# Intersection-Based Geographical Routing Protocol for VANETs: A Proposal and Analysis

Hanan Saleet, *Member, IEEE*, Rami Langar, *Member, IEEE*, Kshirasagar Naik, *Senior Member, IEEE*, Raouf Boutaba, *Senior Member, IEEE*, Amiya Nayak, *Senior Member, IEEE*, and Nishith Goel

***Abstract*—This paper presents a class of routing protocols for vehicular ad hoc networks (VANETs) called the Intersection-based Geographical Routing Protocol (IGRP), which outperforms existing routing schemes in city environments. IGRP is based on an effective selection of road intersections through which a packet must pass to reach the gateway to the Internet. The selection is made in a way that guarantees, with high probability, network connectivity among the road intersections while satisfying quality-of-service (QoS) constraints on tolerable delay, bandwidth usage, and error rate. Geographical forwarding is used to transfer packets between any two intersections on the path, reducing the path's sensitivity to individual node movements. To achieve this, we mathematically formulate the QoS routing problem as a constrained optimization problem. Specifically, analytical expressions for the connectivity probability, end-to-end delay, hop count, and bit error rate (BER) of a route in a two-way road scenario are derived. Then, we propose a genetic algorithm to solve the optimization problem. Numerical and simulation results show that the proposed approach gives optimal or near-optimal solutions and significantly improves VANET performance when compared with several prominent routing protocols, such as greedy perimeter stateless routing (GPSR), greedy perimeter coordinator routing (GPCR), and optimized link-state routing (OLSR).**

***Index Terms*—Message routing, performance analysis, quality of service (QoS), vehicular ad hoc networks (VANETs).**

Manuscript received September 30, 2010; revised April 6, 2011, June 5, 2011, and August 17, 2011; accepted September 26, 2011. Date of publication October 25, 2011; date of current version December 9, 2011. This work was supported in part by the Natural Science and Engineering Council of Canada under its discovery program and in part by the World Class University program through the Korea Science and Engineering Foundation funded by the Ministry of Education, Science, and Technology under Project R31-2008-000-10100-0. The review of this paper was coordinated by Prof. A. Jamalipour.

H. Saleet is with the Department of Mechanical and Industrial Engineering, Applied Science University, Amman 11931, Jordan (e-mail: saleet21@yahoo.com).

R. Langar is with the Computer Science Laboratory of Paris 6, University of Pierre and Marie Curie (UPMC), UPMC Sorbonne Universites, 75005 Paris, France (e-mail: rami.langar@lip6.fr).

K. Naik is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: knaik@swen.uwaterloo.ca).

R. Boutaba is with the David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada, and also with the Division of IT Convergence Engineering, Pohang University of Science and Technology, Gyungbuk 790-784, Korea (e-mail: rboutaba@cs.uwaterloo.ca).

A. Nayak is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: anayak@site.uottawa.ca).

N. Goel is with Cistel Technology Inc., Ottawa, ON K2E 7V7, Canada (e-mail: ngoel@cistel.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2011.2173510

## I. Introduction

MUCH existing research considers vehicular ad hoc networks (VANETs) as a vehicle-to-vehicle or a vehicle-to-road-side-unit network architecture that can be easily deployed without relying on expensive network infrastructure. Nevertheless, enabling communication between vehicles and preexisting fixed infrastructure such as gateways to the Internet opens up a plethora of interesting applications to both drivers and passengers. The promising applications and the cost effectiveness of VANETs constitute major motivations behind increasing interest in such networks [1]–[3]. The success of VANETs revolves around a number of key elements such as message routing between the mobile nodes (MNs) and the gateway to the Internet. Without an effective routing strategy, the success of VANETs will continue to be limited.

We classify VANET-based applications into two categories: 1) those that are sensitive to delay, e.g., downloading a multimedia application from the closest Internet gateway, connecting to a virtual personal network (VPN) for video or voice conferencing, and video streaming; and 2) those that are delay tolerant, e.g., sending simple text messages or sending an advertisement. In this paper, we focus on message routing in both classes of applications. The main concern is whether the performance of VANET routing protocols can satisfy the delay requirements of such applications.

Analysis of traditional routing protocols for mobile ad hoc networks (MANETs) demonstrated that their performance is poor in VANETs [4], [5]. The main problem with these topology-based routing protocols (e.g., optimized link-state routing (OLSR) [6], dynamic source routing [7], and ad-hoc on demand distance vector routing (AODV) [8]) in VANET environments is their route instability. Indeed, the traditional node-centric view of the routes (i.e., an established route is a fixed succession of nodes between the source and destination) leads to frequent broken routes in the presence of VANETs' high mobility. Consequently, many packets are dropped, and the overhead due to route repairs or failure notifications significantly increases, leading to low delivery ratios and high transmission delays [9].

An alternative approach is offered by geographical routing protocols, such as distance routing effect algorithm for mobility (DREAM) location service (DLS) [10], greedy perimeter stateless routing (GPSR) [11], and greedy perimeter coordinator routing (GPCR) [12], which decouples forwarding from the nodes' identity. They do not establish routes but use the position of the destination and the position of the neighboring nodes to forward data. Despite better path stability, geographical forwarding does not perform well in a city environment either [4], [13]. Its problem is that, many times, it cannot find a next hop (i.e., a node closer to the destination than the current node). The recovery strategies proposed in the literature are often based on planar graph traversals, which were shown not to be as effective in VANETs due to radio obstacles and high node mobility [4].

A number of road-based routing protocols [4], [5], [13] have been designed to address this issue. However, they fail to factor in vehicular traffic flow by using the shortest road path between source and destination [14]. It is possible indeed that the road segments on the shortest path are empty.

To overcome these limitations, we propose in this paper an Intersection-based Geographical Routing Protocol (IGRP) consisting of successions of road intersections that have, with high probability, network connectivity among them. Geographical forwarding is still used to transfer packets between any two intersections within the path, reducing the path's sensitivity to individual node movements. The selection of the road intersections is made in a way that maximizes the connectivity probability of the selected path while satisfying quality-of-service (QoS) constraints on the tolerable delay within the network, bandwidth usage, and error rate.

To achieve this, we mathematically formulate the QoS routing problem as a constrained optimization problem. Specifically, analytical expressions of connectivity probability, tolerable end-to-end delay, hop count, and bit error rate (BER) for a two-way road scenario are derived. Then, we propose a genetic algorithm (GA) to solve our NP-complete optimization problem. Numerical and simulation results show that the proposed protocol achieves an optimal or a near-optimal solution, particularly in sparse networks. Therefore, it stands out as a promising candidate compared to the well-known protocols: GPSR [11], GPCR [12], and OLSR [6].

The remainder of this paper is organized as follows. Section II presents an overview of the related works, followed by a description of our proposed IGRP in Section III. In Section IV, we present the analytical framework used to evaluate the QoS routing problem. In Section V, we formulate the QoS routing problem as an optimization problem and present a GA to solve it. Numerical and simulation results are presented in Section VI. Finally Section VII contains our concluding remarks.

## II. Related Work

As we previously mentioned, message routing protocols are classified into two categories, i.e., topology and position based [29]–[37]. In topology-based protocols, it is assumed that each node has information about the entire network topology before the node begins forwarding messages. In position-based routing protocols, messages are routed based on knowledge of the geographical location of the source, intermediate nodes, and final destination. One advantage of geographical routing protocols is that they can find a suboptimal route from source to destination without the use of routing tables; therefore, there is no need to flood the network and store routing information at each node. This section reviews a number of the prominent existing routing protocols and discusses the drawbacks that make these protocols unsuitable for VANETs.

### *A. OLSR*

OLSR [6] is considered as a topology-based routing protocol. Nodes using OLSR periodically broadcast their routing table to the rest of the nodes in the network, which incurs a large communication overhead. OLSR limits the number of nodes that forward the control messages using multipoint relays. It uses two primary control messages: 1) topology control messages and 2) HELLO messages. Topology control messages are forwarded across the network. HELLO messages are sent to each one-hop neighbor. If a node does not receive HELLO messages from one neighbor during a certain time period, then the link is considered down. The source using this link to forward messages is not aware that the route is broken until that intermediate node broadcasts its next topology control message. In VANETs, the movement of nodes may cause the network topology to frequently change, which causes deterioration in network performance as it introduces congestion in the communication channel. These limitations of the topology-based protocols make them unsuitable for VANETs.

### *B. GPSR*

GPSR [11] assumes that each node in the network has a local table in which all neighboring nodes are listed by name and position. The entry of the local table is soft stated and updated after the related timer expires, where beacons broadcast information of the new neighbor(s). GPSR also assumes that each source node knows the location of the destination with the aid of a location service. GPSR has two working modes: 1) a greedy forwarding mode and 2) a perimeter mode.

Greedy forwarding is the default mode, where the packet is forwarded to the node that is geographically closer to the destination. Greedy forwarding works well if there are no holes, meaning voids, in the network. Voids may be caused by physical obstacles, such as mountains or large buildings. If there is a void between the forwarding node and destination node, then the greedy forwarding may get deadlocked at the perimeter of the void. Thus, the forwarding node may not find a neighbor that is geographically closer to the destination than itself.

In such a scenario, the forwarding node switches to perimeter mode where it chooses the neighbor as the next forwarder based on the right-hand rule. As soon as that neighbor finds a node that is closer to the destination than itself, it returns to greedy forwarding mode. However, if such neighbor is not available, then the packet continues in perimeter mode, moving along the perimeter of the voids.

Because GPSR lacks information about the network topology, it can potentially go through loops. This occurs in the case of perimeter routing when the protocol routes the message in the wrong direction, resulting in performance degradation [28].

### *C. GPCR*

GPCR [12] assigns the routing decision to the nodes located at the street intersections, and at the same time, it uses the greedy forwarding strategy to route the message between the street intersections. Like GPSR, GPCR does not make use of road maps for routing the messages, which may result in loops and introduce many hops in the route. In addition, GPCR does not take into consideration the quality of the routes nor does it have a method to select the best path.

### D. MURU

The MUltihop Routing protocol for Urban VANETs (MURU) [36] assumes that each node has a static street map and that there is a location service that gives the source node information about the location of destinations. To find a route, therefore, the source node calculates the shortest path to the destination based on a static street map and the location of both the source and the destination. MURU provides routes that minimize the hop count. At the same time, it proposes the "expected disconnection degree (EDD)" to estimate the quality of the routes. The EDD of a given route represents the probability that this route will fail during a given time period. MURU uses the EDD to construct an optimal path based on predicted speed, location, and road geometry. Each node broadcasts route request packets, which are routed on paths that are constrained by node movement trajectory. However, since MURU uses the local information available to the forwarding node, it is susceptible to local optimum [32], which would significantly decrease the scalability of the routing protocol.

### E. Delay-Bounded Routing in VANETs

In [37], a carry-and-forward algorithm to enable the vehicles to deliver messages during a limited time period, which is specified by the VANET's application, is proposed. It is assumed that each vehicle has access to a digital map that is preloaded with historical statistical data about the traffic on the roads. This traffic information is utilized to form the routes. One drawback of this scheme is that it assumes that each node can update statistical data about traffic conditions once it comes into contact with an access point. However, given the fact that the access points cannot be densely distributed in the network, they may not be found at all times. In addition, the traffic pattern changes throughout the day, resulting in frequent obsolete information that leads to incorrect routing decisions.

Several other routing protocols for VANETs have been proposed. However, many of them do not consider the characteristics of VANETs, such as the vehicles' movement on the roads where they face radio obstacles. In addition, they do not consider the staleness of information about the network, which causes the selected routes to be unstable. To overcome these limitations, we propose IGRP, which solves the QoS routing problem in VANETs. As opposed to existing approaches, the constructed routes are not based on the MNs. Instead, IGRP chooses the routes based on fixed points, which are the road intersections (i.e., junctions). This increases the stability of the constructed routes. Specifically, IGRP chooses the path that maximizes connectivity probability while satisfying the QoS constraints regarding hop count, BER, and end-to-end delay. Between any two intersections on the selected path, geographical forwarding is used to transfer packets, thus reducing the path's sensitivity to individual node movements. To do so, IGRP makes use of a central control unit, which is the gateway. This latter node has indeed detailed information about the MNs in its vicinity using a location-aware service and uses a GA to choose the optimal routes. Note that our proposed GA converges to the optimal or the near-optimal solution after a few iterations, as will be shown in Section VI.

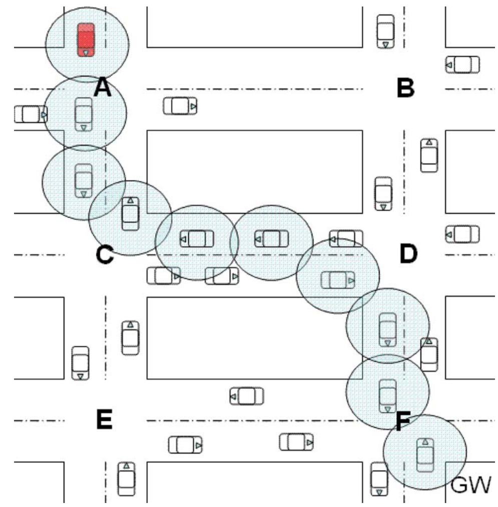



Fig. 1. Message routing in VANETs using IGRP.

## III. INTERSECTION-BASED GEOGRAPHICAL ROUTING PROTOCOL

In this section, we introduce our proposed IGRP. First, we present the system model used to build our framework. Then, we present the functionality of IGRP.

### A. System Model

We envision a VANET environment that consists of roads with intersections, which is a typical scenario in urban areas. We assume location-aware vehicles that obtain their geographical position from a global positioning system (GPS) receiver or other location service such as in [15]. Vehicles also have access to a digital map of the area using an onboard navigation system to determine the position of its neighboring road intersections. Such kind of digital map has already been commercialized. The latest one is developed by MapMechanics [16], which includes road speed data and an indication of the relative density of vehicles on each road. Yahoo is also working on integrating traffic statistics in its new product called SmartView [17], where real traffic reports for major U.S. cities are available.

The street map is abstracted as a graph $G(V, E)$. For any two intersections $A$ and $B$, $(A, B) \in G$ if and only if there is a road segment connecting $A$ and $B$ and vehicles can travel on that segment.

In the urban scenario we are considering, the network consists of MNs (vehicles) and stationary Internet gateways that do not provide full city coverage (see Fig. 1). When a message is generated at an MN, depending on its location, it may need to be relayed multiple times through several vehicles before reaching the closest gateway.

### B. Functionality of IGRP

Recent studies in multihop routing in VANETs [3]–[5], [36] have shown that, with the GPS and digital map, geographic routing, in which data packets are forwarded from the source to the destination with the aid of the nodes' location information, has high end-to-end packet delivery ratio, low end-to-end delay, and low control overhead. All these protocols assume that an efficient location management service is available to provide the source node with the destination's location. Hence, a good location management scheme in VANETs is important to support geographic routing and other location-based applications, as adopted in IGRP.

Specifically, in IGRP, a source node needs to know the route that it should use to forward data packets to the Internet gateway. This information is provided by the Internet gateway, which has an up-to-date view of the local network topology. Indeed, this gateway acts as a location server where it is responsible for saving current location information about all vehicles in its vicinity. This can be addressed using our previously proposed location service management protocol called Region-based Location-Service-Management Protocol (RLSMP) [15]. Specifically, each vehicle reports its location information to the gateway each time it moves one transmission range farther from its previous location. This information contains the node ID, transmission range $T_r$, $X$ and $Y$ coordinates of the node location, time of the last update, and the velocity and direction of the node's movement. Based on these location information, the Internet gateway constructs a set of routes between itself and the MNs. Nevertheless, one should note that, if these routes consist of intermediate MNs, these routes cannot be considered to be stable due to intermediate nodes' mobility. To increase their stability, IGRP builds routes based on intermediate and adjacent road intersections toward the gateway. These routes, which are called backbone routes, are represented as sequences of intersections, as shown in Fig. 1. This figure shows, for example, three feasible backbone routes: A-B-D-F, A-C-D-F, or A-C-E-F.

Based on the constructed backbone routes, the Internet gateway will select the path that has, with high probability, the most "connected" road segments. A connected road segment is a segment between two adjacent intersections with enough vehicular traffic to ensure network connectivity. The selected path will be then sent to the source node and will be stored in the data packet headers to allow the intermediate nodes to geographically forward packets between intersections. Indeed, the forwarding process can be described as follows: When the MN moves along the same junction, it chooses the next hop based on the geographical forwarding algorithm, where the next intermediate MN is chosen to be the node closest to the intersection that terminates the backbone link. When the MN is approaching an intersection, it selects a node closest to the next intersection (i.e., next hop in the backbone route) using geographical routing. Note that the "next intersection" is known by the intermediate MN since this information is stored in the received data packet header, as mentioned earlier. Note also that, in our approach, the gateway selects the most connected backbone path, and hence, the probability of finding an intermediate MN toward the "next intersection" is high.

TABLE I
$T_r$ SETTINGS

| $\gamma$ | $0-0.004$ | 0.004-0.006 | $0.006-0.008$ | $\geq 0.008$ |
|---|---|---|---|---|
| $T_r(m)$ | 750 | 550 | 350 | 250 |

It is worth noting that the path selection process is achieved while ensuring the QoS requirements of the VANET application, mainly tolerable delay, bandwidth usage (represented by the hop count), and BER constraints. In view of this, to meet the end-to-end delay requirements, the selected backbone routes should have high connectivity probability. One way to increase the connectivity probability in low-density roads is to increase the transmission range $T_r$ of the MNs. On the other hand, when road density increases, $T_r$ should be reduced to avoid high interference and then reduce the error rate without deteriorating the network connectivity. Hence, in IGRP, the gateway will decide on the transmission range that the source node (i.e., the vehicle requesting the optimal backbone path) should use to achieve high route connectivity and, at the same time, low error rate. Table I illustrates the node density ranges and the corresponding $T_r$ values. More details will be presented in Section VI.

To meet the various QoS requirements of users in a highly dynamic environment, such as VANETs, one should avoid establishing a backbone path for each MN. Instead, in our approach, the backbone path is established for each group of users satisfying the same QoS requirements and located in the vicinity of each other (forming a cluster). Indeed, each vehicle first queries its neighboring nodes about the optimal backbone route before forwarding its messages. If the required information is available, a positive response will be sent back to the source node, including the optimal route. Otherwise, the query will be relayed to the gateway to select the optimal path according to the new user's QoS requirement. Hence, users with the same QoS requirements and belonging to the same cluster share the same backbone route. The QoS granularity can be determined according to the traffic flow type (audio, video, or background data). This can be translated into upper bounds on the end-to-end delay and the number of hops that a packet can cross to reach the gateway, respectively, as will be shown in Section V. Doing so, we can resolve the scalability problem of routing. Note also that the backbone routes are recomputed by the gateway only if significant changes in the node density are observed.

To illustrate the functionality of IGRP, let us consider the simple example presented in Fig. 1. Assume that the red car moves southward. To send its messages to the gateway, there are three feasible backbone routes, i.e., $A-B-D-F$, $A-C-D-F$, or $A-C-E-F$. In this case, the local gateway selects the $A-C-D-F$ path since it is the most connected path. Indeed, forwarding the packet through this backbone route would be faster than through other routes since they present some disconnection parts. The reason is that, in case of disconnection, the packet has to be carried by the vehicle, whose moving speed is significantly slower than the wireless communication. More formally, Algorithm 1 illustrates the functionality of IGRP.

TABLE II
LIST OF PARAMETERS

| Parameter | Description |
|---|---|
| $BER_{th}$ | thresholds on the bit error rate |
| $D$ | delay in the backbone route |
| $G$ | road graph |
| $H_{th}$ | thresholds on the hop count |
| $L$ | road segment length |
| $m$ | number of road intersections in one route |
| $n$ | number of road segments in one route |
| $N$ | number of nodes in the network |
| $N_g$ | number of generations for the genetic algorithm |
| $P_{cth}$ | thresholds on the connectivity probability |
| $p_z$ | population size |
| $t_p$ | the time needed for a node to process and transmit a message |
| $t_s$ | the simulation time |
| $T_r$ | transmission range |
| $\widehat{S}$ | average speed of nodes on road segment $j$ |
| $\alpha$ | the ratio between road segment length $L$ and transmission range $T_r$ |
| $\beta$ | portion of road segment that does not have any node to forward the message |
| $\gamma$ | node density |
| $\mu$ | mutation rate |
| $\theta$ | crossover probability |

In the following, we present the analytical framework that is used to derive the connectivity probability, end-to-end delay, hop count, and BER. Table II describes the parameters used in the analysis.

**Algorithm 1** IGRP
**In the network**
**if** (a gateway) **then**
  **if** There is a significant change in the node density **then**
    Recalculate the transmission range;
    Recalculate the routes between the different intersections and the gateway;
    Send this data to the nodes in the network;
  **end if**
**end if**
**if** (an MN) **then**
  **if** Has data to transmit **then**
    Queries its neighbors about the optimal backbone route before forwarding its messages.
    **if** the required information is available **then**
      A positive response will be sent back to the source node including the optimal route.
    **else**
      The query will be relayed to the local gateway using normal geographical routing.
      Receive the required information from the gateway;
    **end if**
  **end if**
  Save the updated route information;
  Adjust the transmission range;
  Use this route to forward the data packets to the required destination;
**end if**

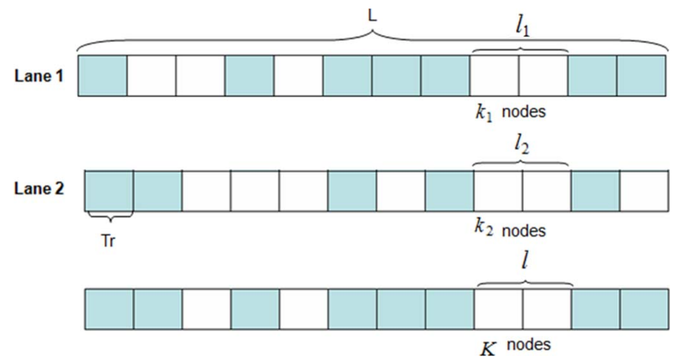



Fig. 2. Two-lane road segment.

## IV. ANALYTICAL FRAMEWORK

As stated before, we model the road network as a graph $G = (V, E)$ consisting of road intersections (i.e., junctions) $v \in V$ and road segments $e \in E$ connecting these junctions. We consider a two-way road scenario, where vehicles are moving in two opposite directions on each road segment and then the message route may contain vehicles moving in the opposite direction. Each road segment has two lanes, as shown in Fig. 2. The road segment is divided into equal slots. Each slot corresponds to one transmission range $T_r$. That is, the two-lane road is divided into slots according to the transmission range of the nodes.

In IGRP, the local gateway needs to have an up-to-date view about the local network topology, so that it can update the estimated statistics about each segment in road graph G. These statistics include the following: 1) the average speed of nodes on segment $j$ (denoted by $\widehat{S}$) and 2) the average spatial node density (denoted by $\gamma_1$ and $\gamma_2$ for lanes 1 and 2, respectively). The average node density is the number of vehicles per lane per kilometer.

In the following, we derive analytical expressions for connectivity probability $P_c$, the BER, delay $D$, and hop count $H_c$ of a backbone route $y$ in a two-way road scenario. Backbone route $y$ consists of a number of intersections $v_1, v_2, \ldots, v_m$, which are connected by a set of road segments $e_1, e_2, \ldots, e_n$, where $n = m - 1$.

### A. Connectivity Probability $P_c$

To compute $P_c$, let us first derive the connectivity probability $P_{cj}$ of the road segment $j$ ($j \in \{e_1, e_2, \ldots, e_n\}$). In this paper, data packets are relayed in the same direction as the vehicles' movement direction, as opposed to the strategy proposed in [18]. To increase the connectivity probability, one may be able to take advantage of the vehicles moving in the opposite direction on a two-way road scenario (see Fig. 2).

In this context, let us define a broken link between two consecutive vehicles $N_i$ and $N_{i+1}$ within a road segment $j$ as a link with length $l = X_i > T_r$. This broken link is fixable if there are vehicles in the opposite direction within the transmission range of each other and connecting $N_i$ to $N_{i+1}$. This implies that the distance between any two consecutive vehicles of the new path on lane 2 must be smaller than transmission range $T_r$.

Let $k_1$ and $k_2$ be random variables denoting the number of vehicles that are present in an interval of length $T_r$ on lanes 1 and 2, respectively (see Fig. 2). Assuming that the vehicles on both lanes are uniformly distributed with node spatial density $\gamma_1$ for lane 1 and $\gamma_2$ for lane 2, then $k_1$ and $k_2$ are Poisson distributed with the probability mass function given as follows:

$$f(k_1) = \frac{(\gamma_1 T_r)^{k_1}}{k_1!} e^{-\gamma_1 T_r} \tag{1}$$

$$f(k_2) = \frac{(\gamma_2 T_r)^{k_2}}{k_2!} e^{-\gamma_2 T_r}. \tag{2}$$

Using (2), the probability $P_f$ that a broken link between two consecutive vehicles $N_i$ and $N_{i+1}$ is fixable can thus be given by

$$\begin{aligned} P_f &= \prod_{k=1}^{\lfloor X_i/T_r \rfloor} (1 - f(k_2 = 0)) \\ &= (1 - e^{-\gamma_2 T_r})^{\lfloor X_i/T_r \rfloor}. \end{aligned} \tag{3}$$

Note that the number of vehicles on lane 1 follows a Poisson distribution and that the distance $X_i$ between $N_i$ and $N_{i+1}$ is exponentially distributed with parameter $\gamma_1$. To compute $P_{cj}$, one should note that more than one broken link on lane 1 can occur. Let $Q$ be a random variable denoting the number of broken links on lane 1. Road segment $j$ will be considered as connected if all the $Q$ links are fixable. Let $P_{c|Q}$ be the conditional connectivity probability, given that there are $Q$ broken links. $P_{c|Q}$ can be written as

$$\begin{aligned} P_{c|Q}(q) &= \prod_{i=1}^{q} P_f \qquad \forall q = 0, 1, \ldots, C_j - 1 \\ &= (1 - e^{-\gamma_2 T_r})^{\sum_{i=1}^{q} \lfloor X_i/T_r \rfloor} \\ &= (1 - e^{-\gamma_2 T_r})^{\left(\alpha - \frac{(C_j - 1 - q)}{\gamma_1 T_r}\right)} \end{aligned} \tag{4}$$

where $C_j$ denotes the number of nodes on lane 1 of road segment $j$. To obtain the total connectivity probability of segment $j$, it is important to know the probability mass function of $Q$ (i.e., $P_Q(q), \forall q = 0, 1, \ldots, C_j - 1$). Recall that a link is broken if the distance between any two consecutive vehicles is larger than $T_r$. Let $P_b$ be the probability that a link $q$ is broken. Since the distance between any two consecutive vehicles is exponentially distributed, it follows that

$$P_b = Pr\{X_i > T_r\} = e^{-\gamma_1 T_r}. \tag{5}$$

Hence

$$P_Q(q) = \binom{C_j - 1}{q} \times P_b^q \times (1 - P_b)^{(C_j - 1 - q)}. \tag{6}$$

Therefore, the total connectivity probability of road segment $j$ can be expressed as

$$P_{cj} = \sum_{q=0}^{C_j - 1} P_{c|Q}(q) \times P_Q(q). \tag{7}$$

Finally, the connectivity probability of the backbone route, which is formed by $n$ road segments, is given by

$$P_c = \prod_{j=1}^{n} P_{cj}. \tag{8}$$

### *B. BER*

A measure of the route quality is the BER, which is mainly affected by the transmission range. Increasing the transmission range increases the BER because of the channel fading and interference. The BER on each link between two consecutive nodes can be given as [18]

$$BER_l = \frac{1}{2}\left(1 - \sqrt{\frac{2\sigma_f^2 \alpha_1 P_t / z^2}{P_{\text{therm}} + 2\sigma_f^2 \alpha_1 P_t / z^2}}\right) \tag{9}$$

where $\alpha_1$ is a constant, $P_t$ is the transmission power, $P_{\text{therm}} = \alpha_2 R_b$ is the thermal noise power, $\alpha_2$ is a constant, $R_b$ is the transmission data rate, and $2\sigma_f^2$ is the mean square value of the signal envelope described by the Rayleigh density function [19]. $z$ is the hop length between two consecutive nodes. Given that the distance $Z$ between two vehicles is exponentially distributed, the probability density function (pdf) of $Z$ can be written as follows:

$$f(Z) = \begin{cases} \frac{\rho e^{-\rho z}}{1 - e^{-\rho T_r}}, & \text{if } 0 \le z \le T_r \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

which represents the conditional pdf of the distance between two consecutive vehicles, given that the distance between them is less than or equal to transmission range $T_r$. Therefore, the expected BER for one link between two consecutive vehicles can be calculated as

$$E\left[BER_l(Z)\right] = \int_0^{T_r} BER_l(z) f_Z(z) dz. \tag{11}$$

In addition,, the BER $BER_j$ of the street segment $j$ is given as follows:

$$BER_j = 1 - (1 - E\left[BER_l(Z)\right])^{(C_j - 1)}. \tag{12}$$

Finally, the $BER$ of a backbone route $y$ formed by $n$ road segments is given by

$$BER = \prod_{j=1}^{n} BER_j. \tag{13}$$

### *C. Delay D*

The end-to-end delay $D$ of a backbone route $y$ defines the time it takes for a data packet to arrive at the gateway from the time it was sent out from the MN. Given the fact that route $y$ from an MN to the gateway consists of a total number of road segments $n$ and each road segment $j$ has an estimated delay $D_j$, then $D$ can be expressed as

$$D = \sum_{j=1}^{n} D_j. \tag{14}$$

Delay $D_j$ depends on the number of MNs $C_j$ traveling on road segment $j$ and on the time required for a message to be transmitted between the two MNs $N_i$ and $N_{i+1}$, which are traveling on road segment $j$. The time required for a message to travel from node $N_i$ to node $N_{i+1}$ depends on the strategy that $N_i$ uses to forward the message. If $N_i$ uses hop-by-hop greedy forwarding, the delay will be the time needed to process and transmit the message, which are denoted as $t_p$. On the other hand, if $N_i$ uses the carry-and-forward strategy, the message carried by $N_i$ will travel with the same speed $S_i$ as that of MN $N_i$. Thus, the delay depends on $S_i$ and the distance traveled by $N_i$ while carrying the message until it is able to forward the message to the next MN $N_{i+1}$, i.e., when it comes within the transmission range of $N_{i+1}$. To estimate delay $D$, two cases are considered. Let $\alpha$ be defined as $\alpha = (L/T_r)$.

*Case 1:* One vehicle is allowed to forward the message along the road segment. This case occurs if segment length $L$ is less than one transmission range $T_r$. In this case, $\alpha \leq 1$. The delay of that segment will be $t_p$, where $t_p$ is the time that the vehicle needs to process and transmit the message. In our study, we assumed an average value of $t_p$ to reflect the behavior of a multichannel VANET. Indeed, in such networks, where interferer wireless links operate on different channels, multiple contentionless parallel transmissions can occur. In doing so, collisions and interferences between transmissions over interferer links are avoided. This assumption has been used by several works such as [38]–[41]. It is worth noting that the elaborated analytical model can reflect the real behavior of the VANET as long as the use of different channels is ensured. In addition, an efficient channel assignment algorithm needs to be used to avoid contention and collisions, and to enable optimal spatial reuse of available channels.

*Case 2:* More than one vehicle are allowed to forward messages along the road segment. This case occurs when the road segment length is larger than the transmission range (i.e., $\alpha \geq 1$), which is likely to be the case in real networks. In this context, more than one hop is needed to forward the message along that segment.

Let $K$ be a random variable denoting the number of vehicles present in the interval of length $T_r$ on both lanes. Likewise, $K$ follows a Poisson distribution with the following probability mass function:

$$f(K) = \frac{((\gamma_1 + \gamma_2)T_r)^K}{K!} e^{-(\gamma_1+\gamma_2)T_r}. \tag{15}$$

To compute the delay on the road segment, the strategy that the MN uses to forward the message is considered. If the message is forwarded hop by hop, the delay on such a link will be $t_p$ as in the first case. On the other hand, if the message is carried and forwarded by nodes, an estimate of the portion $\beta$ of the road segment that does not have any node to forward the message is needed. In this case, the last node on that portion receiving the message is allowed to carry and forward the message along that portion. The vehicle will not transmit the message until it comes within the transmission range of another vehicle. This portion $(\beta)$ can be estimated as

$$\beta = f(K = 0) = e^{-(\gamma_1+\gamma_2)T_r}. \tag{16}$$

In this case, the average delay can be computed using the average speed of nodes on road segment $j$; recall that $C_j$ is the number of nodes on lane 1 of road segment $j$.

Thus, the average delay on road segment $j$ can be given as

$$D_j = \begin{cases} t_p, & \text{if } \alpha \leq 1 \\ \alpha(1-\beta)t_p + \beta\frac{L}{\widehat{S}}, & \text{otherwise} \end{cases} \tag{17}$$

where $\widehat{S}$ is the average speed of nodes on lane 1 of road segment $j$ given as

$$\widehat{S} = \frac{\sum_{k=1}^{C_j} S_k}{C_j}. \tag{18}$$

### D. Hop Count $H_c$

For a given backbone route $y$, the number of hops the message travels on one road segment $j$ is controlled by the length $L$ of the road segment and the transmission range $T_r$ of the nodes traveling on that road segment. If $L$ is less than $T_r$ (, i.e., $\alpha \leq 1$), then one hop will be enough to transmit the message on that road segment. On the other hand, if $L$ is larger than $T_r$ (, i.e., $\alpha \geq 1$) the message can be transmitted hop by hop, or it can be carried and forwarded. Thus, the average hop count on road segment $j$ can be given as

$$H_{cj} = \begin{cases} 1, & \text{if } \alpha \leq 1 \\ \alpha(1-\beta) + \beta C_j, & \text{otherwise.} \end{cases} \tag{19}$$

Accordingly, the hop count of a backbone route $y$ formed by $n$ road segments is given by

$$H_c = \sum_{j=1}^{n} H_{cj}. \tag{20}$$

### E. Estimating the Transmission Range

It is worth noting that, using (7), the gateway will estimate the transmission range $T_r$ that each vehicle should use along each road segment $j$ to achieve high connectivity. To do so, it uses the node density value of road segment $j$ to decide on $T_r$ that guarantees a probability of connectivity approaching 1.

To illustrate this, let us consider Figs. 3 and 4. Fig. 3 shows the relationship between road density and transmission range for different values of probability of connectivity. As we can see, when the node density is low, we need to increase the transmission range to achieve high connectivity. On the other hand, when the node density is high, a small transmission range value is enough to guarantee high connectivity.

From Fig. 4, we notice that the connectivity probability increases with the increase in transmission range. For example, when the density is 10/750, a transmission range of 300 m gives a connectivity probability approaching 1, which shown as point 1. On the other hand, when the density decreases to 4/750, the transmission range should be increased to 600 m to achieve a connectivity approaching 1, which is shown as point 2.

Likewise, the BER increases when increasing the transmission range. To achieve a low BER, the transmission range

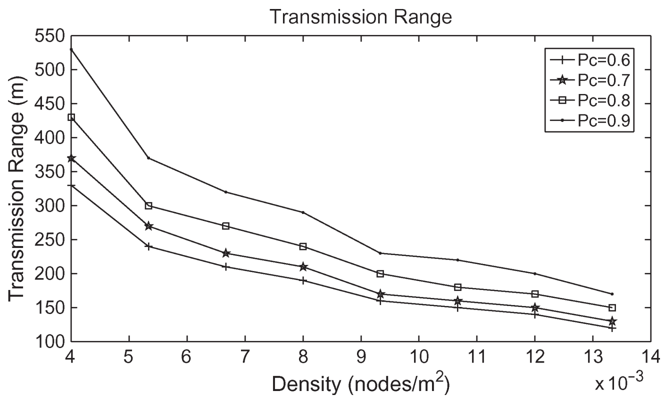



Fig. 3. Transmission range as a function of node density.

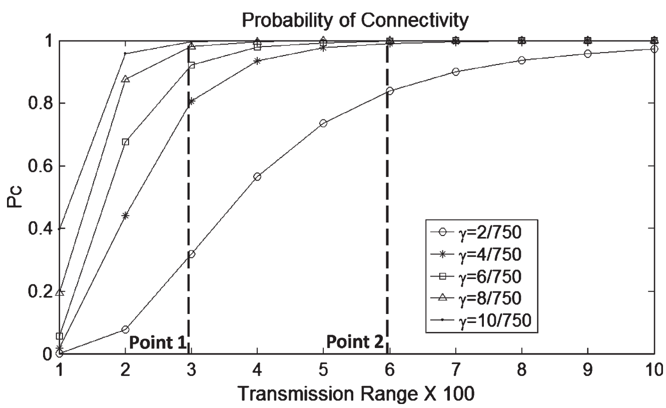



Fig. 4. Probability of connectivity changes with both transmission range and node density.

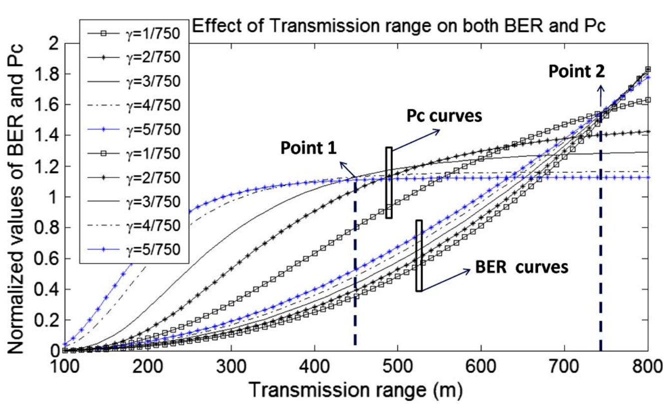



Fig. 5. BER and connectivity probability change with both transmission range and node density.

should then be decreased. Therefore, $T_r$ should be selected, so that a tradeoff between increasing the connectivity probability and decreasing the BER is achieved. Fig. 5 shows the effect of increasing $T_r$ on both the connectivity probability and BER.

Fig. 5 shows the effect of increasing $T_r$ on both the connectivity probability and BER for different node density values. As we can see, for low node density, $T_r$ is selected to be the point of intersection between the two curves. As the node density increases, the connectivity probability reaches 1 at low $T_r$ values. Therefore, in our simulations, $T_r$ is selected to be the value that results in connectivity 1 and, at the same time, results in the lowest BER. For example, when node density $\gamma$ is 5/750 (which is a high density), the $T_r$ value is selected to be 450 m, which is shown as point 1 in Fig. 5. On the other hand, when the node density $\gamma$ is 1/750 (which is a low density), the $T_r$ value is selected to be 750 m, which is shown as point 2 in Fig. 5.

## V. Formulating Message Routing as an Optimization Problem

In this section, we address the problem of finding the optimal or the near-optimal backbone route $y$, which consists of a number of intersections $v_1, v_2, \ldots, v_m$ connected by a set of road segments $e_1, e_2, \ldots, e_n$; $n = m - 1$. Note that intersection $v_1$ is the first intersection in the backbone route that is connected to the source node and that $v_m$ is the last intersection in the route that is connected to the gateway.

The optimal or the near-optimal backbone route is the route that maximizes the probability of connectivity while satisfying the constraints on tolerable end-to-end delay, hop count, and BER. The gateway uses this objective function to decide on the backbone routes used by the MNs in its vicinity to forward their data packets. Note that the delay constraint is translated into an upper bound $D_{\mathrm{th}}$, whose values depend on the intended VANET applications. For instance, assigning low values for $D_{\mathrm{th}}$ corresponds to delay-sensitive applications. However, high values of $D_{\mathrm{th}}$ refer to delay-tolerant applications.

Hence, our approach can be formulated as an optimization problem with the objective function given as

$$\max_{y} P_c(y) \tag{21}$$

$$P_c(y) = \prod_{j=1}^{n} P_{cj}(y) \tag{22}$$

subject to

$$D(y) = \sum_{j=1}^{n} D_j(y) \leq D_{\mathrm{th}} \tag{23}$$

$$H_c(y) = \sum_{j=1}^{n} H_{cj}(y) \leq H_{\mathrm{th}} \tag{24}$$

$$BER(y) = \prod_{j=1}^{n} BER(y) \leq BER_{\mathrm{th}} \tag{25}$$

where $P_c(y)$ is the connectivity probability of route $y$, and $D_{\mathrm{th}}$, $H_{\mathrm{th}}$, and $BER_{\mathrm{th}}$ are thresholds on the tolerable end-to-end delay, hop count, and BER, respectively.

It is worth noting that our problem previously described is nonprobabilistic hard [20]. Hence, to solve it, we propose a GA, which is described in the following section, since this kind of heuristic methods yields better results for routing problems [21]–[25].

Fig. 6 shows the flowchart of the proposed GA, which includes the following components: solution representation, initialization, evaluation, selection, crossover, mutation, and termination.

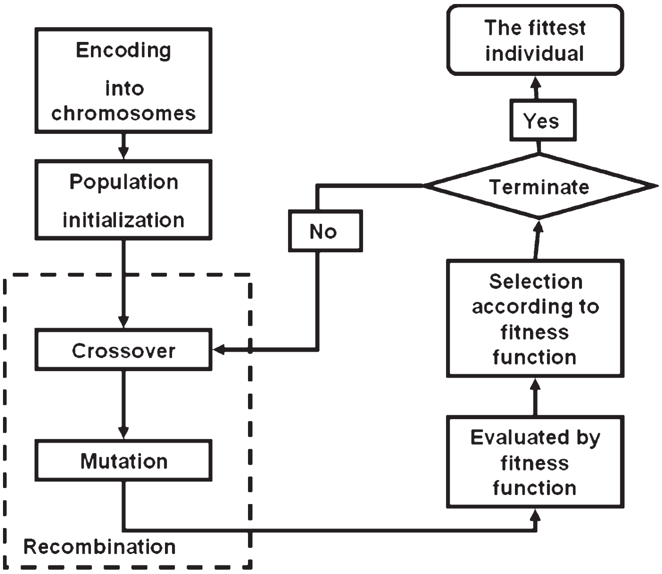



Fig. 6. Flowchart of the proposed GA.

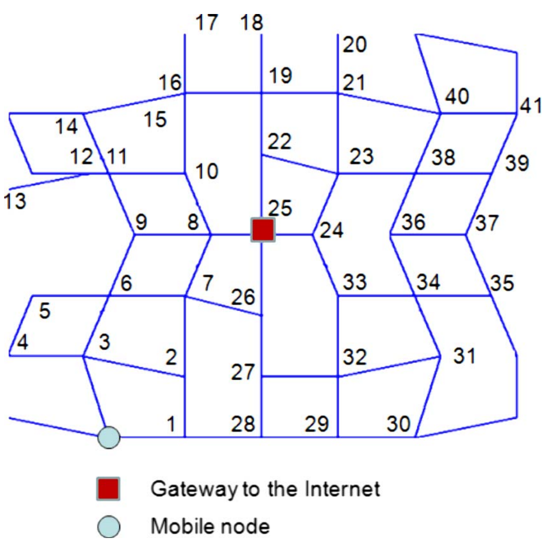



Fig. 7. Road map used in the simulation.

### A. Solution Representation and Initialization

Choosing an appropriate representation to encode the feasible solutions is the first step in applying GAs. This representation should be suitable for the fitness function and the genetic operations. In our approach, a natural encoding scheme would be to define each intersection in the backbone route as a gene. The backbone route consists of the identification number of each selected intersection. Then, the ordered intersections in one route can be represented as a chromosome. Therefore, each feasible solution $y$ consists of one chromosome, which is denoted as $v_1, v_2, \ldots, v_m$. For example, routes 1-2-7-8-25, 1-28-27-26-25, and 3-6-9-8-25 in Fig. 7 are chromosomes. Thus, an individual (or chromosome) is a vector containing the ordered intersections.

Our GA search is conducted from a population of solutions. The initial population is generated by randomly selecting feasible solutions. Each solution or chromosome begins with the intersection adjacent to the MN. The next gene is constructed from a randomly selected intermediate intersection. Then, the process randomly chooses the next intermediate intersection in the backbone route, and the process stops when the next intersection corresponds to that adjacent to the Internet gateway. It is important to ensure that the solution is feasible, i.e., it satisfies the following two conditions: 1) Each of the two consecutive intersections in the route are connected by a backbone link. 2) The route satisfies the QoS constraints. A population of individuals can be constructed by continuing this process until generating a certain number of chromosomes called population size $p_z$.

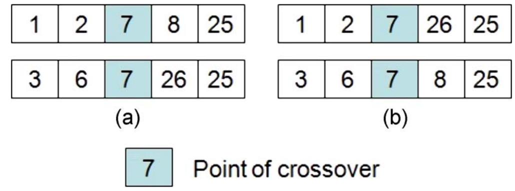



Fig. 8. One point crossover operator. (a) Two chromosomes with 7 as crossover point. (b) Two new offsprings.

### B. Evaluation

A value for fitness function $f(y)$ is assigned to each chromosome $y$, depending on how it is close to solving the problem. Then, the best individuals are selected, depending on their fitness function. Since our objective is to maximize the connectivity probability given in (21), fitness function $f(y)$ can be defined as follows:

$$f(y) = P_c(y). \tag{26}$$

### C. Selection

During the selection operation, the quality of the population is improved by giving the high-quality solutions a better chance to produce offsprings, which will be part of the next generation. In our implementation, we use the roulette wheel selection strategy. Doing so, the chromosomes are selected based on a probability that is proportional to its normalized fitness value, i.e., the probability of choosing a chromosome $y$ corresponds to

$$P_{\text{selection}} = \frac{f(y)}{\sum_{y=1}^{p_z} \left(f(y)/p_z\right)} \tag{27}$$

where $p_z$ is the population size.

### D. Crossover

The crossover operation is usually executed with a probability $\theta$. One possible crossover operator is the *one point crossover*, where two chromosomes are selected from the current population, and then, a common intermediate gene is randomly selected. That is, the *one point crossover* operator finds an intermediate intersection called point of crossover, which is common to the two selected routes. Then, it swaps the second part of each selected route beyond the point of crossover to form two new offsprings. Fig. 8(a) shows two randomly selected chromosomes with 7 as crossover point, and Fig. 8(b) shows two new offspring. Note that it is important to check that the new individuals are feasible.

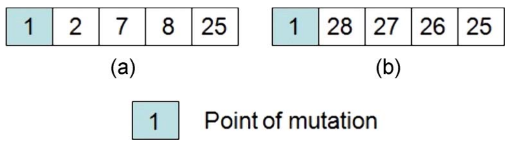



Fig. 9. Uniform mutation operator. (a) Chromosome with 1 as a point of mutation. (b) New offspring.

### *E. Mutation*

Mutation is an operator that causes random changes in the genes inside one chromosome. Therefore, mutation causes diversion in the genes of the current population, which prevents the solution from being trapped in a local optimum. Mutation is performed on the current population with rate $\mu$. In our implementation, we use a *uniform mutation* operator. Thus, after choosing any individual from the population with equal probabilities, we randomly pick an intermediate gene (intersection) and then randomly choose the adjacent intersection (see Fig. 9). It is important to verify that the new individual is a feasible solution.

### *F. Termination*

The termination criteria, which is shown in Fig. 6, can be based on the total number of generations, maximum computing time, an acceptable threshold of the standard deviation between solutions in one population, or a hybrid termination criteria among them. In our implementation, we use the maximum number of generations as a termination criteria.

## VI. Numerical and Simulation Results

In this section, we compare our proposal with respect to three benchmark routing protocols, i.e., GPSR [11], GPCR [12], and OLSR [6]. To this end, we developed our own discrete-event simulator using Matlab. We used the IEEE 802.11p physical layer (PHY), which defines an international standard for wireless access in vehicular environments. We started from an available MATLAB/SIMULINK model, i.e., the IEEE 802.11a, to obtain IEEE 802.11p PHY. We used also multipath Rayleigh fading.

To implement IGRP, we implemented first the location-service management protocol RLSMP [15]. The overhead generated by this protocol has already been presented in our previous work [15]. In our experiments, we consider different scenarios representing morning rush hours (i.e., dense network), noontime having intermediate density, and nighttime with low density (sparse network). To do so, we use different numbers of vehicles, given that the area of the simulated network is fixed. The number of nodes is varied between 150 and 620 nodes. In addition, the mobility of nodes is modeled based on a given street map where the mobility generator SUMO [27] is used to generate vehicle mobility traces. The parameters settings in our experiments are listed in Table III, where $t_s$ denotes the simulation time and $N_g$ is the number of generations for our GA. Additional GA parameters are mutation rate $\mu$, crossover rate $\theta$, and population size $p_z$.

TABLE III
Parameter Settings

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $t_s$ | 1000 sec | $T_r$ | 250m |
| $t_p$ | 3 msec [26] | $\mu$ | 0.3 |
| $D_{th}$ | 110 $\sim$ 170 (default 130) | $\theta$ | 0.8 |
| $H_{th}$ | 32 $\sim$ 44 (default 40) | $N_g$ | 20 |
| $S_k$ | 50 $\sim$ 65 km/h (default 50 km/h) | $p_z$ | 10 |

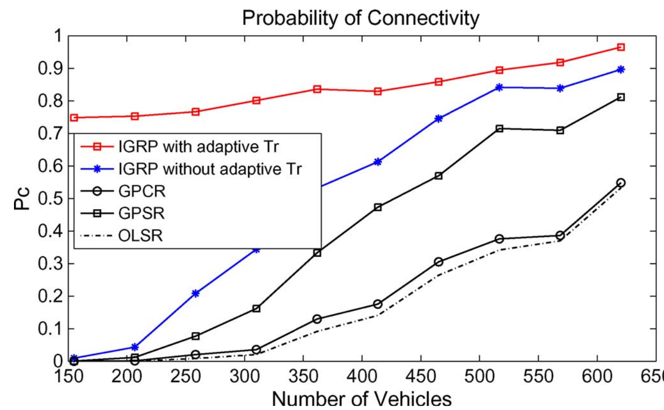



Fig. 10. Connectivity probability.

To get an insight into our mathematical model, we first compare between the routes chosen by the gateway using the mathematical model and the simulation environment. In both scenarios, there is a 97.5% confidence interval that the chosen routes are the same, which demonstrates the accuracy of our analytical model.

We ran experiments for IGRP under two scenarios. The first one concerns nodes with fixed transmission range $T_r$ and used to simulate the basic performance of IGRP. In this case, $T_r$ equals to 250 m. In the second scenario, the $T_r$ values are no longer constant and are adapted to the changes in nodes' densities on the different road segments. In this case, we used the values depicted in Table I. Compared with the basic IGRP (i.e., with fixed $T_r$ values), IGRP with adaptive $T_r$ can achieve higher connectivity probability, less delay, less number of hops, and less BER, as shown in Figs. 10 and 12.

In general, the performance of the basic IGRP approaches that of IGRP with adaptive $T_r$ when increasing the number of nodes due to the increase in the nodes' density. This behavior is shown in Fig. 10 and Fig. 12(a) and (b) since, in a high-density environment, the transmission range of IGRP with adaptive $T_r$ is reduced.

Let us now focus on the comparison of the performance of IGRP with that of GPCR, GPSR, and OLSR. Fig. 10 shows the connectivity probability for all protocols as a function of the number of nodes in the network. As expected, IGRP chooses routes that have high connectivity to relay messages with delay that is below the maximum tolerable delay threshold, particularly in low-density networks. Indeed, IGRP selects routes with higher number of nodes to achieve higher connectivity probability and, at the same time, meet the delay, hop count, and BER constraints. On the other hand, GPCR and GPSR select the nodes on routes that have minimum distance from the gateway. Therefore, they select the path with a minimum number of

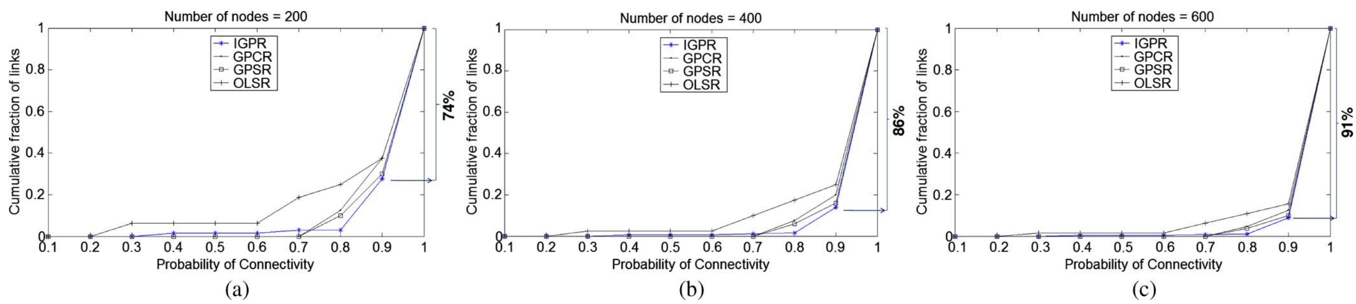



Fig. 11. CDF of the probability of connectivity for $N$ nodes. (a) $N = 200$. (b) $N = 400$. (c) $N = 600$.

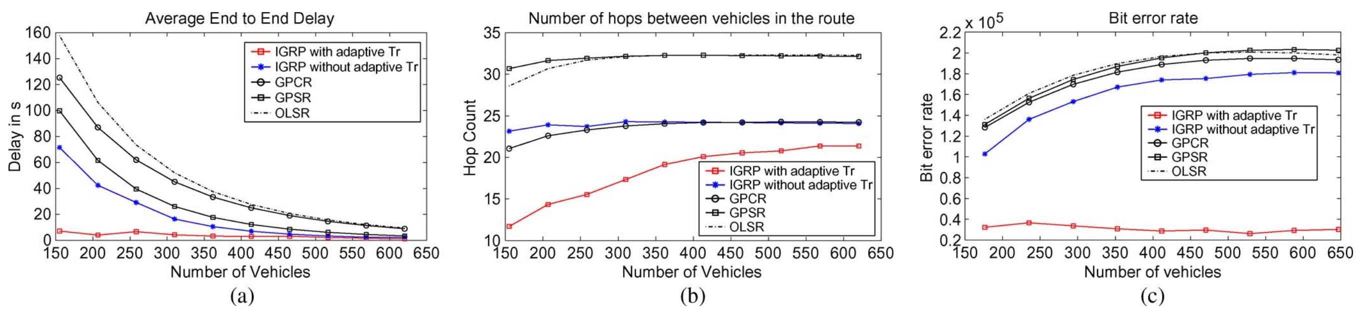



Fig. 12. Impact of variation of the number of nodes on the end-to-end delay, hop count, and BER. (a) End-to-end delay. (b) Hop count. (c) BER.

intersections, without taking into consideration the connectivity degree. As such, in GPCR and GPSR, more nodes are allowed to use the store-and-forward mechanism, which decreases the probability of connectivity and increases the delay (see Fig. 10). For OLSR, since the intermediate nodes build routes based on the topology information in the local table, the message forwarded on these routes may face a route failure due to the stale information resulting from the high mobility of nodes. At that time, the intermediate node waits for a route maintenance reply to begin to forward the message again, which deteriorates the probability of connectivity.

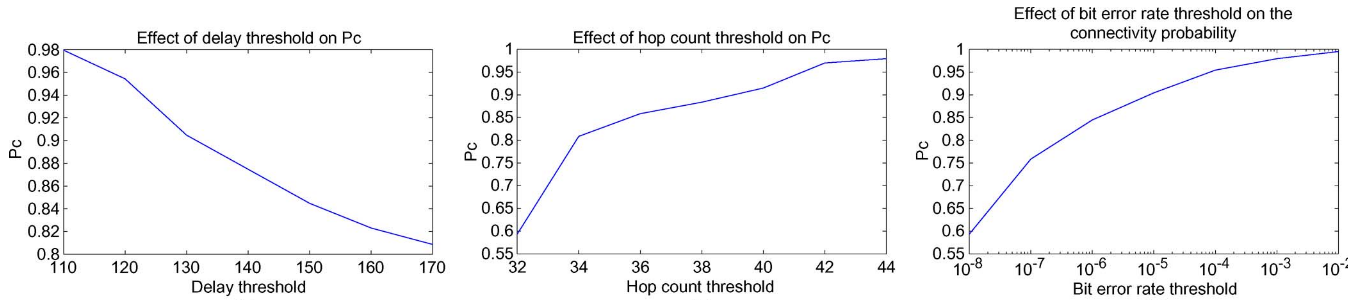
Fig. 11 compares the cumulative distributed function (cdf) of the probability of connectivity for the links used by the aforementioned protocols for 200, 400, and 600 nodes. From this figure, we can notice that IGRP often uses links with high probability of connectivity (i.e., higher than 0.9) compared to with the remaining protocols. Indeed, 74%, 86%, and 91% of the links with $0.9 < P_c < 1$ are used by IGRP for the 200-, 400-, and 600-node cases, respectively, against 62%, 75%, and 84% for GPSR; 62%, 80%, and 86% for GPCR; and 70%, 84%, and 90% for OLSR.

Fig. 12(a)–(c) shows the end-to-end delay, hop count, and BER for all protocols as a function of the number of nodes in the network, respectively. Note that the results of Fig. 12 are obtained assuming a multichannel environment. Indeed, in our simulations, different nonoverlapping channels are assigned to interferer links, thus allowing multiple contentionless parallel transmissions. This can be realized using multiple radio interfaces and efficient interference-aware channel assignment (e.g., [42]). In such scenario, we can notice that the delay decreases with the increase in network density. The reason is that, in case of low-density networks, the packet has to be carried by the vehicle, whose moving speed is significantly slower than the wireless communication. On the other hand, in high-density networks, wireless transmission over different channels is more often used. This can indeed be realized since, in a high-density environment, the transmission range of IGRP with adaptive $T_r$ is reduced. As such, the average number of interferer links will be reduced. This significantly decreases the end-to-end delay, as observed in Fig. 12(a). It is worth noting that, for the case of single radio interface, interference between vehicles may significantly degrade the end-to-end delay.

Moreover, Fig. 12(b) shows that IGRP constructs routes with enough number of nodes to avoid disconnectivity but, at the same time, does not choose routes that have a very high number of nodes (i.e., high density), which results in less network contention and, then, lower BER, as shown in Fig. 12(c).

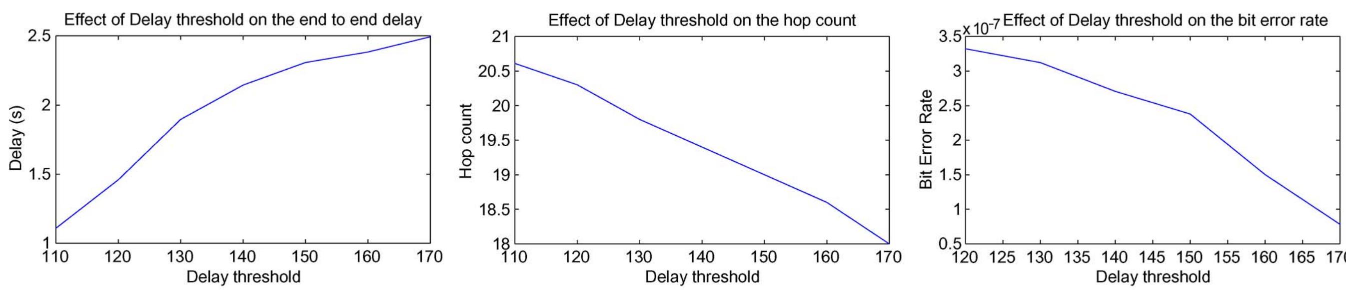
On the other hand, GPCR chooses the next road intersection without considering if there are enough nodes to relay the message. As a result, less number of nodes are selected [as shown in Fig. 12(b)], but a relatively higher delay is experienced due to frequent use of the carry-and-forward strategy [as seen in Fig. 12(a)] and a relatively lower BER due to the low dense routes [as shown in Fig. 12(c)]. Considering OLSR, routes fail often and may encounter loops very often, which increases the number of nodes in the route toward the gateway and, consequently, increases the delay and then the BER. Regarding GPSR, since it is a position-based routing protocol, it selects routes that have nodes close to each other, which results in higher connectivity, less delay, and a very high number of nodes and BER, compared with both GPCR and OLSR.

Fig. 13(a)–(c) shows the impact of the delay threshold $(D_{\rm th})$, the hop count threshold $(H_{\rm th})$, and the BER threshold $(BER_{\rm th})$ on the connectivity probability. From Fig. 13(a), we can notice that the connectivity probability decreases when increasing $D_{\rm th}$. This is related to the fact that more vehicles are allowed to carry the message, which will be transmitted with the same speed as that of the vehicles. As such, routes will



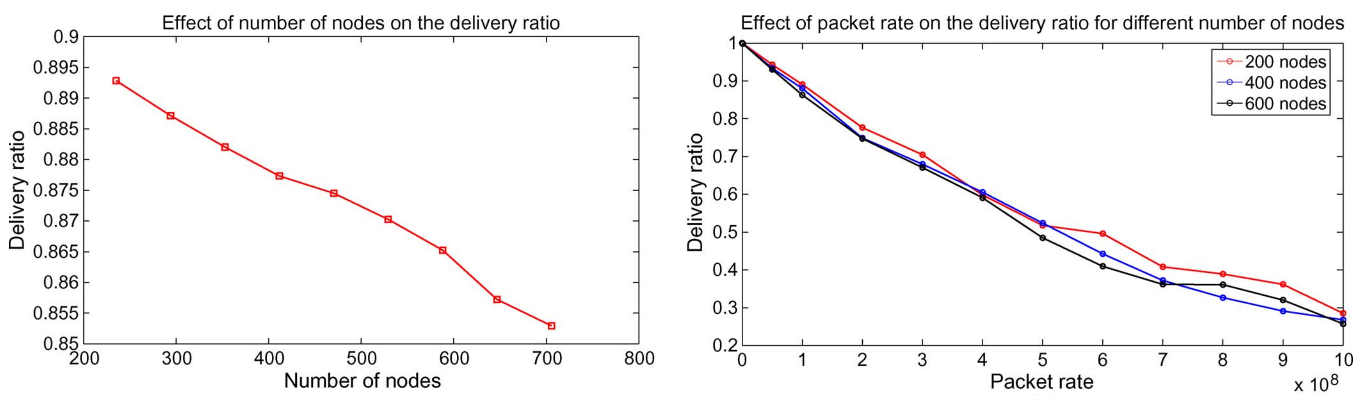

(a) (b) (c)

Fig. 13. Impact of the threshold levels on the probability of connectivity $P_c$. (a) Impact of $D_{\mathrm{th}}$. (b) Impact of $H_{\mathrm{th}}$. (c) Impact of $BER_{\mathrm{th}}$.




(a) (b) (c)

Fig. 14. Impact of $D_{\mathrm{th}}$ on the end-to-end delay, hop count, and BER of a backbone route. (a) End-to-end delay. (b) Hop count. (c) BER.




(a) (b)

Fig. 15. Delivery ratio of IGRP when varying the number of nodes and the packet rate. (a) Impact of variation of the number of nodes. (b) Impact of packet rate variation.

have more nodes that are distant by more than the transmission range, thus decreasing connectivity probability.

Fig. 13(b) shows that the connectivity probability increases when increasing $H_{\mathrm{th}}$. This is due to the fact that routes with more and more vehicles are allowed to be selected. This enforces hop-by-hop forwarding and may result in higher connectivity probability.

Fig. 13(c), on the other hand, shows that the connectivity probability increases when increasing $BER_{\mathrm{th}}$. Indeed, increasing the BER threshold allows the selected routes to have more and more nodes, which causes contention in the network but, at the same time, increases the connectivity probability of these selected routes.

Note that variations of the threshold levels (i.e., $D_{\mathrm{th}}$, $H_{\mathrm{th}}$, or $BER_{\mathrm{th}}$) do not affect the performance of GPSR, GPCR, or OLSR since they do not consider these parameters in the routing process.

Fig. 14(a)–(c) shows the effect of increasing the tolerable delay threshold on the delay, hop count, and BER of a selected backbone route, respectively. Increasing the tolerable delay threshold will allow the routes to have more and more nodes that carry and forward the messages, which increases the delay needed to deliver the messages, as shown in Fig. 14(a). However, this will decrease the number of nodes in the selected routes, which results in less hops, as shown in Fig. 14(b), and, at the same time, will decrease the contention in the network, thus decreasing the BER [see Fig. 14(c)].

Fig. 15(a) and (b) shows the effect of both the number of nodes and the packet rate on the delivery ratio, respectively. From those figures, we can see that the delivery ratio decreases when increasing the number of nodes and the packet rate due to the increase in the network contention.

It is worth mentioning that our GA does not guarantee optimality but rather gives optimal or near-optimal solutions. To

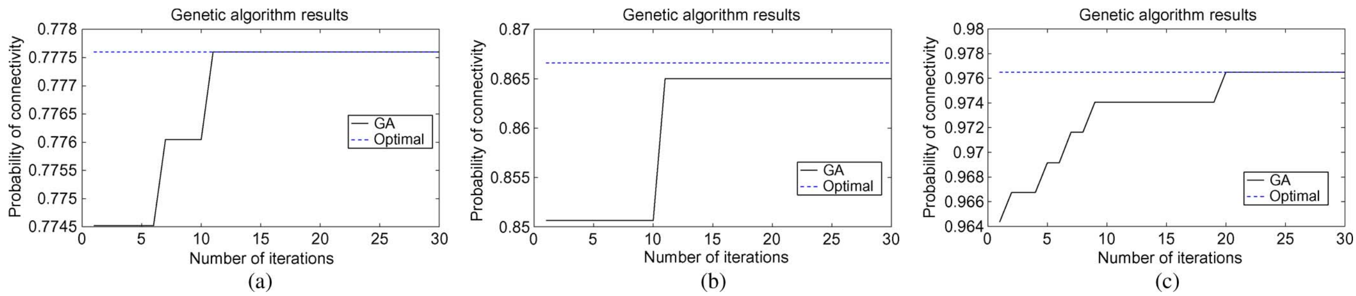



Fig. 16. GA scenarios. (a) Scenario 1. (b) Scenario 2. (c) Scenario 3.

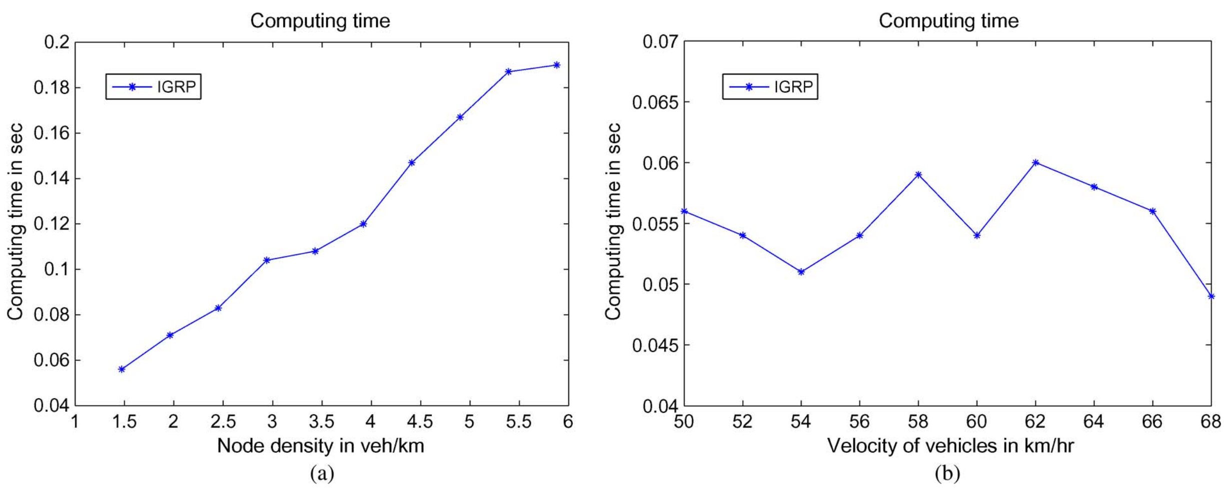



Fig. 17. Computation time of the proposed GA when varying the node density and node velocity. (a) Impact of node density variation. (b) Impact of node velocity variation.

illustrate this, Fig. 16 shows the convergence of our proposed GA using different scenarios. Specifically, Fig. 16(a) shows a scenario where our algorithm converges to the optimal solution after 11 iterations. On the other hand, Fig. 16(b) shows a second scenario where our algorithm could not reach the optimal solution, but it could find a near-optimal solution that is just 0.16% less than the optimal value. Finally, Fig. 16(c) shows that our GA could reach the optimal solution but after a large number of iterations (after 20 iterations). In this case, it is up to the decision maker if he/she would like to have exact solution after 20 iterations or be satisfied by the quality of the solution, which is just 0.24% less than the optimal value, which can be achieved after only nine iterations.

Finally, Fig. 17(a) and (b) shows the processing time needed by the gateway to compute the backbone routes as a function of the node density and the node velocity, respectively.

In Fig. 17(a), node density is varied between 1.5 and 6 veh/km. From that figure, we can see that the gateway computation time increases with the increase in node density. It is about 60 and 200 ms at low and high densities, respectively. This is due to the fact that the gateway will need more processing time to be able to consider more nodes in its decision when the node density increases. Fig. 17(b), on the other hand, shows the impact of the average node velocity on the computation time for the 200-node-network case. In each simulation, the average vehicle speed is chosen between 50 and 65 km/h and remains the same for all road segments. In addition, the node density is kept constant by replacing every vehicle, leaving the road segment by a new one entering the road segment. We can see that the gateway computation time does not significantly vary when we vary the average node velocity and lies between 50 and 60 ms. This confirms that the constructed backbone routes are not affected by the individual nodes' mobility but rather depend on the node density [as shown in Fig. 17(a)].

## VII. Conclusion

In this paper, we have proposed a new approach for routing messages in city-based environments that takes advantage of the roads layouts to improve the performance of routing in VANETs. Our proposal IGRP tends to satisfy QoS constraints on four performance metrics: 1) tolerable end-to-end delay; 2) connectivity probability; 3) bandwidth usage; and 4) BER. To achieve this, we have formulated the QoS routing problem as a constrained optimization problem. We have also derived analytical expressions for the four performance metrics in a two-way street scenario. Using both analytical and simulation approaches, we have compared our proposal with GPSR, GPCR, and OLSR. We have found that IGRP achieves better performance. Indeed, it selects routes that are connected and, at the same time, satisfies thresholds on the end-to-end delay, hop count, and BER. As such, our solution stands out as a promising candidate for large-scale ad hoc networks, such as VANETs.

## REFERENCES


[1] Y. Toor, P. Muhlethaler, A. Laouiti, and A. Fortelle, "Vehicular ad hoc networks: Applications and related technical issues," *IEEE Commun. Surveys Tutorials*, vol. 10, no. 3, pp. 74–88, 3rd Quarter, 2008.

[2] H. Hartenstein and K. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 164–171, Jun. 2008.

[3] J. Zhao and G. Cao, "VADD: Vehicle-assisted data delivery in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1910–1922, May 2008.

[4] C. Lochert, "A routing strategy for vehicular ad hoc network in the city environments," in *Proc. IEEE Intell. Veh. Symp.*, 2003, pp. 156–161.

[5] V. Naumov and T. Gross, "Connectivity-Aware Routing (CAR) in vehicular adhoc networks," in *Proc. IEEE INFOCOM*, 2007, pp. 1919–1927.

[6] T. Clausen and P. Jacquet, "RFC 3626: The optimized link-state routing protocol," Internet Eng. Task Force Request For Comments, 2003.

[7] D. Johnson and D. Maltz, "Dynamic source routing in ad hoc wireless networks," *Mobile Comput.*, vol. 353, pp. 153–181, 1996.

[8] C. Perkins and E. Royer, "Ad-hoc on demand distance vector routing," in *Proc. 2nd IEEEWorkshop Mobile Comput. Syst. Appl.*, 1999, pp. 90–100.

[9] C. Lochert, B. Scheuermann, C. Wewetzer, A. Luebke, and M. Mauve, "Data aggregation and roadside unit placement for a VANET traffic information system," in *Proc. ACM VANET*, 2008, pp. 58–65.

[10] T. Camp, J. Boleng, and L. Wilcox, "Location information services in mobile ad hoc networks," in *Proc. IEEE ICC*, 2001, pp. 3318–3324.

[11] B. N. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *Proc. ACM MOBICOM*, 2000, pp. 243–254.

[12] C. Lochert, M. Mauve, H. Fusler, and H. Hartenstein, "Geographic routing in city scenarios," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 9, no. 1, pp. 69–72, Jan. 2005.

[13] T. Li, S. K. Hazra, and W. Seah, "A position-based routing protocol for metropolitan bus networks," in *Proc. IEEE VTC-Spring*, 2005, pp. 2315–2319.

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

[15] H. Saleet, O. Basir, R. Langar, and R. Boutaba, "Region-based location-service-management protocol for VANETs," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 917–931, Feb. 2010.

[16] GB Traffic Volumes, May 2005. [Online]. Available: www.mapmechanics.com

[17] SmartView. [Online]. Available: www.yahoo.com

[18] S. Panichpapiboon and W. Pattara-atikom, "Connectivity requirements for self-organizing traffic information systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3333–3340, Nov. 2008.

[19] H. Cheng and Y.-D. Yao, "Power adaptation for multihop networks with end-to-end BER requirements," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3445–3454, Sep. 2010.

[20] Z. Wang and J. Crowcroft, "Quality-of-service routing for supporting multimedia applications," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 7, pp. 1228–1234, Sep. 1996.

[21] X. Cheng, "Multiple description video communications in wireless ad hoc networks," M.S. thesis, Fac. Virginia Polytech. Inst., State Univ., Blacksburg, VA, 2005.

[22] J. W. Mann and G. D. Smith, "A comparison of heuristics for telecommunications traffic routing," in *Modern Heuristic Search Methods*, V. J. Rayward-Smith, I. H. Osman, C. R. Reeves, and G. D. Smith, Eds. New York: Wiley, 1996.

[23] J. Schaffer and L. Eschelman, "Combinatorial optimization by genetic algorithms: The value of the genotype/phenotype distinction," in *Modern Heuristic Search Methods*, V. Rayward-Smith, I. Osman, C. Reeves, and G. Smith, Eds. New York: Wiley, 1996.

[24] D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[25] J. M. G. Lopez, M. Imine, and O. B. Madsen, "Network planning using GA for regular topologies," in *Proc. IEEE ICC*, 2008, pp. 5258–5262.

[26] S. Haykin and M. Moher, *Modern Wireless Communication*. Englewoods, NJ: Prentice-Hall, 2005.

[27] SUMO: Simulation of Urban MObility. [Online]. Available: http://sumo.sourceforge.net

[28] B. Ducourthial, Y. Khaled, and M. Shawky, "Conditional transmissions: Performance study of a new communication strategy in VANET," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3348–3357, Nov. 2007.

[29] C. Perkins and E. Royer, "Ad-hoc on demand distance vector routing," in *Proc. 2nd IEEE Workshop Mobile Comput. Syst. Appl.*, 1999, pp. 90–100.

[30] C. Perkins, *Ad Hoc Networking*. Reading, MA: Addison Wesley, 2001.

[31] S. Basagni, I. Chlamtac, V. Syrotiuk, and B. Woodward, "A distance routing effect algorithm for mobility (DREAM)," in *Proc. ACM MOBICOM*, 1998, pp. 76–84.

[32] J. Bernsen and D. Manivannan, "Routing protocols for vehicular ad hoc networks that ensure quality of service," in *Proc. ACM ICWMC*, 2008, pp. 1–6.

[33] F. Bai, N. Sadagopan, and A. Helmy, "The important framework for analyzing the impact of mobility on performance of routing protocols for ad hoc networks," *Ad Hoc Netw.*, vol. 1, no. 4, pp. 383–403, Nov. 2003.

[34] T. Taleb, M. Ochi, N. Kato, and Y. Nemoto, "An efficient vehicle-heading based routing protocol for VANET networks," in *Proc. WCNC*, 2006, pp. 2199–2204.

[35] K. C. Lee, J. Haerri, U. Lee, and M. Gerla, "Enhanced perimeter routing for geographic routing protocols in urban vehicular scenarios," in *Proc. IEEE Globecom Workshops*, 2007, pp. 1–10.

[36] Z. Mo, H. Zhu, K. Makki, and N. Pissinou, "MURU: A multi-hop routing protocol for urban vehicular ad hoc networks," in *Proc. 3rd Annu. Int. Conf. Mobile Ubiquitous Syst.—Networking and Services*, 2006, pp. 1–8.

[37] A. Skordylis and N. Trigoni, "Delay-bounded routing in vehicular ad-hoc networks," in *Proc. MOBIHOC*, 2008, pp. 341–350.

[38] T. K. Mak, K. P. Laberteaux, R. Sengupta, and M. Ergen, "Multichannel medium access control for dedicated short-range communications," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 349–366, Jan. 2009.

[39] X. Xie, B. Huang, S. Yang, and T. Lv, "Adaptive multi-channel MAC protocol for dense VANET with directional antennas," in *Proc. 6th IEEE CCNC*, 2009, pp. 1186–1190.

[40] H. Su and X. Zhang, "Clustering-based multichannel MAC protocols for QoS provisioning over vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3309–3323, Nov. 2007.

[41] T. Kim, S. Jung, and S. Lee, "CMMP: Clustering-Based Multichannel MAC Protocol in VANET," in *Proc. 2nd Int. Conf. Comput. Elect. Eng.*, 2009, pp. 380–383.

[42] P. Fazio, F. De Rango, C. Sottile, and C. Calafate, "A new channel assignment scheme for interference-aware routing in vehicular networks," in *Proc. IEEE VTC-Spring*, 2011, pp. 1–5.



**Hanan Saleet** (M'10) received the B.S. and M.S. degrees in industrial engineering from the University of Jordan, Amman, Jordan, and the Ph.D. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 2010.

She is currently an Assistant Professor with the Department of Mechanical and Industrial Engineering, Applied Science University, Amman. Her research interests include modeling, systems design and development, mobility and resource management in vehicular ad-hoc networks and sensor networks, performance evaluation and quality-of-service support in dependable wireless networks, network design, and optimization in wireless communications.

Dr. Saleet is a member of the IEEE Communication Society. She serves as a technical program committee member of the 2012 IEEE International Conference on Communications Ad Hoc, Mesh, and Sensor Networks Symposium.

**Rami Langar** (M'10) received the M.S. degree in network and computer science from the University of Pierre and Marie Curie, Paris, France, in 2002 and the Ph.D. degree in network and computer science from Telecom ParisTech, Paris, in 2006.

In 2007 and 2008, he was with the School of Computer Science, University of Waterloo, Waterloo, ON, Canada, as a Postdoctoral Research Fellow. He is currently an Associate Professor with the Computer Science Laboratory of Paris 6, University of Pierre and Marie Curie. His research interests include mobility and resource management in wireless mesh, vehicular ad-hoc and femtocell networks, performance evaluation, and quality-of-service support.

Dr. Langar is a member of the IEEE Communication Society. He serves as a Co-Chair of the 2012 IEEE International Conference on Communications (ICC) Ad Hoc, Mesh, and Sensor Networks Symposium; a Posters Co-Chair of the 2011 IEEE Global Information Infrastructure Symposium (GIIS); and a Tutorial Chair of the 2009 IEEE GIIS. He has also served as technical program committee member for many international conferences, including the IEEE ICC; the IEEE Global Communications Conference: the IEEE Conference on Personal, Indoor, and Mobile Radio Communications; and Vehicular Technology Conferences.

**Kshirasagar Naik** (SM'11) received the B.Sc. degree from Sambalpur University, Sambalpur, India, and the M.Tech. degree from the Indian Institute of Technology, Kharagpur, India, the M.Math. degree in computer science from the University of Waterloo, Waterloo, ON, Canada, and the Ph.D. degree in electrical and computer engineering from Concordia University, Montreal, QC, Canada.

He has been a faculty member with the University of Aizu, Fukushima, Japan, and Carleton University, Ottawa, ON, Canada. In 2003, he was a Visiting Associate Professor with the Research Institute of Electrical Communications, Tohoku University, Sendai, Japan. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Waterloo. He is currently an Associate Editor for the *Journal of Peer-to-Peer Networking and Applications* and the *International Journal of Parallel, Emergent, and Distributed Systems*. His research interests include dependable wireless communication, resource allocation in wireless, sensor networks, ad hoc networks, mobile computing, peer-to-peer communication, intelligent transportation systems, capability enhancement of smartphones and tablet computers, and communication protocols for smart power grids.

Prof. Naik served as a Program Co-Chair of the Fifth International Conference on Information Technology, held in Bhubaneswar, India, in December 2002. He was a Co-Guest Editor of two special issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS published in June 2005 and January 2007.

**Raouf Boutaba** (SM'01) received the M.Sc. and Ph.D. degrees in computer science from the University Pierre and Marie Curie, Paris, France, in 1990 and 1994, respectively.

He is currently a Professor of computer science with the David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada, and a distinguished Visiting Professor with the Division of IT Convergence Engineering, Pohang University of Science and Technology, Gyungbuk, Korea. His research interests include network, resource, and service management in wired and wireless networks.

Dr. Boutaba served as the founding Editor-in-Chief of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT (2007–2010) and is serving on the editorial boards of several journals. He has received several Best Paper Awards and other recognitions, such as the Premiers Research Excellence Award, the IEEE Hal Sobol Award in 2007, the Fred W. Ellersick Prize in 2008, the Joe LociCero Award, and the Dan Stokesbury Award in 2009.

**Amiya Nayak** (SM'04) received the B.Math. degree in computer science and combinatorics and optimization from the University of Waterloo, Waterloo, ON, Canada, in 1981 and the Ph.D. degree in systems and computer engineering from Carleton University, Ottawa, ON, in 1991.

He is currently a Full Professor with the School of Electrical Engineering and Computer Science, University of Ottawa. He has more than 17 years of industrial experience in software engineering, avionics and navigation systems, and simulation and system-level performance analysis. He is serving on the Editorial Boards of several journals, including the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS; the *International Journal of Parallel, Emergent, and Distributed Systems*; the *International Journal of Computers and Applications*; and the *EURASIP Journal of Wireless Communications and Networking*. His research interests are fault tolerance, distributed systems/algorithms, and mobile ad hoc networks, with more than 150 publications in refereed journals and conference proceedings.

**Nishith Goel** received the B.E. degree in electronics and telecommunications engineering from the University of Jodhpur, Jodhpur, India, in 1978 and the M.A.Sc. degree in electrical engineering and Ph.D. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1978 and 1983, respectively.

He joined Bell-Northern Research in 1984 and moved to Nortel Networks in 1988. He left Nortel in 1995 and founded Cistel Technology, Ottawa, ON. He is very active in research on various areas of telecom and information technology. He is currently the Chair of the Board of Directors of the Queen's Centre for Energy and Power Electronics Research. He is the Chief Executive Officer of Cistel Technology Inc., which is an information technology company that has operations in Canada and the U.S. He is a veteran technology executive and entrepreneur and a co-founder of CHiL Semiconductor, IPine Networks, Technologie SanStream, and Sparq Systems.