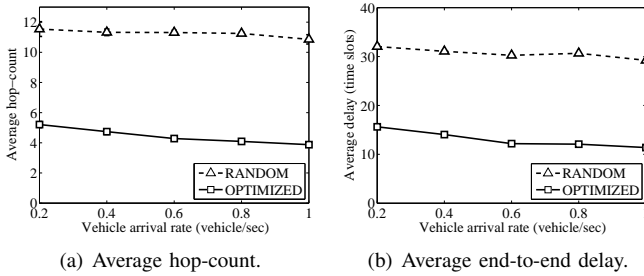


Figure 8. Multi-hop performance in low mobility scenario.

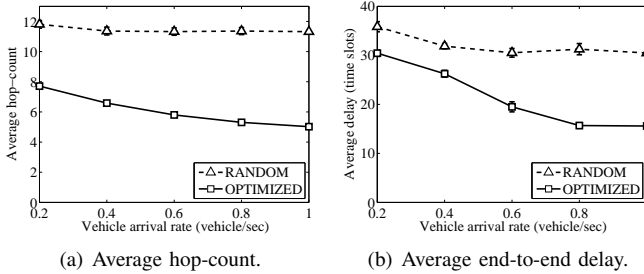


Figure 9. Multi-hop performance in high mobility scenario.

increases (see Figure 7). The two aforementioned observations are the key features characterizing the end-to-end delay behavior of the relaying strategies. A similar behavior was observed for the high mobility case. The end-to-end delay is similar in both cases, especially when the vehicle density is high.

D. Sensitivity to Vehicle Distribution

The simulation results show that the performance is slightly worse in the low density scenarios. That is, the end-to-end delay improves as the vehicle density increases. Looking at a simulation snapshot of the highway and the positions of vehicles (see Figure 10) shows that the statistical distribution of vehicles varies with the arrival rate and the speed parameters of the freeway mobility model. When vehicle density is low, the freeway mobility model disturbs the distribution of nodes because the low speed nodes hinder approaching nodes and very dense regions are created while other regions are empty. Therefore, we suspect that the hypothesis of Poisson distribution of nodes on the highway is violated in low density scenarios.

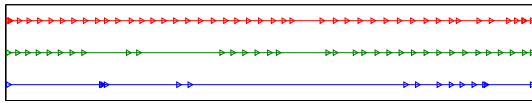


Figure 10. Snapshot of 1500 meter highway for 3 vehicle arrival rates: 0.6 (top), 0.4 (middle), and 0.2 (bottom).

To further investigate this behavior, a chi-square test was performed on a sample of vehicle positions at an arbitrary time on the highway. The chi-square test (or χ^2 test) is widely used for goodness-of-fit test in order to determine whether a sample of observations follows a certain distribution, such as Poisson distribution. Using Matlab, we calculated the p -value

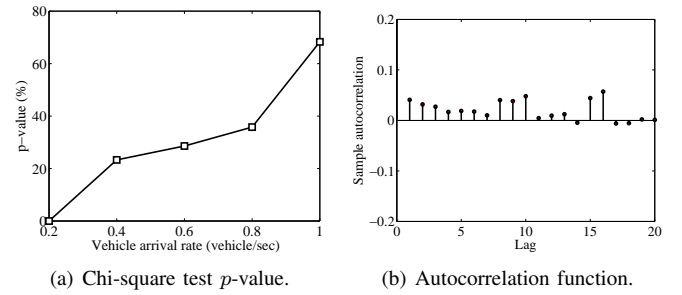


Figure 11. (a) Goodness-of-fit, and (b) independence tests for vehicle distribution on the highway.

of the chi-square goodness-of-fit, which is the probability of observing a Poisson distribution.

Figure 11(a) shows the calculated p -value for different vehicle arrival rates. It is observed that the Poisson distribution hypothesis is rejected due to the very low value of p when the arrival rate is 0.2 (*i.e.*, 13 vehicles per 1000 meters on average). For other arrival rates, the p -value is high enough to accept the hypothesis. However, a significant increase in p -value is observed when density increases, implying that the distribution of vehicles resembles the Poisson distribution at higher arrival rates.

We have also plotted the autocorrelation function for vehicle positions in Figure 11(b) for the high mobility scenario. The autocorrelation is used to test for the independence of observed samples. The figure shows that the autocorrelation function at different lags is extremely low, indicating independent vehicle locations on the highway.

We found that the freeway mobility model disturbs the initial distribution of vehicles completely when density is very low. However, as shown in Figures 8 and 9, OPTIMIZED is relatively insensitive to the arrival rate and the mobility model, and shows an acceptable performance even in the low density scenario.

E. Sensitivity to Distance Estimation Error

In this subsection, we study the effect of distance estimation error on the performance of OPTIMIZED. Recall that OPTIMIZED requires knowledge of the distance between vehicles and the current transmitter in order to choose a relay. In practice, particularly in the absence of onboard GPS, distance estimation may not be accurate. For example, RSSI-based distance estimation techniques are in general inaccurate. To study the sensitivity of OPTIMIZED to distance estimation errors, an *artificial error* is added to vehicle distances that are obtained from the simulator. After this process, each vehicle knows its distance with some error and not accurately. The distance error for each vehicle is randomly picked from a normal distribution with 0 mean and variance σ^2 .

Figure 12 illustrates the single-hop throughput and multi-hop delay for different values of σ . Note that $R = 1000$ m is the transmission range. It is observed that the single-hop throughput is affected only when σ exceeds $R/5$, *i.e.*, 200 m, which represents a significantly high estimation error. However, for a reasonable estimation error, both the single-

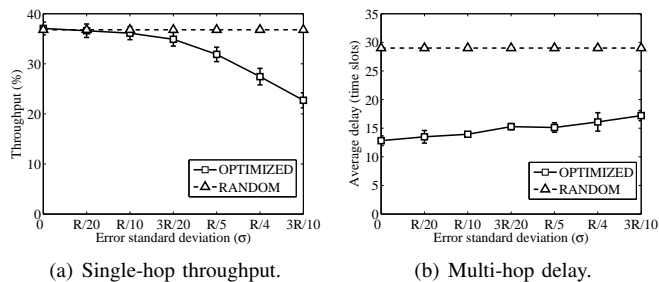


Figure 12. Performance in the presence of distance estimation errors.

hop throughput and end-to-end delay are relatively unchanged. Interestingly, we observe that when the estimation error is extremely high the end-to-end delay approaches that of RANDOM.

VI. CONCLUSION

This paper presented a low-latency distributed routing algorithm called OPTIMIZED for vehicular ad hoc networks. Due to the highly dynamic nature of vehicular ad hoc networks, routing algorithms that require global network state information or centralized coordination are not practical, making OPTIMIZED a desirable candidate for these networks. While OPTIMIZED does not require local or global coordination among vehicles, it achieves throughput and delay comparable to fully coordinated routing algorithms. In particular, through analysis and simulations, we showed that its throughput is only a factor of $1/e$ from the throughput of a scheme that centrally coordinates vehicle transmissions, essentially achieving the same throughput as the popular ALOHA-based RANDOM [9]. However, its end-to-end delay is about *half* of that for RANDOM. We further conjecture that OPTIMIZED achieves an average path length that is close to that of GEOGRAPHIC, which locally coordinates vehicle transmissions for delay minimization, while we know its throughput is superior to that of GEOGRAPHIC.

Our next step would be to extend OPTIMIZED to multicast scenarios, enabling low-latency multicast in vehicular networks. Also, we would like to design a data dissemination middleware based on OPTIMIZED for vehicular networks. This involves designing the appropriate protocols for naming and addressing as well as location and mobility management.

REFERENCES

- [1] "OnStar." [Online]. Available: <http://www.onstar.com>
- [2] "Standard specification for telecommunications and information exchange between roadside and vehicle systems - 5GHz band dedicated short range communications (DSRC) medium access control (MAC) and physical layer (PHY) specifications," ASTM E2213-03, Sep. 2003.
- [3] D. Jiang and L. Delgrossi, "IEEE 802.11p: Towards an international standard for wireless access in vehicular environments," in *Proc. IEEE VTC*, Marina Bay, Singapore, May 2008.
- [4] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc on-demand distance vector (AODV) routing," RFC 3561, IETF, Jul. 2003.
- [5] D. Johnson, Y. Hu, and D. Maltz, "The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4," RFC 4728, IETF, Feb. 2007.
- [6] S. Biswas and R. Morris, "ExOR: Opportunistic multihop routing for wireless networks," in *Proc. ACM SIGCOMM*, Philadelphia, USA, Aug. 2005.

- [7] B. Karp and H. T. Kung, "Greedy perimeter stateless routing for wireless networks," in *Proc. ACM Mobicom*, Boston, USA, Aug. 2000.
- [8] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks," in *Proc. Workshop on Discrete Algorithms and Methods for Mobile Computing and Communication*, Seattle, USA, Aug. 1999.
- [9] M. Johnson, L. De Nardis, and K. Ramchandran, "Collaborative content distribution for vehicular ad hoc networks," in *Allerton Conf. Communication, Control, and Computing*, Urbana Champaign, USA, Sep. 2006.
- [10] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in *the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, New Orleans, LA, 1999, pp. 90–100.
- [11] L. Xiong, L. Libman, and G. Mao, "On cooperative communication in ad-hoc networks: The case for uncoordinated location-aware retransmission strategies," in *IEEE Conference on Local Computer Networks (LCN)*, Zurich, Switzerland, October 2009.
- [12] R. K. Ganti and M. Haenggi, "Analysis of uncoordinated opportunistic two-hop wireless ad hoc systems," *CoRR*, vol. abs/0901.2160, 2009.
- [13] S. Biswas and R. Morris, "Exor: opportunistic multi-hop routing for wireless networks," *SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 4, pp. 133–144, 2005.
- [14] N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks," *Wireless Communications, IEEE*, vol. 14, no. 6, pp. 84–94, 2007.
- [15] R. Chen, W.-L. Jin, and A. Regan, "Broadcasting safety information in vehicular networks: issues and approaches," *Netw. Mag. of Global Internetwkg.*, vol. 24, no. 1, pp. 20–25, 2010.
- [16] *Broadcasting in VANET*, 2007.
- [17] L. Briesemeister and G. Hommel, "Role-based multicast in highly mobile but sparsely connected ad hoc networks," in *MobiHoc '00: Proceedings of the 1st ACM international symposium on Mobile ad hoc networking & computing*. IEEE Press, 2000, pp. 45–50.
- [18] G. Korkmaz, E. Ekici, F. Özgüner, and U. Özgüner, "Urban multi-hop broadcast protocol for inter-vehicular communication systems," in *VANET '04: Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*. ACM, 2004, pp. 76–85.
- [19] G. Korkmaz, E. Ekici, and F. Özgüner, "An efficient fully ad-hoc multi-hop broadcast protocol for inter-vehicular communication systems," in *IEEE ICC '06*. ACM, 2006, pp. 423–428.
- [20] F. Bai, N. Sadagopan, and A. Helmy, "The IMPORTANT framework for analyzing the impact of mobility on performance of routing protocols for adhoc networks," *Ad Hoc Networks Journal*, vol. 1, no. 4, pp. 383–403, Aug. 2003.
- [21] Y. Toor, P. Muhlethaler, A. Laouiti, and A. de La Fortelle, "Vehicle ad hoc networks: applications and related technical issues," *IEEE Communications Surveys and Tutorials*, vol. 10, no. 3, pp. 74–88, 2008.
- [22] A. Senart, M. Bouroche, V. Cahill, and S. Weber, "Vehicular networks and applications," in *Middleware for Network Eccentric and Mobile Applications*. Springer, 2009, pp. 369–381.
- [23] J. Kuruvila, A. Nayak, and I. Stojmenovic, "Hop count optimal position based packet routing algorithms for ad hoc wireless networks with a realistic physical layer," in *Proc. IEEE MASS*, Fort Lauderdale, USA, Oct. 2004.
- [24] R. Parker and S. Valaee, "Vehicular node localization using received-signal-strength indicator," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3371–3380, Nov. 2007.
- [25] D. Hadaller, S. Keshav, T. Brecht, and S. Agarwal, "Vehicular opportunistic communication under the microscope," in *Proc. ACM MobiSys*, San Juan, Puerto Rico, Jun. 2007.
- [26] J. F. C. Kingman, *Poisson processes*. Oxford University Press, 1993.
- [27] S. M. Ross, *Simulation: Statistical Modeling and Decision Science Series*, 4th ed. Academic Press, 2006.
- [28] "OMNeT++." [Online]. Available: <http://www.omnetpp.org/>
- [29] A. Rnyi, "On two mathematical models of the traffic on a divided highway," *Journal of Applied Probability*, vol. 1, no. 2, pp. 311–320, 1964.