

Adaptive Quality of Service based Routing for Vehicular Ad hoc Networks with Ant Colony Optimization

Guangyu Li, Lila Boukhatem, and Jinsong Wu, *Senior Member, IEEE*

Abstract—Developing highly efficient routing protocols for Vehicular Ad hoc Networks (VANETs) is a challenging task mainly due to the special characters of such networks: large-scale sizes, frequent link disconnections and rapid topology changes. In this paper, we propose an adaptive quality of service (QoS) based routing for VANETs called AQRV. This new routing protocol is to adaptively choose the intersections through which data packets pass to reach the destination, and the selected route should satisfy with the QoS constraints and fulfill the best QoS in terms of three metrics, namely connectivity probability, packet delivery ratio and delay. To achieve the above objectives, we mathematically formulate the routing selection issue as a constrained optimization problem, and propose an Ant Colony Optimization (ACO) based algorithm to solve this problem. In addition, a terminal intersection concept is presented to decrease routing exploration time and alleviate network congestion. Moreover, in order to decrease network overhead, we propose Local QoS Models (LQM) to estimate real-time and complete QoS of urban road segments. Simulation results validate our derived LQM models and show the effectiveness of AQRV.

Index Terms—QoS, ACO, Routing protocol, Models, VANETs.

I. INTRODUCTION

AS an important component of smart cities and Intelligent Transportation Systems (ITS), Vehicular Ad hoc Networks (VANETs) have been attracting increasing attentions from both research and industry communities and they are considered as a special type of Mobile Ad hoc Networks (MANETs) [1]. In the environments of VANETs, each vehicle can act as a router to communicate with each other without relying on fixed infrastructure supports. There are two wireless communication ways including Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) [2] for various applications, which are classified into safety and infotainment applications [3], such as collision warning, driver assistance, Internet access and so on. Compared with other MANETs, VANETs have special characteristics [4] involving the high mobility, self-organization, road pattern restrictions, no energy constraints, large-scale network sizes and so on. Although we have achieved remarkable developments in VANETs, there are still

many challenges in the fields of designs and implementations, such as security, privacy, routing, scalability and quality of service (QoS). In this paper, we mainly focus on the network routing problems with QoS constraints.

Position-based routing protocols [5] are considered to be more stable and suitable for large-scale dynamic VANET environments, and they make use of the geographic positions of mobile nodes to relay data packets. However, these routing protocols still suffer from various limitations. VADD [6] and A-STAR [7] make routing decisions utilizing the historical or statistic route QoS, which are not accurate in dynamic VANET environments. In order to solve this problem, SADV makes use of periodic packet forwarding to estimate the real-time QoS of routing paths, but this method may result in network congestions especially in the case of high network loads. GyTAR [8] and TADS [9] employ local traffic parameters (such as vehicle densities and vehicle distributions) to select the next intersection. However, these traffic parameters can not comprehensively reflect the QoS of routing paths, for example, higher vehicle densities can improve network connectivity, but they may exacerbate end-to-end packet delivery ratio in the case of heavy data flows. On the other hand, when choosing the next intersection, GyTAR and TADS routing protocols neglect the corresponding global routing QoS, as a result, data packet transmission may suffer from extreme traffic conditions (network partitions or congestions) in the future routing selection processes. CAR [10] is a source-driven routing protocol and it explores complete end-to-end routing paths by means of the blind flooding scheme, but this method is not efficient, and usually leads to very high network overhead.

Based on the above descriptions, the main challenges we undergo are given as follows: (1) the QoS of route is not accurate, (2) the methods of QoS estimation are not efficient, (3) the routing selections are implemented by means of the incomplete or local QoS, and (4) the routing exploration algorithms are not effective and adaptive. In this paper, an Adaptive QoS-based Routing for VANETs called AQRV is proposed to deal with the aforementioned problems. We firstly formulate this route selection issue as an optimization problem, and propose an ACO-based algorithm to solve this NP-complete issue. Then, based on both global and local pheromone (reflecting the QoS of routing paths and road segments, respectively), AQRV makes use of forward and backward ants to establish the optimal route via an opportunistic method rather than the blind flooding. Once the route selection process is

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

G. Li and L. Boukhatem are with CNRS UMR 8623, Building 660, Laboratoire de Recherche en Informatique (LRI), University of Paris-Sud, 91405 Orsay, France (e-mail: liguangyu0627@hotmail.com, Lila.Boukhatem@lri.fr).

J. Wu is with University of Chile, Av Tupper 2007, 8370451 Santiago, Chile (email: wujs@ieee.org).

Manuscript received MONTH DAY, YEAR; revised MONTH DAY, YEAR.

completed, source vehicles initiate data packets transmission, which is implemented through dynamic intersection selections. Note that the comprehensive real-time local QoS (namely connectivity probability, delay and packet delivery ratio) is derived by our mathematical models. In addition, a simple greedy carry-and-forward mechanism [11] is adopted to relay packets between two adjacent intersections so as to reduce the effects of individual vehicle movements on the routing paths. Thanks to dynamic routing decisions on intersections and closed cooperations of different ants and communication pairs, AQRV is capable of adaptively coping with rapid topology changes in VANET scenarios.

The main contributions of this paper are as follows:

(1) We introduce a Terminal Intersection (TI) model for a source or destination vehicle. Compared with an end-to-end routing path for each communication pair, this model makes a group of communication pairs (sharing the same TIs and QoS requirements) directly forward data packets using the explored backbone paths, and thus this scheme is beneficial to the decrease of routing exploration time and network congestions.

(2) An ACO-based algorithm is proposed to search for the optimal QoS route between two terminal intersections. This algorithm mainly consists of the candidate route derivation process and the optimal route selection process, and it enables different communication pairs and ants closely collaborate with each other, so as to update the latest routing information and adaptively cope with rapid topology changes.

(3) An opportunistic method is proposed in the candidate route derivation process by means of both local and global pheromone. Compared with blind flooding or broadcast mechanisms, our method helps to explore new routing paths, decrease routing exploration time and reduce network overhead. Besides, employing global pheromone can improve route stability and avoid extreme routing conditions in upcoming routing selections.

(4) We derive mathematical models for local road segment QoS (namely connectivity probability, packet delivery ratio and delay). These local QoS models are used to estimate real-time local and global pheromone in AQRV. Compared with other QoS estimation methods, such as forwarding periodic update messages between adjacent intersections [12] and on-the-fly collection process [10], our models can alleviate network congestions and decrease network overhead.

The rest of this paper is organized as follows. The related works are described in Section II. AQRV routing protocol is detailed in Section III. In Section IV, we derive the local QoS models in 1-lane road segment scenarios. Section V presents and analyzes the simulation results. Finally, Section VI concludes the paper.

II. RELATED WORKS

The challenges of network routing protocols in dynamic environments have been attracting more and more research efforts [4], and a number of routing protocols have been proposed to solve related problems and support adverse applications for VANETs [8], [13], [14].

A. QoS-based routing protocols

VANETs support a number of applications for safety and comfortability, so efficient QoS-based routing protocols [15] are necessary to forward packets within the required QoS constraints. GyTAR [8] dynamically chooses the next intersection by considering curvometric distance from each candidate intersection to destination as well as local vehicle densities. Obviously, the mechanism of dynamic intersection selections helps GyTAR find more robust routing paths and cope with rapid topology changes. However, local vehicle densities and global distance are not enough to choose the best routing path, and data packets may suffer from local maximum problems in upcoming road segments. ARP-QD [16] is an QoS-based routing protocol in terms of hop count, link duration and connectivity, so as to cope with dynamic topology and keep the balance between stability and efficiency of the algorithm. However, connectivity is estimated by broadcasting beacon packets, which may cause channel congestions in the case of high traffic loads. In addition, it is not enough to only use a global distance to reflect the overall QoS of a routing path. GEOPPS [17] is a novel delay tolerant routing algorithm that exploits the available information from navigation systems to opportunistically relay data packets, but its delay model is of too high computational complexity to run in real time. MURU [18] uses the Expected Disconnection Degree (EDD) to establish the optimal path based on predicted speeds, node locations and road geometries, but only local information is applied to make routing decisions, thus it is more likely to suffer from local optimums.

B. Intersection-based routing protocols

Intersection-based routing protocols [10], [12], [8], [9], [19] are more effective and stable in VANET urban scenarios. GSR [20] makes use of a city digital map to search the shortest path towards the destination, but it neglects the traffic information along the route and frequently undergo network partition problems in sparse networks. In CAR [10], a source vehicle broadcasts request packets towards its destination to gather the available information (vehicle densities, hop counts and passing-through intersections) along the routes. Once receiving the request packets, the destination confirms the best routing path and then replies this route's information to the source. However, due to maintaining complete end-to-end routing paths, CAR is not adaptive to the routing changes. SADV [12] utilizes static nodes located at road intersections to support packet relaying, so the packets can then be buffered for a while in the static nodes until a suitable vehicle is available along the best delivery path. Note that these static nodes are not connected over a wired network, and the real-time delay between two adjacent intersections is updated by periodic messages, which may lead to network congestions. IEGRP [21] is a hybrid vehicular routing protocol that facilitates V2V or V2I unicast communication by dynamically changing its routing decisions on intersections. Note that these intersections are connected with each other via high bandwidth links. In addition, a simple greedy carry-and-forward scheme

is used to relay data packets between intermediate vehicles and intersections.

C. ACO-based routing protocols

Ant Colony Optimization (ACO) originated from the food foraging behaviors of real ants [22], [23], and it can be used to solve multi-objective problems with high computational complexities, such as traveling salesman problems [24]. ACO can also be suitable to VANET routing problems [25], [26] because of the following reasons: (1) scalability: the number of ants in ACO may vary based on network sizes, and scalability is also promoted based on the interactions among different ant colonies; (2) adaptation: ants can change, die or reproduce themselves based on network changes; (3) parallelism: ants operations are inherently parallel, and this scheme helps to speed up convergence processes; (4) fault tolerance: the losses of a few nodes or links would not result in catastrophic failures, as ACO intrinsically shows an uncentralized control mechanism. Based on above arguments, ACO is superior in terms of excellent adaptation, robustness and decentralized nature, which meet the requirements to solve dynamic routing optimization problems. HAMRV [27] employs ACO to design a multi-objective heuristic algorithm, which is used to search the optimal route based on the shortest distance and lowest probability of disconnection. ANTHOCNET [28] is a hybrid routing algorithm that combines the reactive route establishment with the proactive route maintenance process. Nevertheless, each ant in ANTHOCNET has to maintain all the addresses of visited vehicles, which expand network overhead especially for long-distance routes.

D. V2I-based routing protocols

V2I-based routing protocols [29], [30], [31] for VANETs have been extensively investigated to cope with the significant topology variations. VPGR [32] forwards data packets from source vehicles to fixed infrastructures in urban environments by predicting a sequence of valid intersections. These valid intersections are calculated by means of the geographic positions, velocities, directions and accelerations of vehicles. However, VPGR neglects the real-time traffic information (e.g., vehicle density) within road segments. Based on the movement characteristics of vehicles and the spatial relationship between vehicles and roadside units, PIACO-AODV [33] is proposed to improve the performance of AODV [34] in V2I environments, and in this routing protocol, heuristic information of ACO algorithm scheme is modified so as to obtain a higher convergence rate and increase the stability of VANETs. DUBHE [35] is a reliable and low-latency routing protocol for transmitting data from source vehicles to stationary roadside infrastructures, and it is composed of three components including a delay model, a path choosing algorithm and an improved greedy broadcast algorithm. IGRP [36] is an efficient routing protocol for V2I communication, and it is based on effective road intersection selections, which are implemented in terms of network connectivity, delay, bandwidth and error rate. However, IGRP is a source-driven routing protocol, which

requires a complete route and can not cope with rapid topology changes in VANET environments.

In order to effectively address the aforementioned problems, we propose an adaptive QoS-based routing protocol called AQRV. Firstly, we design a Terminal Intersection (TI) model, and propose an ACO-based algorithm to establish the best routing path by means of both local road segment QoS and global QoS. In addition, so as to improve routing stability, adaptively cope with topology changes and decrease network overhead, data packets are forwarded intersection by intersection, rather than via the complete route stored in the headers of packets as used in other source-driven routing protocols. Furthermore, we derive the mathematical models to estimate local real-time QoS of road segments in terms of connectivity probability, delay and packet delivery ratio.

III. AQRV ROUTING ALGORITHM

In this section, we propose AQRV routing protocol in details. AQRV is an adaptive intersection-based routing with the use of an ACO-based algorithm to search the optimal QoS route. The QoS is measured in terms of connectivity probability, packet delivery ratio and delay. In fact, each source vehicle firstly sends a routing request to its TIS (Terminal Intersection for a Source vehicle) for available routing information before forwarding data packets. If such routing information exists, a positive response is sent back to source vehicle S , which directly initiates data packets forwarding. Otherwise, TIS replies to S with a negative message, and then based on both local and global pheromone, TIS implements optimal route establishment processes between TIS and TID (Terminal Intersection for a Destination vehicle) using ACO. AQRV is mainly composed of two components: the terminal intersection selection and the optimal route establishment.

A. System model

In this paper, we assume that each vehicle is equipped with a Global Positioning System (GPS) facility, digital map and navigation system, which provide the vehicles information (including vehicle speeds, moving direction and geographical positions), the positions of intersections and the length of road segments. Besides, we suppose that source vehicles can obtain the geographic locations of their respective destinations using location services, and different communication pairs have the same QoS requirements. Moreover, a Static Node (SN_i) is installed on each intersection I_i to help packets forwarding and routing information storage.

AQRV aims at finding the optimal route in urban scenarios with the best QoS in terms of connectivity probability, packet delivery ratio and delay while satisfying the delay constraint. Here we assume that the urban street map is abstracted as graph $G(I, E)$, and for any two neighboring intersections, there is a road segment to connect them. Obviously, we can define that the optimal backbone route y consists of a sequence of intersections $\{I_1, I_2, \dots, I_{m-1}, I_m\}$, which are connected by a set of road segments $\{e_1, e_2, \dots, e_{n-1}, e_n\}$, where $n = m - 1$. Note that I_1 is the first intersection connected to source vehicle S , and I_m is the last intersection connected to destination

vehicle D . Here we set I_1 to TIS and I_m to TID. Based on above considerations, AQRV routing issue can be formulated as an optimization problem, and the objective function is given as

$$\max F(y) = \varphi_1 \cdot PC(y) + \varphi_2 \cdot PDR(y) + \varphi_3 \cdot \frac{D_{th} - D(y)}{D_{th}} \cdot \frac{1}{(1 + Dv(y))} \quad (1)$$

where

$$\begin{cases} PC(y) = \prod_{i=1}^n PC(e_i) \\ PDR(y) = \prod_{i=1}^n PDR(e_i) \\ D(y) = \sum_{i=1}^n D(e_i) \\ Dv(y) = \sum_{i=1}^n \frac{Dv(e_i)}{n} \end{cases} \quad (2)$$

subject to

$$D(y) \leq D_{th} \quad (3)$$

where $F(y)$ denotes our objective function and represents the global QoS of route y from TIS to TID. $PC(y)$, $PDR(y)$, $D(y)$ and $Dv(y)$ stand for the connectivity probability, packet delivery ratio, delay and delay variance of route y , respectively. $PC(e_i)$, $PDR(e_i)$, $D(e_i)$ and $Dv(e_i)$ represent road segment e_i 's connectivity probability, packet delivery ratio, delay and delay variance, respectively. D_{th} is the delay threshold. φ_1 , φ_2 and φ_3 are weight values, and $\varphi_1 + \varphi_2 + \varphi_3 = 1$. During the normalization process of the QoS metrics, in order to remove the dimensional characteristic effect of $D(y)$ and avoid the overwhelming influence of $Dv(y)$, we use $\frac{D_{th} - D(y)}{D_{th}} \cdot \frac{1}{(1 + Dv(y))}$ to make its value less than 1.

B. Terminal intersection selection

In this section, we present the terminal intersection (TIS or TID) selection process for the corresponding communication pairs. A terminal intersection is determined by two factors, namely the moving direction of a communication terminal (source vehicle S or destination vehicle D) and the distance from itself to its neighboring intersections. Based on these parameters, a score is allocated to each candidate intersection, and then the one owning the highest score is selected as the terminal intersection. The score expression is given as

$$Score(I_i) = \chi \cdot \left(1 - \frac{d(I_i)}{L}\right) + (1 - \chi) \cdot Dir(I_i) \quad (4)$$

where L is the current road segment length, $d(I_i)$ means the distance between a communication terminal and its neighboring intersection I_i , χ denotes the weight parameter, and $Dir(I_i)$ stands for the communication terminal's moving direction value, which is given as

$$Dir(I_i) = \begin{cases} 1, & \text{the terminal vehicle moves to } I_i \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

C. Optimal route establishment

In this section, we describe the establishment process of the optimal backbone route from TIS to TID. It is worth noting that we regard the optimal route issue as the optimization problem expressed in (1). To solve this problem, we propose an ACO-based algorithm which is skillful in addressing such combinatorial nonlinear optimization problems in complex and dynamic systems.

1) *Candidate route derivation*: Once the terminal intersections for source S and destination D are selected, respectively, S firstly sends a routing request to TIS. If the routing information (satisfying with QoS requirements) towards TID exists at TIS and it is not out of date, TIS replies a positive message to S ; otherwise TIS sends a negative response to S and then starts to search the optimal backbone route.

To derive the candidate routes from TIS to TID, TIS launches a group of forward ants towards TID, here we set the number of forward ants to N_{fant} . When a forward ant arrives at intersection I_i , it firstly records I_i 's ID and its transmission delay, then it makes a stochastic decision to choose the next intersection based on both global and local pheromone stored locally at I_i . Suppose that there are K neighboring intersections of I_i , namely $I_1, I_2, \dots, I_{K-1}, I_K$. The probability p_{ij} with which this forward ant selects I_j as the next intersection is given as

$$p_{ij} = \frac{LP_{ij}^\alpha \cdot GP_{ij}^\beta}{\sum_{m=1}^K LP_{im}^\alpha \cdot GP_{im}^\beta} \quad (6)$$

where α and β denote the weight values. LP_{ij} depicts the Local Pheromone, which reflects the local QoS of road segment e_{ij} from I_i to I_j . GP_{ij} means the Global Pheromone, which implies the global QoS of route y_{ij} from I_i to TID when going through I_j . The expressions of LP_{ij} and GP_{ij} are derived as

$$LP_{ij} = F(e_{ij}) = \varphi_1 \cdot PC(e_{ij}) + \varphi_2 \cdot PDR(e_{ij}) + \varphi_3 \cdot \frac{D_{th} - D(e_{ij})}{D_{th}} \cdot \frac{1}{(1 + Dv(e_{ij}))} \quad (7)$$

$$GP_{ij} = F(y_{ij}) = \varphi_1 \cdot PC(y_{ij}) + \varphi_2 \cdot PDR(y_{ij}) + \varphi_3 \cdot \frac{D_{th} - D(y_{ij})}{D_{th}} \cdot \frac{1}{(1 + Dv(y_{ij}))} \quad (8)$$

where $PC(e_{ij})$, $PDR(e_{ij})$, $D(e_{ij})$ and $Dv(e_{ij})$ denote the connectivity probability, packet delivery ratio, delay and delay variance of road segment e_{ij} , respectively, and these three metrics are deduced in section IV. $PC(y_{ij})$, $PDR(y_{ij})$, $D(y_{ij})$ and $Dv(y_{ij})$ (derived in (2)) represent the same QoS metrics for route y_{ij} . Obviously, LP_{ij} is the heuristic pheromone which can help forward ants get the latest local QoS information and explore new routing paths. GP_{ij} represents the global route QoS, and it makes the forward ants be inclined to select the established optimal route, which is beneficial to the algorithm convergence as well as the avoidance of routing problems (routing holes and/or highly loaded links) on upcoming road segments. Therefore, the probability p_{ij} is used to keep the balance between new route explorations and prior route exploitations by adjusting the values of α and β .

Algorithm 1 AQRV routing algorithm concept

```

1:  $T_{GP}$ : update time interval of global pheromone.
2: Identify TIS and TID using (4).
3:  $S$  sends a routing request message to TIS for the global pheromone towards TID.
4: if global pheromone exists  $\&\&$   $CurrentTime - LastUpdateTime \leq T_{GP}$   $\&\&$   $D(y) \leq D_{th}$  then
5:   A positive response is sent back to  $S$ .
6: else
7:   TIS implements Optimal route establishment (refer to Algorithm 2).
8: end if
9:  $S$  initiates data packet forwarding intersection by intersection.
10: if TIS || TID changes then
11:   TIS implements Optimal route establishment.
12: end if

```

When a forward ant arrives at TID, and if its delay is less than D_{th} , the path passed through by this ant is selected as an available candidate route. An outage situation may occur if all the arriving forward ants do not fulfill the delay requirement, which means that there is no available routing path at the moment, most probably due to a heavy loaded network or an excessively high delay requirement. The problem can be resolved by restarting the routing exploration processes when the network condition improves or by selecting an appropriate delay threshold value.

2) *Optimal route selection*: We make use of backward ants to implement the optimal route selection among the explored candidate routes in terms of updated global pheromone. When a forward ant arrives at TID, and if its transmission delay fulfills the required delay threshold D_{th} , this forward ant is converted into a backward ant, otherwise, it is dropped directly. This backward ant carries an intersection list (give by the respective forward ant), and then it is sent back to TIS using the same but reverse route of the corresponding forward ant. When arriving at a given I_i (moving from I_j), the backward ant firstly calculates the Latest Global Pheromone LGP_{ij} (derived in (8)) by means of carried global QoS metrics of the passing through route (via (2) in terms of local road segment QoS), and then the pheromone GP_{ij} stored at I_i 's table is updated as follows

$$GP_{ij} \leftarrow (1 - \delta) \cdot GP_{ij} + \delta \cdot LGP_{ij} \quad (9)$$

where δ is the weight parameter ($0 < \delta < 1$). Obviously, this pheromone update process can avoid the exploration stagnancy of routing paths and alleviate the effect of LGP_{ij} 's instant value. After that, GP_{ij} is renewed in the field of updated global pheromone of this backward ant.

Once all backward ants arrive at TIS, we compare $F(y)$'s values of all the available candidate routes, and the route with the maximum value of $F(y)$ is selected as the optimal one. Then, TIS sends a positive routing message to source vehicle S to start data packet forwarding.

Note that a pheromone evaporation process is necessary to avoid the rapid algorithm convergence towards a suboptimal

Algorithm 2 Optimal route establishment

```

1: for each ant packet  $A$  arriving at  $I_i$  do
2:   if  $A \rightarrow Type == ForwardAnt$  then
3:     Record  $I_i$ 's ID in  $A$ 's header.
4:     NextIntersection selection using (6).
5:     Send  $A$  to NextIntersection.
6:     if NextIntersection == TID then
7:       if  $D(y) \leq D_{th}$  then
8:          $A \rightarrow Type = BackwardAnt$ .
9:       else
10:        Drop  $A$ .
11:       end if
12:     end if
13:   end if
14:   if  $A \rightarrow Type == BackwardAnt$  then
15:     UpdatePheromone process using (9).
16:     Send  $A$  to NextIntersection.
17:     if NextIntersection == TIS then
18:       UpdatePheromone process using (9).
19:     Drop  $A$ .
20:   end if
21: end if
22: end for
23: PheromoneEvaporation using (10).
24: Output the optimal route according to (1).
25: TIS sends a positive routing message to  $S$ .

```

region. In AQRV, we define the evaporation process as follows

$$GP_{ij}(t + T_{eva}) = \begin{cases} \eta \cdot GP_{ij}(t) & \text{If } GP_{ij}(t) > \tau_{min} \\ \tau_{min} & \text{Otherwise} \end{cases} \quad (10)$$

where η denotes the pheromone evaporation factor and $0 < \eta < 1$, T_{eva} means the evaporation time interval and τ_{min} stands for the lowest threshold of global pheromone.

3) *Data packet forwarding and route maintenance*: In AQRV, we define two forwarding processes for data packets. When forwarded between two neighboring intersections, data packets adopt the greedy carry-and-forward mechanism. When arriving at intersection I_i , data packets are dynamically relayed to next intersection I_j which shows the maximum global pheromone $GP_{ij} = \max_m^K \{GP_{im}\}$, where K denotes the number of I_i 's neighboring intersections.

To cope with the topology changes and maintain routing paths, when the terminal intersection of source or destination vehicle changes or deposited global pheromone is out of date, AQRV implements a new optimal route establishment process. Algorithm 1 and Algorithm 2 summarize the different steps of AQRV.

IV. LOCAL QOS MODELS (LQM) OF A ROAD SEGMENT

In this section, we propose the local QoS models for 1-lane road segment scenarios, and these models play an important role in routing selections for urban environments. We choose connectivity, packet delivery ratio and delay to evaluate the QoS of road segments, as these metrics are closely related to the communication parameters such as wireless channel fading

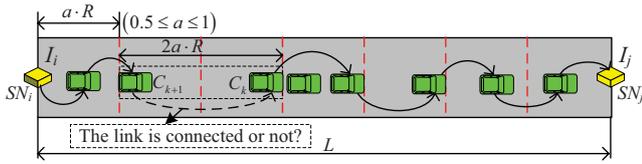


Fig. 1. Vehicle distribution analysis on a 1-lane road segment.

effects, communication ranges, road segment length, vehicle densities, distributions and so on, all of which reflect the almost complete traffic conditions of road segments. Note that these three metrics may be antagonistic in different VANET scenarios. For example, the ascending vehicle density is advantageous to the link connectivity improvement and delay reduction, but may aggravate end-to-end packet delivery ratio due to more influences from channel fading and interferences. To this end, we propose mathematical models for connectivity probability, packet delivery ratio and delay, respectively, so as to evaluate the comprehensive and real-time local QoS of 1-lane road segments.

A. Assumptions

In this section, we present the assumptions to design our models. Firstly, the number of vehicles on a roadway follows Poisson Distribution, and the vehicle spacing density is set to λ . This assumption has been proved in [36], [37]. Secondly, the vehicle velocity is assumed constant and referred to as v , and the length of vehicles is negligible compared with the length of road segment. Thirdly, a simple greedy carry-and-forward mechanism is used to relay packets within road segments, and when packets are carried by a moving vehicle, they do not suffer from packet losses. Finally, we employ the shadowing model [38] as the propagation path loss model, assume the rayleigh fading channel [39] as the wireless channel, adopt RTS/CTS (Request To Send/Clear To Send) mechanism, and utilize simple BPSK modulation technique.

B. Connectivity probability derivation in a road segment

In a 1-lane scenario, we define that a link between two consecutive vehicles is connected as long as this link length is within the communication range. In this section, the multi-hop link connectivity probability is derived for a given 1-lane road segment. Due to the variational number of vehicles, the vehicle topologies on a road segment are timely changed and the exact connectivity probability is hardly derived. In our model, the given road segment is divided into several cells to capture the connectivity characteristics. We define the cell size as $cs = a \cdot R$ and consider two threshold values of the cell size $0.5R$ and R , as shown in Fig. 1, where R denotes the communication range, a is the weight value, L stands for the length of this road segment, I_i and I_j represent two intersections, and SN indicates a static node located on each intersection to assist data packet forwarding. In the case of $cs = 0.5R$, the maximum distance between any two vehicles in two adjacent cells equals R , which guarantees the link connectivity between them. In the case of $cs = R$, the location

distributions of vehicles are random and the distance between vehicles separately located in two consecutive cells may vary from 0 to $2R$, so the link (such as the link between two vehicles C_k and C_{k+1} shown in Fig. 1) may be disconnected. Obviously, the case of $cs = 0.5R$ is a sufficient but not always necessary condition for the connectivity of a link between any two vehicles, which are respectively located in two adjacent cells. On the contrary, $cs = R$ is a necessary but insufficient condition for link connectivity. According to aforementioned interpretations, we set $cs = a \cdot R$ ($0.5 < a < 1$) to confirm the necessary and sufficient condition for link connectivity, and let K be the random variable denoting the number of vehicles in the interval $a \cdot R$. Based on our assumptions, K follows Poisson Distribution and its Probability Mass Function (PMF) is given as

$$P(K = k) = \frac{(aR\lambda)^k}{k!} \cdot e^{-aR\lambda} \quad (11)$$

Obviously, in the case of $cs = a \cdot R$, the probability PC_{1-cell} that one cell contains at least one vehicle is expressed as

$$PC_{1-cell} = 1 - P(K = 0) = 1 - e^{-aR\lambda} \quad (12)$$

As shown in Fig. 1, the multi-hop link between I_i and I_j is connected as long as each link distance between any two consecutive vehicles does not exceed R . This implies that there is at least one vehicle in each cell of the road segment, so the link connectivity probability of the road segment between I_i and I_j is derived as

$$PC = (PC_{1-cell})^N = (1 - e^{-aR\lambda})^N \quad (13)$$

where $N = \lceil \frac{L}{a \cdot R} \rceil$ denotes the number of cells on the road segment between I_i and I_j . From (13), we see that the connectivity probability of a road segment is a function of the vehicle spacing density, communication range, cell size and road segment length.

C. Packet delivery ratio derivation in a road segment

As a key metric to measure the reliability of multi-hop wireless networks, Packet Delivery Ratio (PDR) is deduced for a given 1-lane road segment in this section. Note that, due to channel fading influences and propagation characteristics of the shadowing model, a packet can not be always transmitted successfully even if the distance between the transmitting and receiving vehicles is within communication range R . Obviously, the successful packet delivery is totally different from the aforementioned link connectivity probability in our models. In addition, the statistical and real-time traffic information have demonstrated that vehicles tend to move in clusters on the road [40]. Consequently, based on the scale sizes of vehicles clusters, an analytical PDR model can be deduced via taking into account three different cases illustrated in Fig. 2: the Fully Connected road segment case (FC case, where the Communication Cluster Scale $ccs \geq L$), Fully Broken road segment case (FB case, $ccs = 0$) and Partly Connected road segment case (PC case, $0 < ccs < L$). We define forwarding link length L_f as the distance through

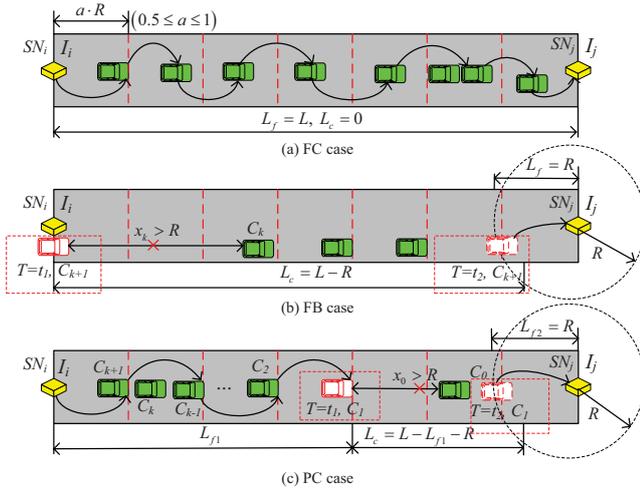


Fig. 2. Analysis of different cases in a road segment scenario.

which packets travel when forwarded via wireless multi-hop communications within the road segment, while carry link length L_c is the distance through which packets pass when carried by a vehicle at speed v .

1) *1-hop packet delivery ratio*: In order to propose the multi-hop packet delivery ratio model of a road segment, we first need to model the packet delivery ratio of a single hop. In this paper, we assume the use of RTS/CTS mechanism in the wireless channel, so the interferences introduced by hidden terminals and concurrent transmissions can be reduced. However, there are still some potential interferences from the neighboring nodes located in the segment that is outside the communication ranges of both the transmitter and receiver but inside the interference range of the receiver, as shown in Fig. 3. Therefore, we consider the components of the channel fading and potential interferences to deduce our PDR model.

Based on the derivation in [38], the overall shadowing model is given as

$$\left[\frac{P_r(x)}{P_r(x_0)} \right]_{dB} = -10 \cdot \beta \cdot \log \left(\frac{x}{x_0} \right) + X_{dB} \quad (14)$$

where $P_r(x)$ delegates the received signal power at the 1-hop distance x from the transmitter, X_{dB} is a random variable following Gaussian Distribution (GD) with zero mean and standard deviation σ_{dB} ($\sigma_{dB} = 4 \sim 12$), β denotes the path loss exponent ($\beta = 2.7 \sim 5$ for shadowed urban area), and $P_r(x_0)$ is a reference value of received signal power given by

$$P_r(x_0) = \frac{P_t}{x_0^2} \cdot \frac{G_t \cdot G_r}{L_s} \cdot \left(\frac{\lambda_c}{4\pi} \right)^2 \quad (15)$$

where P_t is the transmitted signal power, G_t and G_r denote the antenna gains of the transmitter and the receiver, respectively, L_s represents the system loss and λ_c means the carrier wave length.

In addition, in terms of the coherent BPSK modulation and Rayleigh Fading channel, a link BER (Bit Error Rate) can be

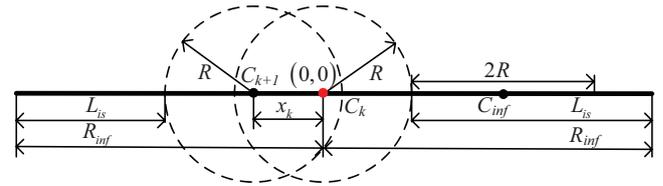


Fig. 3. Analysis of potential interfering nodes. Vehicle C_{k+1} is sending packets to C_k (located in the origin of coordinates $(0,0)$). The distance between C_k and C_{k+1} is x_k , R_{inf} denotes the interference range, R means the communication range, L_{is} represents the length of interfering segment and C_{inf} is an interfering node located in the interfering segments.

expressed as [39]

$$BER = \frac{1}{2} \left(1 - \sqrt{\frac{\gamma_0}{\gamma_0 + 1}} \right) \quad (16)$$

where γ_0 denotes the mean ratio value of the received signal to the interferences plus noise, and it is defined as follows

$$\gamma_0 = 2\sigma^2 \cdot \frac{P_r(x)}{P_{therm} + \sum_{i=1}^{N_{inf}} P_{inf}^i} \quad (17)$$

where $2\sigma^2$ is the mean value of the random variable in Rayleigh Distribution, $P_{therm} = F_k \cdot CTR$ denotes the thermal noise power, F_k is a constant parameter, CTR stands for the channel transmission data rate, P_{inf}^i is the interference from the neighboring node C_i and N_{inf} represents the number of interfering nodes.

As shown in Fig. 3, from the perspective of the receiver C_k , the potential interfering nodes are located in two segments: (R, R_{inf}) and $(-R_{inf}, -R - x_k)$ due to RTC/CTS mechanism, and there is maximum one interfering node in each $2R$ within these segments, so we have: $\max N_{inf} = \left\lceil \frac{R_{inf} - R}{2R} \right\rceil + \left\lceil \frac{R_{inf} - R - x_k}{2R} \right\rceil$.

By substituting (17) into (16), we can derive

$$BER(x) = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{2\sigma^2 P_r(x)}{2\sigma^2 P_r(x) + P_{therm} + \sum_{i=1}^{N_{inf}} P_{inf}^i}} \quad (18)$$

So the 1-hop packet delivery ratio is given as

$$PDR_{1hop}(x) = (1 - (1 - cr) BER(x))^{psize} \quad (19)$$

where $psize$ denotes the packet size, cr stands for the error correction ratio and x depicts the 1-hop distance.

2) *Packet delivery ratio in the FC case*: In this case, the links on the road segment are fully connected, and packets are forwarded from intersection I_i to I_j hop by hop. The hop counts are closely relevant to the connectivity probability of road segments, so we set cell size $cs = a \cdot R$ ($0.5 < a < 1$) to investigate 1-lane road segment scenarios based on the discussions in section IV-B. As shown in Fig. 2(a), forwarding link length $L_f = L$, and carry link length $L_c = 0$, then the hop counts between I_i and I_j in this case is expressed as

$$H_{FC} = \left\lceil \frac{L_f}{a \cdot R} \right\rceil + 1 = \left\lceil \frac{L}{a \cdot R} \right\rceil + 1 \quad (20)$$

where L is the road segment length, R denotes the communication range and a is the cell size parameter.

Therefore, the packet delivery ratio of the 1-lane road segment in the FC case can be given as follows

$$\begin{aligned} PDR_{FC} &= PDR_{1hop}(x)^{H_{FC}} \cdot PDR_C \cdot P_{FC} \\ &= PDR_{1hop}(x)^{H_{FC}} \cdot P_{FC} \end{aligned} \quad (21)$$

where PDR_C is the packet delivery ratio when packets are carried by a forwarding vehicle ($PDR_C = 1$); P_{FC} represents the probability of the FC case and $P_{FC} = PC = (1 - e^{-aR\lambda})^N$ (deduced from (13)); the average 1-hop distance $x = L/H_{FC}$.

3) *Packet delivery ratio in the FB case:* In this case, the road segment is fully disconnected and no vehicle is available to forward packets, so packets from intersection I_i are carried by the forwarding vehicle at vehicle speed v until this vehicle enters the communication range of intersection I_j . Here, we define the link length between two successive vehicles C_k and C_{k+1} as x_k , as shown in Fig. 2(b). Since we assume that vehicle speed v remains constant on a road segment, as long as the link length between two latest arrival vehicles C_{k+1} and C_k satisfies with $x_k > R$, packets hold by C_{k+1} will be carried at v until arriving in the communication range of I_j . Based on the above analyses, $L_c = L - R$, $L_f = R$ and the packet delivery ratio in the FB case is given as

$$PDR_{FB} = PDR_{1hop}(R) \cdot P_{FB} \quad (22)$$

where P_{FB} means the probability of the FB case.

On the basis of the previously mentioned assumptions, the number of vehicles on a road segment follows Poisson Distribution, so the distance X between any two consecutive vehicles obeys exponential distribution [41]. More precisely, the Cumulative Distribution Function (CDF) of X could be expressed as follows:

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x} \quad (23)$$

where λ denotes the vehicle spacing density of the road segment.

Consequently, the probability of the FB case is expressed as

$$P_{FB} = P(X > R) = 1 - F(R) = e^{-\lambda R} \quad (24)$$

Finally, by substituting (24) into (22), the packet delivery ratio in the FB case is derived as

$$PDR_{FB} = PDR_{1hop}(R) \cdot e^{-\lambda R} \quad (25)$$

4) *Packet delivery ratio in the PC case:* In this case, the road segment is partly connected, for example, as shown in Fig. 2(c), the network links from I_i to vehicle C_1 are fully connected, but the link between C_1 and C_0 is broken because of the far distance between them $x_0 > R$. As a result, packets are forwarded hop by hop from I_i to C_1 , and then carried by C_1 until C_1 enters the communication range of I_j . Here, forwarding link length $L_f = L_{f1} + L_{f2}$, L_{f1} is equal to the distance between I_i and C_1 , and $L_{f2} = R$.

The packet delivery ratio in the PC case is illustrated as follows:

$$PDR_{PC} = PDR_{1hop}(x)^{H_{PC1}} \cdot PDR_{1hop}(R) \cdot P_{PC} \quad (26)$$

where $H_{PC1} = \left\lceil \frac{L_{f1}}{aR} \right\rceil$ denotes the hop counts within L_{f1} , $x = L_{f1}/H_{PC1}$ is the average 1-hop distance and P_{PC} is the PC case probability, which is derived as

$$P_{PC} = 1 - P_{FC} - P_{FB} = 1 - (1 - e^{-aR\lambda})^N - e^{-\lambda R} \quad (27)$$

According to (26), once L_{f1} is derived, the packet delivery ratio in the PC case can be deduced.

We define $Cd(m)$ as a constraint condition for the connected cluster, and it means that the link length of any two successive vehicles in this cluster satisfies with $x_k \leq R$ for $k = 1, \dots, m$, and $x_0 > R$. Obviously, we can express L_{f1} as the sum of each link length in the connected cluster, and $L_{f1}(m) = \sum_{k=1}^m x_k$, where m denotes the number of connected links. Based on above illustrations, $E(L_{f1})$ is deduced in (28), where $E[x|x \leq R]$ is the average link length with the constraint $x \leq R$, and M is the largest integer that makes $L_{f1}(m) \leq L$.

Since link length X follows the exponential distribution, the CDF of X with constraint $X \leq R$ is expressed as

$$\begin{aligned} G(X) &= P(X \leq x | X \leq R) = \frac{P(X \leq x, X \leq R)}{P(X \leq R)} \\ &= \begin{cases} \frac{1 - e^{-\lambda x}}{1 - e^{-\lambda R}}, & 0 \leq x \leq R \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (29)$$

As a result, the expression of $E[x|x \leq R]$ can be derived as

$$\begin{aligned} E(x|x \leq R) &= \int_0^R x G'(X) dx = \int_0^R x \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda R}} dx \\ &= \frac{1 - (R\lambda + 1)e^{-\lambda R}}{\lambda(1 - e^{-\lambda R})} \end{aligned} \quad (30)$$

Based on (30), the largest integer M is derived as

$$\begin{aligned} M \cdot E(x|x \leq R) \leq L &\Rightarrow M = \left\lfloor \frac{L}{E(x|x \leq R)} \right\rfloor \\ &= \left\lfloor \frac{L \cdot \lambda (1 - e^{-\lambda R})}{1 - (R\lambda + 1)e^{-\lambda R}} \right\rfloor \end{aligned} \quad (31)$$

Using (30) and (31), the value of $E(L_{f1})$ in (28) can be derived.

Finally, the average packet delivery ratio between I_i and I_j is given as

$$E(PDR) = PDR_{FC} + PDR_{FB} + PDR_{PC} \quad (32)$$

where PDR_{FC} , PDR_{FB} and PDR_{PC} are derived in (21), (25) and (26), respectively.

$$\begin{aligned}
 E(L_{f1}) &= \sum_{m=1}^{\infty} E[L_{f1}(m) | Cd(m)] \cdot P[Cd(m)] \\
 &= \sum_{m=1}^{\infty} E \left[\sum_{k=1}^m x_k | x_k \leq R \text{ for } i = 1, \dots, m, \text{ and } x_0 > R \right] \cdot P[x_k \leq R \text{ for } i = 1, \dots, m, \text{ and } x_0 > R] \\
 &= \sum_{m=1}^{\infty} m \cdot E[x|x \leq R] \cdot P[x_0 > R] \cdot P(x \leq R)^m = \sum_{m=1}^{\infty} m \cdot E[x|x \leq R] \cdot e^{-\lambda R} \cdot (1 - e^{-\lambda R})^m
 \end{aligned}$$

As the road segment is finite and its length equals L , the length of connected cluster should satisfy with $L_{f1}(m) \leq L$. Consequently, $E(L_{f1})$ is rewritten as

$$\begin{aligned}
 E(L_{f1}) &= \left(\sum_{m=1}^M m \cdot E[x|x \leq R] + \sum_{m=M+1}^{\infty} L \right) \cdot e^{-\lambda R} \cdot (1 - e^{-\lambda R})^m = \\
 & \left(M \cdot (1 - e^{-\lambda R})^{M+2} - (M+1) \cdot (1 - e^{-\lambda R})^{M+1} - e^{-\lambda R} + 1 \right) \cdot e^{\lambda R} \cdot E[x|x \leq R] + L \cdot (1 - e^{-\lambda R})^{M+1}
 \end{aligned} \tag{28}$$

D. Delay derivation in a road segment

As a greedy carry-and-forward algorithm is used to relay packets between two adjacent intersections, the packets delay on a road segment is closely correlated to several parameters, namely network connectivity, 1-hop delay, hop count and vehicle speeds. In this section, we propose a delay model by considering three cases identified in section IV-C: the FC case, FB case and PC case.

1) *Delay in the FC case:* In this case, as shown in Fig. 2(a), the road segment between I_i and I_j is fully connected, and packets are forwarded hop by hop, so the delay in the FC case is simply given as

$$D_{FC} = H_{FC} \cdot t_p \cdot P_{FC} \tag{33}$$

where H_{FC} and P_{FC} denote the average hop counts and the probability of the FC case, respectively, and these two parameters are represented in section IV-C2, t_p [42] is the 1-hop delay, which depends on several parameters, such as the packet sizes, channel transmission rates, contention window sizes, traffic loads and so on.

2) *Delay in the FB case:* In the FB case, since network connectivity is totally broken, packets are carried by a forwarding vehicle at speed v until this vehicle enters into the communication range of I_j , the traffic scenario is shown in Fig. 2(b). Obviously, the delay in the FB case consists of two parts including the time consumed in carry link length $L_c = L - R$ and the last hop. So the delay in the FB case is expressed as

$$D_{FB} = \left(\frac{L_c}{v} + t_p \right) \cdot P_{FB} = \left(\frac{L - R}{v} + t_p \right) \cdot P_{FB} \tag{34}$$

where P_{FB} denotes the FB case's probability, which can be derived from (24).

3) *Delay in the PC case:* In the PC case, as shown in Fig. 2(c), packets are relayed from I_i to C_1 via multi-hop wireless links, and then these packets are carried by C_1 until C_1 comes into I_j 's communication range. Forwarding link length L_f in the PC case is composed of two parts:

$L_{f2} = R$ and L_{f1} (given in (28)). Therefore, carry link length $L_c = L - L_{f1} - L_{f2} = L - R - L_{f1}$, and the delay in the PC case is derived as

$$\begin{aligned}
 D_{PC} &= \left(\frac{L_c}{v} + (H_{PC1} + 1) \cdot t_p \right) \cdot P_{PC} \\
 &= \left(\frac{L - R - L_{f1}}{v} + \left(\left\lceil \frac{L_{f1}}{a \cdot R} \right\rceil + 1 \right) \cdot t_p \right) \cdot P_{PC}
 \end{aligned} \tag{35}$$

where P_{PC} denotes the probability of the PC case, which is deduced in (27).

From the above analysis, the delay from I_i to I_j is given as

$$E(D) = D_{FC} + D_{FB} + D_{PC} \tag{36}$$

where D_{FC} , D_{FB} and D_{PC} are deduced from (33), (34) and (35), respectively.

Finally, the delay variance between I_i and I_j can be deduced as:

$$\begin{aligned}
 Dv &= E(D^2) - E(D)^2 \\
 &= \frac{D_{FC}^2}{P_{FC}} + \frac{D_{FB}^2}{P_{FB}} + \frac{D_{PC}^2}{P_{PC}} - (D_{FC} + D_{FB} + D_{PC})^2
 \end{aligned} \tag{37}$$

V. PERFORMANCE EVALUATION AND ANALYSIS

In this section, we firstly verify the proposed local road segment QoS models, and then analyze AQRV's performance compared with different versions of AQRV, two geographic routing protocols (GSR [20] and CAR [10]) and two intersection-assisted routing protocols (SADV [12] and IEGRP [21] (in IEGRP, we assume that 30% intersections of the urban environment are connected via high bandwidth links)). In order to better explain the advantages of our analytical QoS and TI models, we define AQRV_NLQM as a version of AQRV in which the local road segment QoS is obtained by periodic beacons rather than the proposed LQM models, and specify AQRV_NTI as the modified AQRV in which the routing paths are explored between two end-to-end communication vehicles rather than from TIS to TID.

A. Experimental environment

In our simulation experiments, we make use of Vehicular Ad Hoc Networks Mobility Simulator (VanetMobiSim [43]) to generate vehicle mobility. The simulation area is set to 5000 meters (m) \times 5000 m consisting of 57 intersections and 91 1-lane road segments, where the vehicles move at a speed selected between 10 ~ 20 meters per second (m/s). The vehicle spacing density and vehicle safety time are defined as 0.01 ~ 0.04 vehicles per meter and 2 seconds (s), respectively. We use Intelligent Driver Model with Intersection Management (IDM_IM [43]) as the mobility model, where we set that vehicles arriving at intersections follow the rule of "first arrive first go through and turn right". In addition, the initial positions and moving trips of vehicles are randomly selected, and we take advantage of NS-2 [44] as the network simulator and generate stochastically 250 Constant Bit Rate (CBR) sessions. In order to eliminate the effects of random seeds and obtain satisfactory Confidence Intervals (CI), we repeat each simulation 50 times. The remaining simulation parameters are indicated in Table I.

TABLE I
SIMULATION PARAMETERS

Parameters	Value
MAC protocol	IEEE 802.11p
CBR packet rate	2 packets/s
Communication range R	150 ~ 350 m
Interference range R_{inf}	250 ~ 450 m
Data packet size $psize$	128 and 256 Bytes
Forward Ant number N_{fant}	50
T_{GP}	10 s
Delay upper threshold T_{th}	80 s
Cell size parameter a	0.75
Pheromone weight parameter δ	0.3
QoS weight parameters φ_1, φ_2 and φ_3	0.2, 0.2, 0.6
Weight value in TI selection χ	0.5
Weight parameter of local pheromone α	8
Weight parameter of global pheromone β	5
Pheromone threshold τ_{min}	0.3
Pheromone evaporation factor η	0.95
Evaporation time interval T_{eva}	1 s

B. Validation and analysis of local QoS models

1) *Connectivity probability of road segment*: Fig. 4 and Fig. 5 illustrate the impacts of road segment length L , cell size cs , vehicle spacing density λ and communication range R on the road segment connectivity probability. We set $\lambda = 0.015$ vehicles/m in Fig. 4, which displays the upper and lower thresholds of road segment connectivity when $cs = R$ and $cs = 0.5R$, respectively. When $cs = 0.75R$, the average analytical results are closely consistent with the simulation results ($CI = 95\%$). In Fig. 5, we set $L = 1500$ m and $cs = 0.75R$, and from this figure, we can see that the average analytical results coincide with the simulation results for all R and λ . In addition, when the communication range and vehicle spacing density are increasing, the road segment network partitions can be repaired, which improve the connectivity. These two figures validate the accuracy of our proposed connectivity probability model.

2) *Packet delivery ratio of road segment*: In the scenario of Fig. 6, we set $\lambda = 0.015$ vehicles/m and this figure shows that the average error between the theoretical results and simulation results is only 4.85% when $cs = 0.75R$. In addition, when the road segment length rises, the packet delivery ratio decreases due to more communication hops. Moreover, the packet delivery ratio increases with the decline of interference range because of the less interference effects from neighboring nodes. Fig. 7 validates the availability of the packet delivery ratio model with varying vehicle spacing density, channel transmission data rate (CTR) and packet size. As expected, the packet delivery ratio declines with the increase of CTR, which leads to lower SINR and higher BER. Besides, the packet delivery ratio becomes worse with the increase of vehicle spacing density. The reason is that higher vehicle spacing density improves network connectivity and then makes data packets be forwarded mainly by wireless communications, which result in more interferences from the wireless channel fading.

3) *Delay of road segment*: Fig. 8 and Fig. 9 describe the correlation among the delay, vehicle spacing density, channel transmission data rate and cell size. When $cs = 0.75R$, the analytical and simulation results of these two figures are in good agreement, which demonstrates the accuracy of our road segment delay model. Besides, Fig. 8 shows that the road segment delay decreases noticeably, since the network partitions can be repaired with the increase of vehicle spacing density. In Fig. 9, we set $\lambda = 0.03$ vehicles/m. This figure shows that the road segment delay decreases when the channel transmission data rate rises, which makes data packets be transmitted rapidly without the requirement to wait for a long time.

According to above investigations, when $cs = 0.75R$, all average errors between analytical and simulation results are less than 5%, which verifies the correctness of our proposed analytical QoS models for road segments. In addition, for example, when vehicle spacing density increases, connectivity probability is enhanced, delay reduces but packet delivery ratio degrades. These results justify the need for combining the three QoS metrics to meet the complete QoS routing performance.

C. Routing protocol performance analysis

1) *End-to-end packet delivery ratio analysis*: Fig. 10 shows the average packet delivery ratio for different vehicle spacing densities. We can see that all protocols show higher packet delivery ratio for lower vehicle spacing density. The reason is that, when vehicle spacing density increases, data packets prefer to be transmitted through wireless communications, which make them suffer from more channel fading effects than those carried by moving vehicles. In addition, we notice that AQRV achieves the highest packet delivery ratio compared with other routing protocols. There are two main reasons to explain such results. First of all, AQRV dynamically selects the optimal routing path with the highest communication QoS including packet delivery ratio. Secondly, AQRV searches available routing paths based on ACO, which makes different communication pairs and ants cooperate with each other

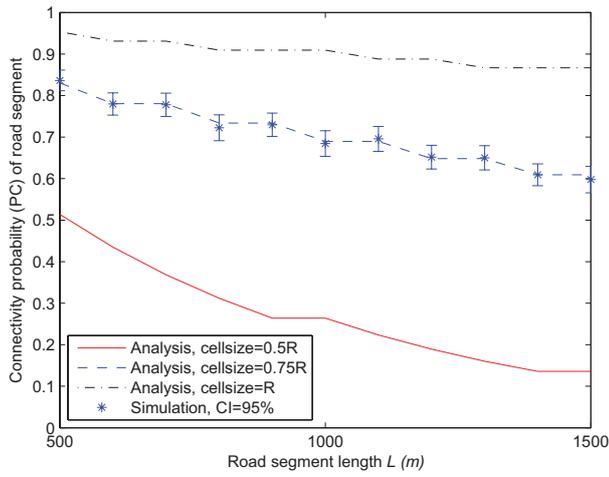


Fig. 4. Road segment connectivity.

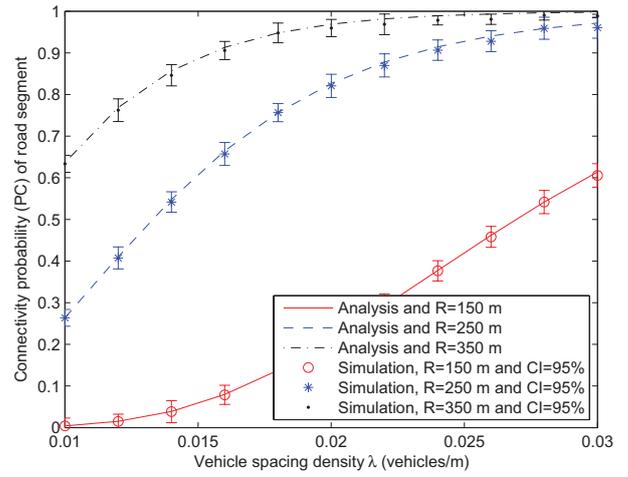


Fig. 5. Road segment connectivity.

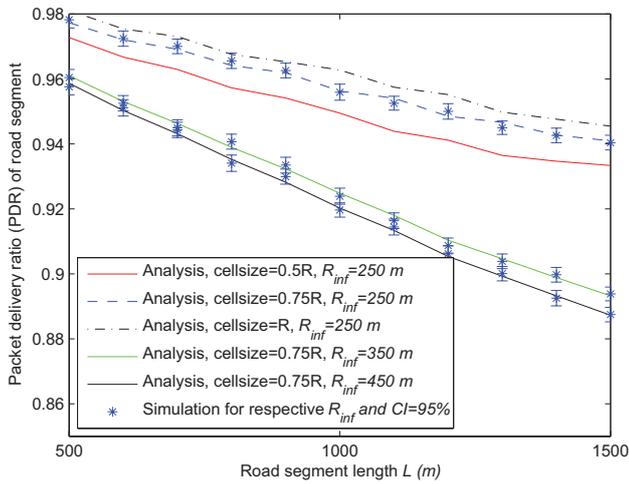


Fig. 6. Road segment packet delivery ratio.

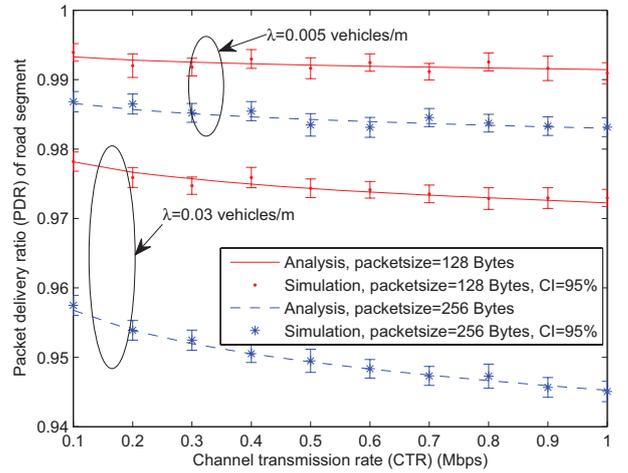


Fig. 7. Road segment packet delivery ratio.

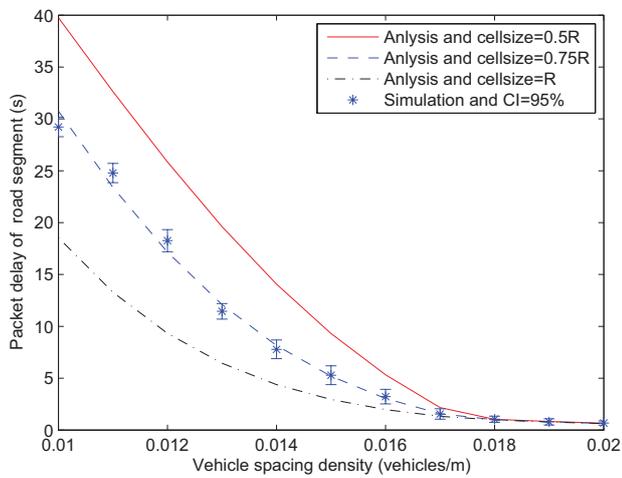


Fig. 8. Road segment delay.

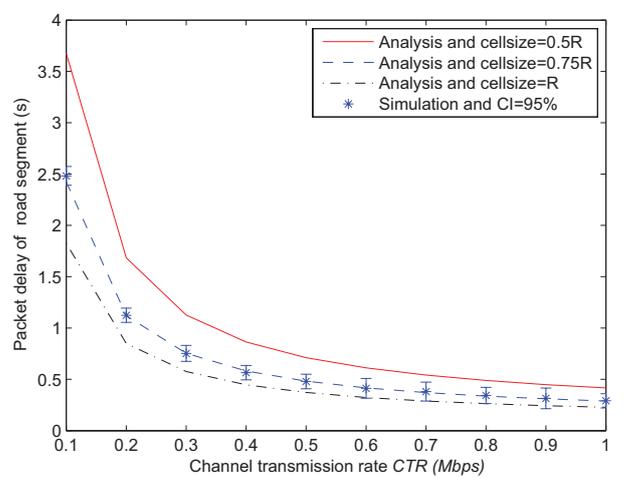


Fig. 9. Road segment delay.

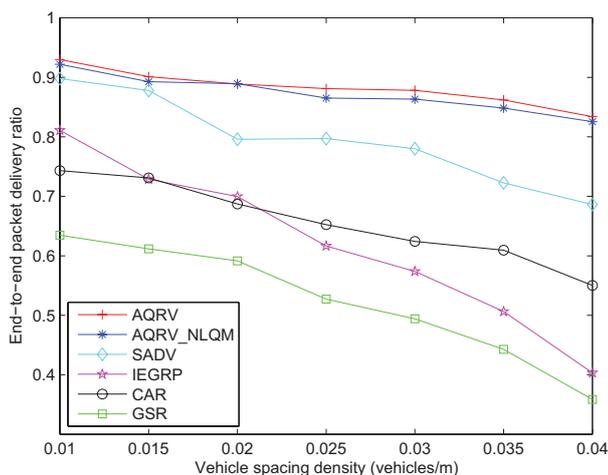


Fig. 10. End-to-end packet delivery ratio.

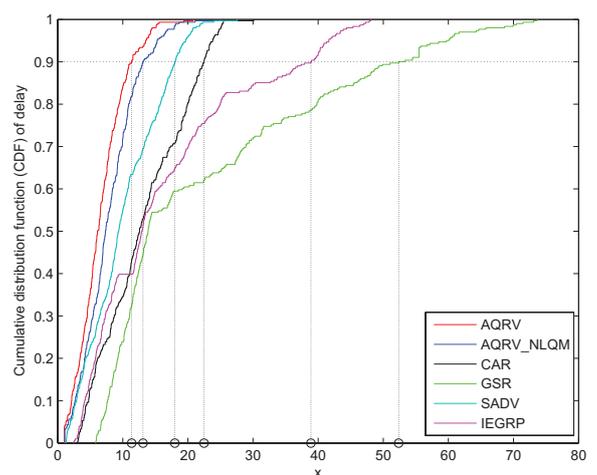


Fig. 11. CDF of end-to-end delay.

to update the latest pheromone immediately. Obviously, it is advantageous to coping with rapid topology changes in VANET environments and improving packet delivery ratio. CAR does not maintain any backup routes and provides only one complete end-to-end routing path, which is pretty much vulnerable in VANET environments and causes higher packet losses. GSR is a source driven routing protocol, and it makes routing decisions using the distance to the destination while neglecting other available routing information, so some data packets may be dropped when encountering extreme network conditions. In SADV routing protocol, packet forwarding delay between two adjacent intersections is estimated by periodically relaying packet, and delay matrixes stored in intersections are revised via delay update message broadcast, these two schemes may result in serious network congestions and low packet delivery ratio. IEGRP can efficiently forward data packets by means of wire-connected intersections, but the rest of routes are selected using a simple greedy carry-and-forward algorithm, which ignores traffic information and may lead to packet loss. In addition, compared with AQRV_NLQM, AQRV indicates a little higher packet delivery ratio. The reason is that LQM models enable AQRV to avoid redundant overhead and alleviate network congestions, so as to improve packet delivery ratio.

2) *End-to-end delay analysis:* Fig. 11 compares the Cumulative Distribution Function (CDF) of end-to-end delay for different routing protocols, where we set vehicle spacing density $\lambda = 0.015$ vehicle/m. We notice that the delay of AQRV is lower compared to the remaining protocols. Indeed, 90% delay statistics of AQRV are within 11.36 s, and 90% delay records of AQRV_NLQM are less than 13.11 s, while for CAR, GSR, SADV and IEGRP, the delay thresholds responding to the CDF value 0.9 are 27.49 s, 52.37 s, 17.99 s and 38.39 s, respectively. The reasons of the best delay performance of AQRV are as follows. In the first place, due to the ACO paradigm, AQRV makes different communication pairs collaborate with each other to update available routing information, which is advantageous to coping with the rapid network topology changes and adaptively choosing the

low-delay routes. Secondly, AQRV dynamically selects the next intersection using global routing information, and this scheme helps to make routing decisions based on the latest traffic information and effectively avoid network holes on upcoming communication road segments. Finally, compared with AQRV_NLQM, AQRV benefits the network congestion alleviation and packet retransmission reduction, which decrease one hop delay. Fig. 11 also represents that the curve of AQRV offers a more rapid convergence, which implies that the delay variance of AQRV is smaller than that of other routing protocols. Actually, AQRV takes delay variance into account in the process of routing exploration and selection, and it also explores available routing paths using both the latest local and global QoS, which is conducive to maintain stable routes. GSR and IEGRP ignore traffic information along routing paths, and data packets are forwarded by means of the distance-greedy rule, so the transmission delay changes dramatically as the traffic conditions vary along the route. CAR is a min-delay routing protocol but can not update routing information in real time, so the upcoming data packets may suffer from serious network partitions leading to higher delay. In SADV routing protocol, due to transmission time of delay update messages, delay matrixes are not updated in real time and the selected routing path may not be optimal, which then cause more delay.

3) *Overhead analysis:* Fig. 12 represents the generated overhead by all routing protocols as a function of different vehicle spacing densities. Note that for a fair comparison, we define the overhead as the ratio between the total control packet bytes and the cumulative bytes of successfully received data packets. From this figure, we find that the overheads of all routing protocols increase when the vehicle spacing densities go up, since the hello control packets occupy an important proportion of the whole overheads and their number is mainly determined by the vehicle spacing densities. In addition, Fig. 12 shows that the overheads of AQRV are lowest compared with those of other routing protocols. The reasons are as follows: (1) AQRV only explores the optimal route between two terminal intersections by ant unicast forwarding rather than between end-to-end communication

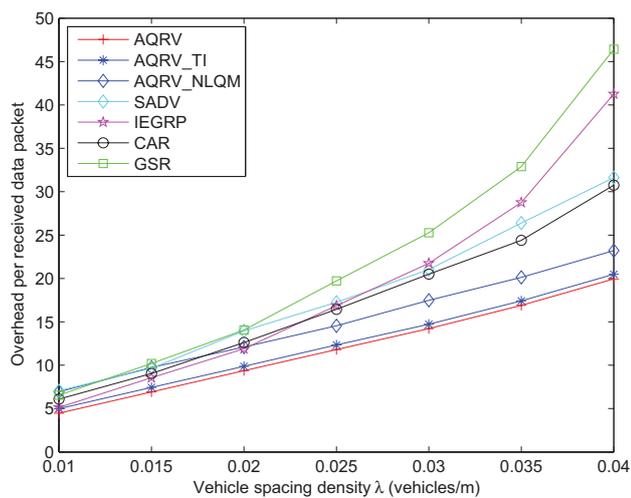


Fig. 12. Overhead for per successfully received data byte.

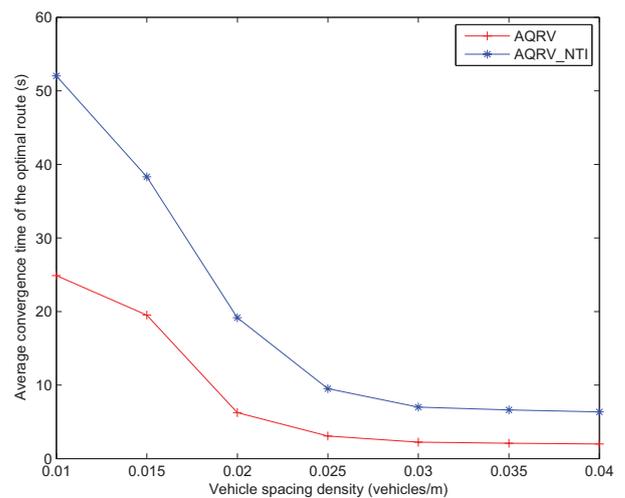


Fig. 13. Average convergence time of the optimal route.

pairs by packet broadcast, (2) based on the existing optimal route, AQRV can directly carry out data forwarding without new routing exploration, and (3) AQRV employs the latest routing information to dynamically choose the optimal next intersection, which enables higher packet delivery ratio and thus reduces the overhead. Compared to the overhead of CAR, GSR, SADV and IEGRP, AQRV_NLQM shows a lower value due to its higher packet delivery ratio and unicast ant forwarding scheme of AQRV_NLQM. Finally, the overhead of AQRV_NTI is higher than that of AQRV, as AQRV_NTI has to initiate a new optimal route establishment process for each communication pair via ants forwarding, which in turn generates more overhead.

4) *Optimal Route convergence time analysis:* Fig. 13 depicts the average optimal route's convergence time for AQRV and AQRV_NTI. From this figure, we observe that the average convergence time of the optimal route in AQRV is shorter than that in AQRV_NTI. Obviously, using already explored optimal route, AQRV can directly implement data forwarding for different communication pairs with same terminal intersections and QoS requirements, so this scheme helps to decrease the convergence time, while AQRV_NTI has to initiate an optimal route establishment process for each communication pair, which may lead to network congestions and end-to-end delay increase. Besides, with the increment of vehicle spacing density, network partitions can be repaired and convergence time is reduced.

5) *Influence of global pheromone update interval and number of forward ants:* Table II shows the normalized QoS and overhead of selected optimal route, in which we vary the update interval of global pheromone T_{GP} and the number of forward ants N_{fant} . Here we define the normalized QoS as $F(y)$ expressed in (1). From table II, we can observe that the normalized QoS, overhead and average convergence time decrease for larger T_{GP} , which implies that AQRV does not need to send forward ants in time to keep up with the topology changes, and data packets can be directly forwarded by means of explored routes in a long period, obviously, it reduces the

TABLE II
THE INFLUENCE OF T_{GP} AND N_{fant}

T_{GP}	N_{fant}	QoS	Overhead	Convergence time
10 s	10	0.8768	13.1241	0.9521 s
	30	0.9001	13.9102	1.2235 s
	50	0.9105	14.1499	1.4814 s
	100	0.9112	15.3222	3.7516 s
	200	0.9147	16.8726	5.5950 s
30 s	10	0.8693	12.7760	0.6150 s
	30	0.8916	12.9304	0.7419 s
	50	0.9008	12.9997	1.4707 s
	100	0.9017	13.6592	2.5700 s
	200	0.9018	14.7214	3.9909 s
50 s	10	0.8535	12.3167	0.5443 s
	30	0.8629	12.4942	0.2695 s
	50	0.8891	12.8158	0.3388 s
	100	0.8900	13.1210	0.5983 s
	200	0.8900	13.9870	1.7605 s

number of required forward ants and the convergence time of the optimal path, but the routing information obtained may be out of date. In addition, Table II indicates that the normalized QoS is proportional to the number of forward ants, because more forward ants can increase the routing exploration randomness and enhance AQRV's global exploration ability, which is beneficial to the optimal route search but increases the average convergence time. When the number of forward ants rises from 50 to 200, the QoS only improves a little, but the overhead increases rapidly. In this paper, in order to enable AQRV to keep a good balance among the QoS, overhead and average convergence time, we set T_{GP} and N_{fant} to 10 s and 50, respectively. Moreover, in order to investigate the correlation between the required forward ant number and traffic network size for a given desired QoS, we carry out experiments for various network size configurations. The results show that the duplicate network size does not need to double the number of required forward ants for a fixed QoS value. For example, based on the following two traffic environments (2500 m × 5000 m vs 5000 m × 5000 m), and for a targeted QoS equal to 0.9105, 32 forward ants on average are required per route establishment in the smaller

TABLE III
THE INFLUENCE OF η AND T_{eva}

η	T_{eva}	QoS	Overhead	Convergence time
0.85	1.0 s	0.8723	14.8789	1.3064 s
	1.5 s	0.8853	14.6822	1.5456 s
0.90	1.0 s	0.8802	14.7579	1.4243 s
	1.5 s	0.9089	14.1812	1.9145 s
0.95	1.0 s	0.9105	14.1499	1.4814 s
	1.5 s	0.9107	14.1476	2.5827 s

traffic environment while 50 forward ants are used in the other one (which corresponds to an increase of 56.25%). This behavior of AQRV is inherited from the adaptive and cooperative features of ants, which collaborate with each other to accelerate routing establishments based on already discovered local and global pheromone, thereby reducing the use of new ants.

6) *Influence of pheromone evaporation factor and evaporation time interval:* Table III presents the effects of pheromone evaporation factor η and evaporation time interval T_{eva} on normalized QoS, overhead and average convergence time. From this table, we can see that both QoS and convergence time increase with the rises of η and T_{eva} , which result in a slight decline in overhead. The reasons are given as follows: higher values of η and T_{eva} cause less pheromone to be evaporated for a period of time. They also implies that the influences of the latest pheromone are reduced and the random effects of routing exploration are enhanced, which are beneficial to deep routing searches and better QoS achievements but at the expense of higher convergence time. In addition, according to the mentioned definition, overhead is a ratio between the total control packet bytes and the cumulative bytes of successfully received packets, so better QoS implies higher packet delivery ratio which induces lower overhead.

VI. CONCLUSION

In this paper, we have proposed a new protocol AQRV for adaptively selecting the best routing path in vehicular urban environments. The challenge of dynamic QoS-based routing selection is firstly regarded as a multi-objective optimization problem, which we have solved through proposing an ACO-based algorithm. By employing local and global pheromone, available candidate routing paths between terminal intersections are explored by forward ants, then the corresponding backward ants are used to choose the optimal route and update the latest global pheromone. Besides, we have derived local QoS models in 1-lane road segment scenarios via considering three metrics, namely connectivity probability, packet delivery ratio and delay. Finally, we have validated our analytical local QoS models through extensive simulations, and demonstrated that AQRV outperforms the reference protocols (GSR and CAR). The simulation results also represent the effectiveness of the LQM and TI concept. For future works, we would further investigate the effects of AQRV weight parameters on the relaying quality of road segments, and compare the performance of ACO to those of other optimization methods.

REFERENCES

- [1] S. Ian and O. Tonguz, "Enhancing vanet connectivity through roadside units on highways," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 8, pp. 3586–3602, Aug. 2011.
- [2] G. Zhioua, N. Tabbane, H. Labiod, and S. Tabbane, "A fuzzy multi-metric qos-balancing gateway selection algorithm in a clustered vanet to lte advanced hybrid cellular network," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 2, pp. 804–817, Feb. 2015.
- [3] A. Mohammad and M. Robert, "A new hybrid location-based ad hoc routing protocol," in *Proc. IEEE GLOBECOM*, Miami, USA, Dec. 2010, pp. 6–10.
- [4] S. Sultan, M. Doori, A. Bayatti, and H. Zedan, "A comprehensive survey on vehicular ad hoc network," *Journal of Network and Computer Applications*, vol. 37, no. 1, pp. 380–392, Jan. 2014.
- [5] M. Altayeb and I. Mahgoub, "A survey of vehicular ad hoc networks routing protocols," *International Journal of Innovation and Applied Studies*, vol. 3, no. 3, pp. 829–846, Jul. 2013.
- [6] J. Zhao and G. Cao, "Vadd: Vehicle-assisted data delivery in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 3, pp. 1910–1922, May 2008.
- [7] B. Seet, G. Liu, B. Lee, C. Foh, and K. Wong, "A-star: A mobile ad hoc routing strategy for metropolis vehicular communications," in *Proc. IEEE NETWORKING*, Athens, Greece, May 2004, pp. 989–999.
- [8] M. Jerbi, T. Rasheed, and Y. Doudane, "Towards efficient geographic routing in urban vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 5048–5058, Nov. 2009.
- [9] C. Ma and N. Liu, "Traffic-aware data delivery scheme for urban vehicular sensor networks," *International Journal of Distributed Sensor Networks*, vol. 9, no. 4, pp. 1–12, Aug. 2013.
- [10] V. Naumov and T. Gross, "Connectivity-aware routing (car) in vehicular ad hoc networks," in *Proc. IEEE INFOCOM*, Anchorage, USA, May 2007, pp. 1919–1927.
- [11] W. Zhang, Y. Chen, Y. Yang, and G. Mao, "Multi-hop connectivity probability in infrastructure-based vehicular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 1–9, May 2012.
- [12] Y. Ding and X. Li, "Sadv: Static-node-assisted adaptive data dissemination in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2245–2255, Jun. 2010.
- [13] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "Vanet routing on city roads using real-time vehicular traffic information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3609–3626, Sep. 2009.
- [14] P. Sahu, E. Wu, J. Sahoo, and M. Gerla, "Ddor: Destination discovery oriented routing in highway/freeway vanets," *Telecommunication Systems*, vol. 50, no. 4, pp. 267–284, Aug. 2012.
- [15] S. Bitam and A. Mellouk, "Qos swarm bee routing protocol for vehicular ad hoc networks," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, p. 11C5.
- [16] Y. Sun, S. Luo, Q. Dai, and Y. Ji, "An adaptive routing protocol based on qos and vehicular density in urban vanets," *International Journal of Distributed Sensor Networks*, vol. 1, no. 1, pp. 1–14, Mar. 2015.
- [17] I. Leontiadis and C. Mascolo, "Geopps: Geographical opportunistic routing for vehicular networks," in *Proc. IEEE WOWMOM*, Espoo, Finland, Jun. 2007, pp. 1–6.
- [18] Z. Mo, H. Zhao, K. Makki, and N. Pissinou, "Muru: A multi-hop routing protocol for urban vehicular ad hoc networks," in *Proc. IEEE MOBIQ*, California, USA, Jul. 2006, pp. 1–8.
- [19] P. Sahu, E. Wu, J. Sahoo, and M. Gerla, "Bahg: Back-bone-assisted hop greedy routing for vanets city environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 199–213, Mar. 2013.
- [20] C. Lochert, H. Hartenstein, D. Hermann, and M. Mauve, "A routing strategy for vehicular ad hoc networks in city environments," in *Proc. IEEE IVS*, Columbus, USA, Jun. 2003, pp. 156–161.
- [21] A. O'Driscoll and D. Pesch, "An infrastructure enhanced geographic routing protocol for urban vehicular environments," in *Proc. IEEE WiVeC*, Dresden, Germany, 2013, pp. 1–5.
- [22] S. Wankhade and M. Ali, "Ant based techniques for qos routing in mobile ad hoc network: An overview," *International Journal of Advanced Networking and Applications*, vol. 3, no. 3, pp. 1094–1097, Jun. 2011.
- [23] S. Kamali and J. Opatrny, "A position based ant colony routing algorithm for mobile ad-hoc networks," *Journal of Networks*, vol. 3, no. 4, pp. 459–462, Apr. 2008.
- [24] L. Gambardella and M. Dorigo, "Ant colony system: a cooperative learning approach to the traveling salesman problem," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 53–66, Apr. 1997.

- [25] G. Li and L. Boukhatem, "Adaptive vehicular routing protocol based on ant colony optimization," in *Proc. IEEE VANET*, Taipei, China, May 2013, pp. 95–98.
- [26] J. Bell and P. McMullen, "Ant colony optimization techniques for the vehicle routing problem," *Advanced Engineering Informatics*, vol. 18, no. 1, pp. 41–48, Jul. 2004.
- [27] R. Silva, H. Lopes, and W. Godoy, "A heuristic algorithm based on ant colony optimization for multi-objective routing in vehicle ad hoc networks," in *Proc. IEEE BRICS-CCI-CBIC*, Ipojuca, Brazil, Sep. 2013, pp. 435–440.
- [28] G. Caro, F. Ducatelle, and L. Gambardella, "Anthocnet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks," *European Transactions on Telecommunications*, vol. 16, no. 5, pp. 443–455, Sep. 2005.
- [29] S. Bilal, C. Bernardos, and C. Guerrero, "Position-based routing in vehicular networks: A survey," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 685–697, Mar. 2013.
- [30] D. Borsetti and J. Gozalvez, "Infrastructure-assisted geo-routing for cooperative vehicular networks," in *Proc. IEEE VNC*, Jersey City, USA, Dec. 2010, pp. 255–262.
- [31] A. Bazzi, B. Masini, A. Zanella, and G. Pasolini, "Vehicle-to-vehicle and vehicle-to-roadside multi-hop communications for vehicular sensor networks: Simulations and field trial," in *Proc. IEEE ICC*, Budapest, Hungary, Jun. 2013, pp. 515–520.
- [32] R. Shrestha, S. Moh, I. Chung, and D. Choi, "Vertex-based multihop vehicle-to-infrastructure routing for vehicular ad hoc networks," in *Proc. IEEE HICSS*, Honolulu, USA, Jan. 2010, pp. 1–7.
- [33] D. Hao, X. Zhao, L. Qu, X. Chi, and X. Cui, "Multi-hop routing optimization method based on improved ant algorithm for vehicle to roadside network," *Journal of Bionic Engineering*, vol. 11, no. 3, pp. 490–496, Jul. 2014.
- [34] C. Perkins and E. Royer, "Ad-hoc on-demand distance vector routing," in *Proc. IEEE WMCSA*, Los Angeles, USA, Feb. 1999, pp. 90–100.
- [35] L. Zhang and B. Jin, "Dubhe: A reliable and low-latency data dissemination mechanism for vanets," *Journal of Distributed Sensor Networks*, vol. 2013, no. 4, pp. 1–17, Jun. 2013.
- [36] H. Saleet, R. Langar, K. Naik, R. Boutaba, A. Nayak, and N. Goel, "Intersection-based geographical routing protocol for vanets: A proposal and analysis," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4560–4574, Dec. 2011.
- [37] S. Panichpapiboon and W. Pattara-atikom, "Connectivity requirements for self-organizing traffic information systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3333–3340, Nov. 2008.
- [38] T. Rappaport, *Wireless Communications: Principles and Practice*. California, USA: Prentice Hall PTR, 1996.
- [39] S. Haykin and M. Moher, *Modern Wireless Communications*. Massachusetts, USA: Prentice Hall, 2004.
- [40] A. Agarwal, D. Starobinski, and T. Little, "Phase transition of message propagation speed in delay-tolerant vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 249–263, Mar. 2012.
- [41] P. Athanasios and P. Unnikrishna, *Probability, random variables and stochastic processes*. New York, USA: McGraw-Hill, 2002.
- [42] R. Oliveira, L. Bernardo, and P. Pinto, "Modelling delay on IEEE 802.11 mac protocol for unicast and broadcast nonsaturated traffic," in *Proc. IEEE WCNC*, Hongkong, China, Mar. 2007, pp. 463–467.
- [43] J. Harri, F. Filali, C. Bonnet, and M. Fiore, "Vanetmobsim: generating realistic mobility patterns for vanets," in *Proc. IEEE VANET*, Los Angeles, USA, Sep. 2006, pp. 96–97.
- [44] A. Ghandour, M. Felice, H. Artail, and L. Bononi, "Modeling and simulation of wave 1609.4-based multi-channel vehicular ad hoc networks," in *Proc. ACM SIMUTOOLS*, Sirmione-Desenzano, Italy, Mar. 2012, pp. 19–23.



G. Li received the M.Sc. degree in communication and information system from Tongji University, Shanghai, China, in 2011. He is currently working towards the Ph.D. degree in the ROCS (Networks and combinatorial and stochastic optimization) team of LRI laboratory, University of Paris-Sud 11.

His main research interests include wireless network protocols, vehicular ad hoc networks, smart antennas, vehicle electro magnetic compatibility and their applications in intelligent transportation systems.



L. Boukhatem received the computer science engineering degree in 1997 from INI Institute, M.Sc. degree in 1998 from the University of Versailles Saint Quentin-en-Yvelines and Ph.D. degree in 2001 from the University of Pierre et Marie Curie (Paris 6).

She is currently an associate professor in the ROCS (Networks and combinatorial and stochastic optimization) team of LRI laboratory (University of Paris-Sud 11). She joined the LRI laboratory in 2002 where she is developing various research works in mobile and wireless networks including: Cross-layer design, modeling and performance evaluation, radio resource allocation and optimization (TDMA and OFDMA systems), mobility management, routing and energy in ad hoc networks. She has lead and actively participated to several national and international projects and programs such as PHC Sakura, FP7-NoE NEWCOM, EIT ICT-Labs Digital Cities. Her current research interests are routing in vehicular networks and interference mitigation in OFDMA systems.



J. Wu (SM'11) received the Ph.D. degree in electrical engineering from Queens University, Canada.

He is the Founder and Founding Chair of Technical Committee on Green Communications and Computing (TCGCC), IEEE Communications Society, which was established in 2011 as an official Technical Subcommittee (TSCGCC) and elevated as TCGCC in 2013. He is also the co-founder and Founding Vice-Chair of the Technical Subcommittee on Big Data (TSCBD) in IEEE Communications Society. He is Associate Editor of IEEE Communications Surveys & Tutorials, Associate Editor of IEEE Systems Journal, Associate Editor of IEEE Access. He is the founder and Series Editor on Green Communication and Computing networks of IEEE Communications Magazine. He has served as the leading Guest Editor of Special Issue on Green Communications, Computing, and Systems in IEEE Systems Journal, and Associate Editor of Special Section on Big Data for Green Communications and Computing in IEEE Access. He was the leading Editor and a co-author of the comprehensive book, entitled *Green Communications: Theoretical Fundamentals, Algorithms, and Applications*, published by CRC Press in September 2012.