Routing Protocols in Vehicular Ad Hoc Networks: A Survey and Future Perspectives

YUN-WEI LIN[1], YUH-SHYAN CHEN[2], AND SING-LING LEE[1]

[1]*Department of Computer Science and Information Engineering*
National Chung Cheng University, Chia-Yi, Taiwan, R.O.C.
E-mail: {*jyneda@gmail.com, singling@cs.ccu.edu.tw*}

[2]*Department of Computer Science and Information Engineering*
National Taipei University, Taipei, Taiwan, R.O.C.
E-mail: yschen@mail.ntpu.edu.tw

Vehicular Ad hoc Network (VANET), a subclass of mobile ad hoc networks (MANETs), is a promising approach for the intelligent transportation system (ITS). The design of routing protocols in VANETs is important and necessary issue for support the smart ITS. The key difference of VANET and MANET is the special mobility pattern and rapidly changeable topology. It is not effectively applied the existing routing protocols of MANETs into VANETs. In this investigation, we mainly survey new routing results in VANET. We introduce unicast protocol, multicast protocol, geocast protocol, mobicast protocol, and broadcast protocol. It is observed that carry-and-forward is the new and key consideration for designing all routing protocols in VANETs. With the consideration of multi-hop forwarding and carry-and-forward techniques, min-delay and delay-bounded routing protocols for VANETs are discussed in VANETs. Besides, the temporary network fragmentation problem and the broadcast storm problem are further considered for designing routing protocols in VANETs. The temporary network fragmentation problem caused by rapidly changeable topology influence on the performance of data transmissions. The broadcast storm problem seriously affects the successful rate of message delivery in VANETs. The key challenge is to overcome these problems to provide routing protocols with the low communication delay, the low communication overhead, and the low time complexity. The challenges and perspectives of routing protocols for VANETs are finally discussed.

Keywords: vehicular ad hoc network, carry-and-forward, routing, min-delay routing, delay-bounded routing

1. INTRODUCTION

The growth of the increased number of vehicles are equipped with wireless transceivers to communicate with other vehicles to form a special class of wireless networks, known as vehicular ad hoc networks or VANETs [1]. To enhance the safety of drivers and provide the comfortable driving environment, messages for different purposes need to be sent to vehicles through the inter-vehicle communications. *Unicast* routing is a fundamental operation for vehicle to construct a source-to-destination routing in a VANET as shown in Fig. 1(a). *Multicast* is defined by delivering multicast packets from a single source vehicle to all multicast members by multi-hop communication. *Geocast* routing is to deliver a geocast packet to a specific geographic region. Vehicles located in this specific geographic region should receive and forward the geocast packet; otherwise, the packet is dropped as shown in Fig. 1 (b). *Broadcast* protocol is utilized for a source vehicle sends broadcast message to all other vehicles in the network as shown in Fig. 1(c).

Many results [2][3][4] on MANETs have been proposed for unicast, multicast and geocast, and broadcast protocols. However, VANETs are fundamentally different to MANETs, such as the special mobility pattern and rapid changed topology. This key differentiation causes the existing routing protocol on MANETs can not be directly applied to VANETs. In this investigation, the recent new results for VANET routing mechanism are first surveyed. Fig. 2 shows that the survey is structured into three broad categories;

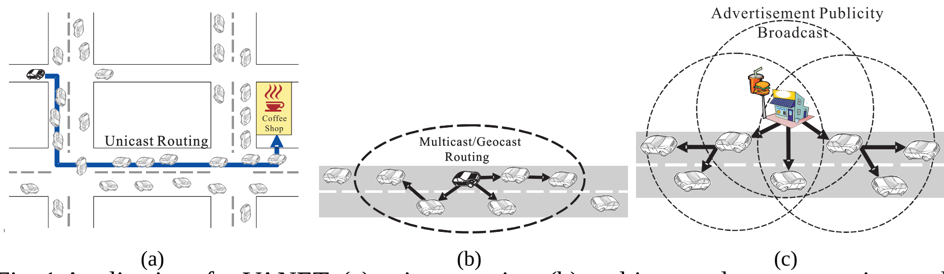


Fig. 1 Applications for VANET: (a) unicast routing, (b) multicast and geocast routing, and (c) broadcast routing

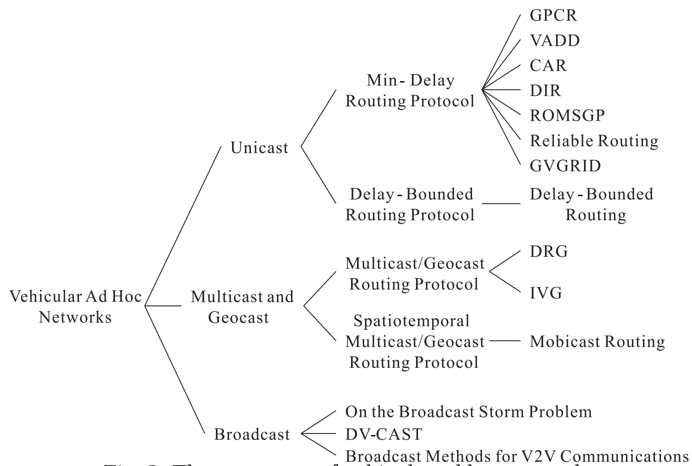


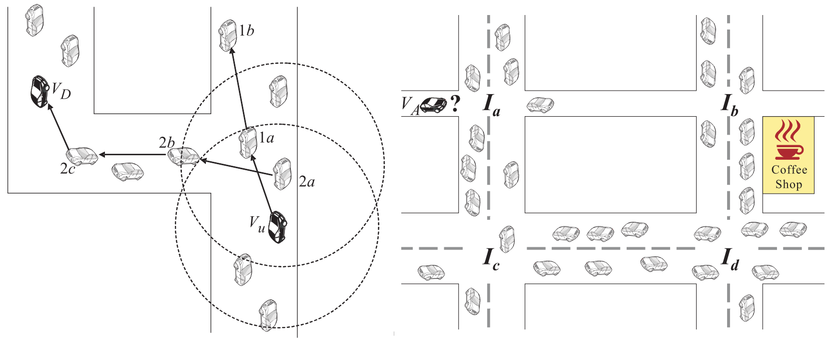
Fig. 2. The taxonomy of vehicular ad hoc networks

unicast [5][6][7][8][9][10][11][12], multicast and geocast [13][14][15], and broadcast approaches[16][17][18]. The key ideas of representative technologies in each category are described.

The remainder of this paper is organized as follows. Section 2 reviews unicast routing protocols in VANETs. Section 3 introduces multicast and geocast routing protocols in VANETs. Section 4 describes broadcast routing protocols in VANETs. Section 5 concludes this paper and gives some possible future perspectives for VANETs.

2. UNICAST ROUTING PROTOCOL

This section introduces the unicast routing protocols in VANETs. The main goal of unicast routing in VANETs is to transmit data from a single source to a single destination via wireless multi-hop transmission or carry-and-forward techniques. In the wireless multi-hop transmission technique, or called as multi-hop forwarding, the intermediate vehicles in a routing path should relay data as soon as possible from source to destination. In the carry-and-forward technique, source vehicle carries data as long as possible to reduce the number of data packets. The delivery delay-time cost by carry-and-forward technique is normally longer than wireless multi-hop transmission technique. Two categories of routing protocol designing are classified, *min-delay* routing protocol and *delay-bounded* routing protocol. Min-delay routing protocol aims to minimize the delivery delay-time from



(a) (b)

Fig. 3. (a) Geographic routing protocol and (b) VADD routing protocol

source to destination. Delay-bounded routing protocol attempts to maintain a low level of channel utilization within the constrained delivery delay-time. This section describes existing unicast routing protocols in VANETs as follows.

2.1 Min-Delay Routing Protocol

The goal of min-delay routing protocols is to transmit data packets to destination as soon as possible. The transmission delay time is the major concern and the shortest routing path is usually adopted. However, the shortest routing path may be not the quickest path with the minimum delay time in VANETs. The shortest routing path may be found in a low density area, packets can not transmit by the multi-hop forwarding since that there is no neighboring vehicle can forward packets. Packets should be delivered by carry-and-forward scheme. The delay time is greatly growing if the multi-hop forwarding can not be utilized. Efforts will be made as finding a routing path with multi-hop forwarding. The min-delay routing protocols [5][6][7][8][9][10][11] are reviewed as follows.

2.1.1 Greedy Perimeter Coordinator Routing Protocol

Lochert *et al*. [5] proposed GPCR (greedy perimeter coordinator routing) which is a position-based routing for urban environment. GPCR protocol is very well suited for highly dynamic environments such as inter-vehicle communication on the highway or city. GPCR traverses the junctions by a restricted greedy forwarding procedure, and adjusts the routing path by the repair strategy which is based on the topology of streets and junctions. Fig. 3(a) shows that vehicle V_u tries to send packets to vehicle V_D. Vehicle $1a$ is selected as the next hop of V_u if greedy forwarding scheme is used. After vehicle $1a$ received the packets, vehicle $1a$ detects destination V_D is not located at north. Vehicle $1a$ then moves packets backward vehicle $2a$, then the packet is forwarded to V_D.

2.1.2 VADD: Vehicle-Assisted Data Delivery Routing Protocol

Data delivery routing protocol is developed by Zhao *et al*. [6], called as VADD. VADD protocol adopted the idea of carry-and-forward for data delivery from a moving vehicle to a static destination. The most important issue is to select a forwarding path with the smallest packet delivery delay. To keep the low data transmission delay, VADD protocol transmits packets through wireless channels as much as possible, and if the packet has to be carried

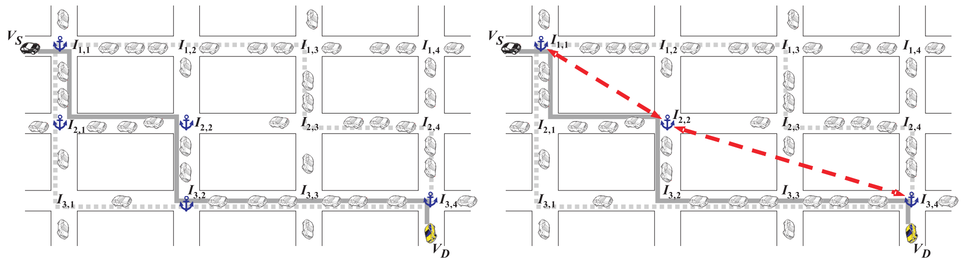


Fig. 4 (a) CAR is an anchor-based routing protocol and (b) DIR protocol sets fewer anchors than CAR protocol

through roads, the road with higher speed is chosen firstly. VADD protocol assumes that vehicles are equipped with pre-loaded digital maps, which provide street-level map and traffic statistics such as traffic density and vehicle speed on roads at different times of the day. According to the information provided by digital maps, VADD protocol proposed a *delay* model to estimate the data delivery delay in different roads as follows,

$$d_{ij} = (1 - e^{-R \times \rho_{ij}}) \times \frac{l_{ij} \times c}{R} + e^{-R \times \rho_{ij}} \times \frac{l_{ij}}{v_{ij}},$$

where d_{ij} is the expected packet forwarding delay from intersection I_i to intersection I_j, R is the communication range of vehicle, c is a constant used to adjust expected packet forwarding delay to a more reasonable value, r_{ij} is the road from intersection I_i to intersection I_j, ρ_{ij} is the vehicle density on r_{ij}, l_{ij} is the Euclidean distance of r_{ij}, and v_{ij} is the average vehicle velocity on r_{ij}. With the delay model, VADD protocol estimates the best road with the lowest data delivery delay based on the current kept traffic patterns. Fig. 3(b) illustrates that vehicle V_A tries to send packets to the coffee shop, while the coffee shop is at the fixed location. Intersections I_a, I_b, I_c, and I_d are considered as the candidate intermediate intersections. After evaluating the expected forwarding delay, intersections I_a, I_c, and I_d are chosen. This is because that the density of vehicle is high between intersections I_a, I_c, and I_d, although it is not the shortest path.

2.1.3 Connectivity-aware Routing Protocol

To overcome the limitation of the static destination, Naumov *et al.* [7] proposed Connectivity-Aware Routing (CAR) protocol. CAR protocol establishes a routing path from source to destination by setting the anchor points at intermediate junctions. CAR protocol sends the searching packets to find the destination. Each forwarding vehicle records its ID, hop count, and average number of neighbors in searching packets. Once the searching packets reach the destination, the destination chooses a routing path with the minimum delivery delay time and replies it to the source. While destination sends the reply packet to the source, the junctions passed through by the reply packet are set as the anchor point. After the path set up, data packets are forwarded in a greedy method toward the destination through the set of anchor points. Fig. 4(a) gives that vehicle V_S tries to send data to vehicle V_D, the anchor points are set at $I_{1,1}$, $I_{2,1}$, $I_{2,2}$, $I_{3,2}$, and $I_{3,4}$. Data is forwarded according to order in the list of anchor points.

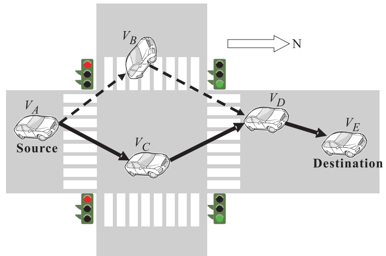


Fig. 5. ROMSGP routing protocol

2.1.4 DIR: Diagonal-Intersection-Based Routing Protocol

To improve the CAR protocol, Chen *et al.* [8] developed a diagonal-intersection-based routing (DIR) protocol. The key difference of CAR and DIR protocols is that DIR protocol [8] constructs a series of diagonal intersections between the source and destination vehicles. The DIR protocol is a geographic routing protocol. Based on the geographic routing protocol, source vehicle geographically forwards the data packet toward the first diagonal intersection, the second diagonal intersection, and so on, until the last diagonal intersection, and finally geographically reaches to the destination vehicle. For given a pair of neighboring diagonal intersections, two or more disjoint sub-paths exist between them. The novel property of DIR protocol is the auto-adjustability, while the auto-adjustability is achieved that one sub-path with low data packet delay, between two neighboring diagonal intersections, is dynamically selected to forward data packets. To reduce the data packet delay, the route is automatically re-routed by the selected sub-path with lowest delay. Fig. 4(b) shows that DIR protocol constructs a series of diagonal intersections between vehicles V_S and V_D. Observe that, DIR protocol may set the fewer number of anchors than CAR protocol [7]. DIR protocol can automatically adjust routing path for keeping the lower packet delay, compared to CAR protocol [7].

2.1.5 ROMSGP Routing Protocol

To improve the routing reliability, Taleb *et al*. [9] proposed ROMSGP (Receive on Most Stable Group-Path) routing protocol in a city environment. Taleb *et al.* indicate that an unstable routing usually occurred due to the loss of connectivity if one vehicle moves out of the transmission range of a neighboring vehicle. In ROMSGP protocol, all vehicles are split into four groups based on the velocity vector. A routing is said as a stable routing if the two vehicles are categorized in the same group; otherwise, the routing is an unstable routing. A vehicle belongs to a group if the velocity vector has the maximum projection vector with this group. Fig. 5 illustrates the ROMSGP routing protocol. Two routing paths are established, $\{V_A V_B,\ V_B V_D\}$ and $\{V_A V_C,\ V_C V_D\}$. If V_A, V_B, V_C, and V_D belong to the same group, the two routing paths are both stable. Packet is delivered via $\{V_A V_B,$ $V_B V_D\}$ or $\{V_A V_C, V_C V_D\}$. If V_B turns into another road, the projection vector is changed. V_B belongs to the other group. Then the routing path $\{V_A V_C, V_C V_D\}$ is the only choice.

2.1.6 Reliable Routing for Roadside to Vehicle Communications

In contrast with routing results developed in the highway or the city environments, it is very interest that Wan *et al*. [11] specially proposed a reliable routing protocol in the *rural* environment. Wan *et al*. [11] proposed two reliable routing strategies for roadside to vehicle (R2V) communication. The challenge of R2V communication in the rural environment

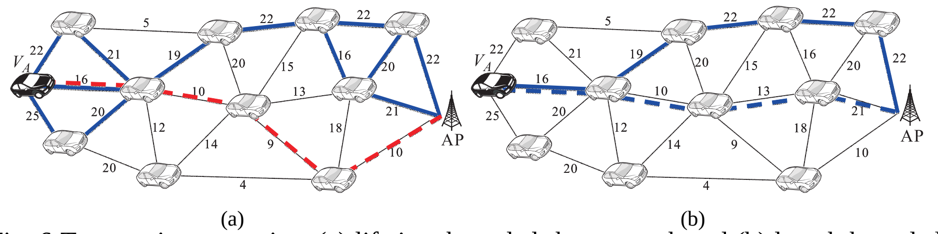
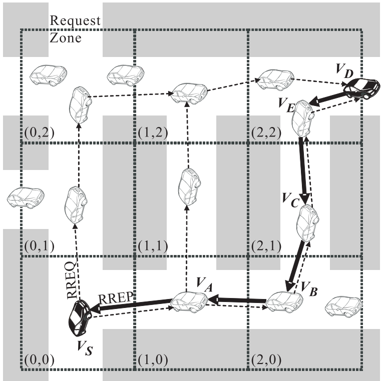


Fig. 6 Two routing strategies: (a) lifetime-bounded shortest path and (b) length-bounded maximum lifetime path

is the *terrain* factor. For instance, a vehicle moving along the rural highway occasionally loses the line of sight (LOS) to the neighbor vehicle or to access points (APs) due to the obstacle-property caused by the curve roadway and mountains. In addition, almost no fixed communication infrastructure is available. Multi-hop inter-vehicle communication connecting to AP is the main solution of the R2V communication. The link lifetime is very important issue for designing the reliable routing. The link lifetime is predicted by two conditions. Once the communication is established, the link lifetime halts if (1) LOS between a pair of vehicles is lost, or (2) one vehicle moves out of the communication range of the neighboring vehicle. A link established in a shorter distance usually has longer link lifetime. The link lifetime is used to predict the lifetime of a route. A route is constructed by a series of links. The lifetime of a route is the minimum link lifetime in a route. Long lifetime of a route improves the routing reliability if considered the lifetime-bounded shortest path. In addition to the lifetime of a routing path, the length-bounded maximum lifetime path is considered. To construct a length-bounded maximum lifetime path, reducing hops can improve the delivery delay-time. A routing path with fewer hops means the links are established in the long distance. Establishing a routing path with longer lifetime implies that the length of this routing path is long. Fig. 6(a) illustrates the example of lifetime-bounded shortest path. The dotted line is current routing path and the link lifetime is going to end, where the minimum link lifetime is 9. The solid line is the candidate path. The link lifetime of solid line is greater than the threshold (=16). The routing path changes to solid line by AP assignment. Fig. 6(b) illustrates the example of length-bounded maximum lifetime path. The dotted line is the routing path with minimum hops to AP (hops=4). The solid line is the selected path (hops=5).

2.1.7 GVGrid: A QoS Routing Protocol

To improve delivery delay-time and routing reliability, Sun *et al*. proposed GVGrid protocol [10] which is a QoS routing protocol for VANETs. GVGrid constructs a routing path from source to destination by grid-based approach, which divides the map into several grids. The RREQ and RREP packets are delivered through different grid to find a routing path through minimum number of grid. A grid is chosen based on the direction and the distance between vehicle and intersection and is selected as next grid if the direction of grid is the same as current grid or the grid is closed to the intersection. Then the intermediate grids between source and destination are recorded in the routing table. An appropriate vehicle which has the fewest number of disconnections in each grid is chosen to forward packets to next grid. A formula of evaluating the expected number of disconnections is derived in [10]. The routing table records in terms of the source vehicle, destined grid, an appropriate vehicle as next hop with minimum the expected number of disconnections, a vehicle as previous hop, and the grid sequence. Once the routing path is broken, GV-



(a)

Routing table on V_A

src	dst	fwd	bwd	grid sequence
V_S	(2,2)	V_B	V_S	(2,2):(2,1):(2,0):(1,0):(0,0)

Routing table on V_D

src	dst	fwd	bwd	grid sequence
V_S	(2,2)	BCAST	V_E	(2,2):(2,1):(2,0):(1,0):(0,0)

(b)

Fig. 7. (a) An example of GVGrid routing protocol and (b) routing tables on vehicle V_A and V_D

Grid just finds another vehicle in the grid instead of the previous vehicle. The routing path dose not require finding again. Fig. 7(a) shows that vehicle V_S floods RREQ message to find vehicle V_D and vehicle V_D replies RREP message to V_S. Fig. 7(b) demonstrates that not only the grid sequence but also the information of the next vehicle are recorded in the routing table.

2.2 Delay-Bounded Routing Protocol

Skordylis *et al*. [12] proposed a delay-bounded routing protocol in VANETs, which provides a routing scheme that satisfy user-defined delay requirements while at the same time maintaining a low level of channel utilization. The delay-bounded routing protocol [12] focuses on the development of carry-and-forward schemes that attempts to deliver data from vehicles to static infrastructure access point in an urban environment. Two routing algorithms, D-Greedy (Delay-bounded Greedy Forwarding) and D-MinCost (Delay-bounded Min-Cost Forwarding), evaluate traffic information and the bounded delay-time to carefully opt between the Data Muling and Multihop Forwarding strategies to minimize communication overhead while satisfying with the delay constraints imposed by the application. D-Greedy algorithm adopts only local traffic information to make routing decisions. D-Greedy algorithm chooses the shortest path to destined AP form the map information, and then allocates the constrained delay-time to each street within the shortest path according to the length of streets. If packets can be delivered under the constrained delay-time in a street, Data Muling strategy is utilized. Packets are carried by a vehicle and forwarded at the vehicle's speed to destined AP. Otherwise, Multihop Forwarding strategy is applied if packets cannot be delivered within the constrained delay-time. Packets are delivered by multi-hop forwarding. D-MinCost algorithm considers the global traffic information in a city to achieve the minimum channel utilization within the constrained delay-time. According to the global traffic information, the cost and delay of each street can be pre-computed. The cost represents the number of message transmissions in a street. The delay denotes the time required to forward a message in a street. To achieve the minimum cost within the constrained delay, DSA (Delay Scaling Algorithm) [19] is applied to select the best routing path with minimum channel utilization under the constrained delay-time. Fig. 8 shows that Data Muling strategy is applied if the packet can be delivered form V_A to AP within the constrained delay-time. Otherwise, the packet is delivered by Multihop Forwarding

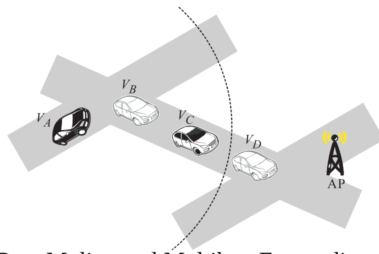


Fig. 8. Data Muling and Multihop Forwarding strategies

strategy.

2.3 Challenges and Future Perspectives

In this section, we have reviewed existing unicast routing protocols. Table 1 gives a detailed comparison of these protocols. Prior forwarding method describes the first routing decision of a protocol when there are packets to be forwarded. Observe that, delay-bounded routing protocol is different from other protocols, which carry-and-forward is the first considered routing decision to reserve the wireless media resource. Destination location method shows how a protocol discovers the routing path and destination, which are categorized into two types, specialized method and integrated method. Specialized method only indicates the destination location and the routing path is discovered while the packet is forwarding. The integrated method integrates the path discovery process into destination finding process. The routing path and destination location are simultaneously discovered. Generally, the integrated method has less routing setup time as well as increment of implementation complexity. Forwarding strategy is usually in greedy or optimum fashion, which expresses what information is considered when a protocol establishes the routing path. Greedy forwarding only considers the local information to make the routing decision. Optimum forwarding considers the entire information in a network to choose the best routing path. Normally, optimum forwarding has better performance; however, lots overheads are required. Recovery strategy describes route recovery strategy if the routing path is failed. All existing protocols adopt carry-and-forward method except delay-bounded routing protocol. Delay-bounded routing protocol adopts multi-hop forwarding to reduce the packet delivery time if the expected packet delivery time cannot satisfy the user-defined delay. Path maintenance shows that a protocol maintains the routing path in passive or active fashion. The passive path maintenance is performed only when a routing path is expired. The active path maintenance is performed if a routing path is inefficient. To actively maintain a routing path, the realistic traffic flow information is usually necessary to update the routing path. Most of protocols were developed in urban areas under the assumption of high network density. Therefore, some future perspectives should include the following:

1. A possible future work is how to design min-delay unicast routing approaching under low network density. The impact of intense density variability should be incorporated into the protocol design.
2. A major challenge in protocol design in VANETs is how to improve reliability of min-delay unicast routing protocols to simultaneously reduce delivery delay time and the number of packet retransmissions.

Table 1. Comparison of unicast routing protocols

Protocols	GPCR	VADD	CAR	DIR	ROMSGP	Reliable Routing	GVGRID	Delay-Bounded
Prior forwarding method	wireless multi-hop forwarding	wireless multi-hop forwarding	wireless multi-hop forwarding	wireless multi-hop forwarding	wireless multi-hop forwarding	wireless multi-hop forwarding	wireless multi-hop forwarding	carry-and -forward
Destination location method	specialized	specialized	integrated	integrated	specialized	integrated	integrated	specialized
Forwarding strategy	greedy forwarding	optimum forwarding	greedy forwarding	greedy forwarding	greedy forwarding	optimum forwarding	greedy forwarding	greedy forwarding
Recovery strategy	carry-and -forward	carry-and -forward	recompute anchors	recompute anchors	carry-and -forward	carry-and -forward	carry-and -forward	multi-hop forwarding
Path main-tenance	passive	active	active	active	passive	active	passive	passive
Digital map needed	no	yes	yes	yes	no	no	yes	no
Realistic traffic flow	no	yes	yes	yes	yes	yes	yes	no
Scenario	urban	urban	urban	urban	urban	rural area	urban	urban

3. Driver behavior should be considered for designing of delay-bounded unicast routing protocols since carry-and-forward method is the mainly approach to deliver packets.

4. To design a routing protocol in a city, the interference by tall buildings along roads should be considered. A robustness routing protocol against interference is appropriate developed in a city environment.

5. Scalability is also an important factor of routing protocol designing. The VANETs could be a large and metropolitan-scale networks. Protocols should consider that many unicast routing requests are operating simultaneously. The conflict of routing requests between vehicles should take into consideration, especially in the intersection.

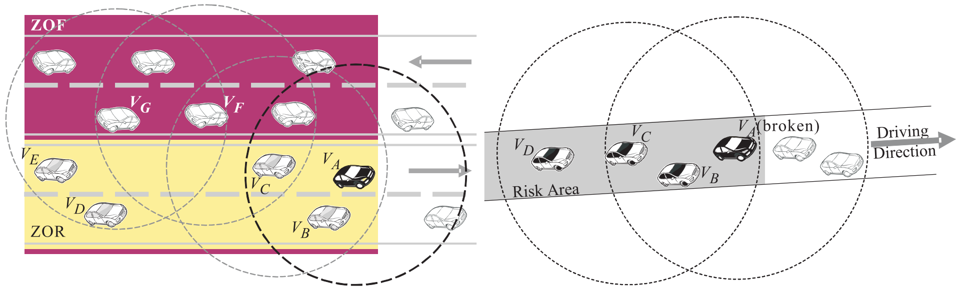
Consequently, how to utilize the driver behavior, along with consideration of density variability and unreliable transmission, will possibly be the next challenge in the design of min-delay and delay-bounded unicast routing protocols. In the next section, the different category of routing protocols, multicast and geocast routing protocols, are discussed.

3. MULTICAST AND GEOCAST ROUTING PROTOCOL

Multicast and geocast routing are the other important routing operations in VANETs. One of the challenges is how to develop the efficient multicast and geocast protocol over VANETs with the highly changeable topology. Some results [13][14][15] have recently investigated the multicast and geocast protocols in a VANET. According to the property of geographic region, existing results can be classified into multicast/geocast protocol and spatiotemporary multicast/geocast routing protocols. This section reviews the existing results for VANETs as follows.

3.1 Distributed Robust Geocast Multicast Routing Protocol

Joshi *et al*. [13] had proposed a distributed robust geocast protocol for inter-vehicle communication. The goal of distributed robust geocast multicast routing protocol is to deliver



(a) (b)
Fig. 9. (a) DRG multicast routing protocol and (b) IVG multicast protocol

packets to vehicles located in a specific static geographic region. A vehicle should receive packets or drop only depended on its current location. If a vehicle is located within this specific geographic region, this vehicle receives packets. Otherwise this vehicle drops packets. The zone of relevance (ZOR) is defined in [13] as a geographic region which vehicles in this region should receive the geocast messages. To enhance the reliability of receiving geocast messages under frequently changeable topology, the zone of forwarding (ZOF) is defined in [13] as the geographic region which vehicles in this region should forward the geocast messages to other vehicles in the ZOR. Notice that, ZOF usually surrounds ZOR to ensure the geocast messages can be delivered to vehicles inside ZOR. A periodic retransmission mechanism is proposed in [13] to overcome the network fragmentation. Fig. 9(a) shows that the temporary network fragmentation problem is overcame with the assistance of vehicles of ZOF, for instance, V_G and V_F.

3.2 Multicast Protocol in Ad Hoc Networks Inter-Vehicle Geocast

Bachir *et al*. [14] proposed a multicast protocol in ad hoc networks inter-vehicle geocast, called IVG protocol [14]. The IVG protocol is used to inform all the vehicles in a highway if any danger is occurred; such as an accident. The *risk* area is determined in terms of driving direction and positioning of vehicles. Vehicles located in the risk area form a multicast group. The multicast group is defined temporarily and dynamically by the location, speed, and driving direction of vehicles. IVG protocol uses periodic broadcasts to overcome temporary network fragmentation for delivering messages to multicast members. The re-broadcast period is calculated based on the maximum vehicle speed. Besides, IVG protocol reduces the hops of delivering message by using the deferring time. A vehicle which has the farthest distance to source vehicle waits for less deferring time to re-broadcast. Fig. 9(b) shows an example for the IVG protocol. Vehicle V_A encounters car-function-failure problem and sends this notification to all vehicles in the risk area. Vehicles V_B, V_C, and V_D form a multicast group since they are located in the risk area. Vehicle V_C is the next hop of V_A since the V_C is farther from V_A than V_B. After vehicle V_C sending out packets, vehicle V_B not forwards packet.

3.3 Spatiotemporary Multicast/Geocast Routing Protocol

The spatiotemporary multicast/geocast routing protocol is a new and very interest routing problem. Unlike ordinary multicast and geocast routing protocols, the spatiotemporary multicast and geocast routing protocol should take the time factor into account. The distinctive feature of this new form of spatiotemporary multicast and geocast routing protocol

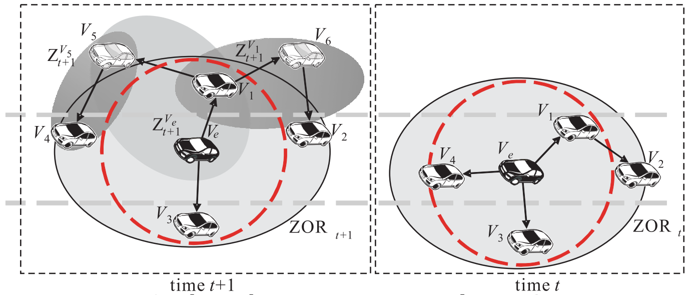


Fig. 10. The mobicast routing protocol in a VANET.

is the delivery of information to all nodes that happen to be in a prescribed region of space at a particular point in time. Chen *et al.* [15] present a "spatiotemporary multicast", called a "mobicast", protocol for supporting applications which require spatiotemporary coordination in VANETs. The spatiotemporary character of a mobicast is to forward a mobicast message to vehicles located in some geographic zone at time t, where the geographic zone is denoted as zone of relevance (ZOR$_t$). Vehicles located in ZOR$_t$ at the time t should receive the mobicast message. Many interesting and useful applications on VANETs can be supported by mobicast routing protocol, such as emergency event, online game, and video advertisement.

To ensure the mobicast message can be sent to all vehicles in ZOR$_t$, vehicles located in ZOR$_t$ at the time t must keep the connectivity to maintain the real-time data communication between all vehicles in ZOR$_t$. The connectivity of ZOR$_t$ is lost if any vehicle in ZOR$_t$ suddenly accelerates or decelerates its velocity. The temporary network fragmentation problem is occurred such that vehicle in ZOR$_t$ cannot successfully receive the mobicast messages. To solve the problem, Chen *et al.* proposed a new mobicast protocol [15] to successfully disseminate mobicast messages to all vehicles in ZOR$_t$ via a special geographic zone, called as zone of forwarding (ZOF$_t$). ZOF$_t$ indicates which vehicle should forward the mobicast message to other vehicles located in the ZOR$_t$. All vehicles in ZOF$_t$ must forward the received mobicast message; even those vehicles are not located in ZOR$_t$. Normally, the size of ZOF$_t$ may be larger or smaller than the optimal size of ZOF$_t$. If the size of ZOF$_t$ is larger than the optimal size of ZOF$_t$, some irrelevant vehicles are asked to uselessly forward the mobicast message. If the size of ZOF$_t$ is smaller than the optimal size of ZOF$_t$, the temporary network fragmentation problem is incompletely overcome. Observe that, the size of ZOF$_t$ is difficult to predict and determined under the high speed environment, such that it easily wastes the network resources. Therefore, the mobicast routing protocol [15] is proposed to dynamically estimate the accurate ZOF$_t$ by zone of approaching (ZOA$_t^{V_i}$ or Z$_t^{V_i}$) to successfully disseminate mobicast messages to all vehicles in ZOR$_t$.

ZOA$_t^{V_i}$ or Z$_t^{V_i}$ is an elliptic zone of approaching to forward the mobicast message more closed to a destined vehicle and Z$_t^{V_i}$ is initiated by vehicle V_i at time t. Any vehicle in the Z$_t^{V_i}$ has the responsibility of forwarding the mobicast message sent from vehicle V_e. Z$_t^{V_i}$ bounds the mobicast message propagation, vehicles in the Z$_t^{V_i}$ can only forward the mobicast message to other vehicles located in the Z$_t^{V_i}$. If a vehicle cannot successfully forward the mobicast message to any neighbor vehicle in the Z$_t^{V_i}$ which is more closed to the destined vehicle, a new approaching zone is initiated. Multiple Z$_t^{V_i}$ are initiated to forward the mobicast message, such that ZOF$_t$ is finally formed by all initiated Z$_t^{V_i}$. Fig. 10 illustrates at time t, V_1, V_2, V_3, and V_4 are located in ZOR$_t$ and receive the mobicast

Table 2. Comparison of multicast and geocast routing protocols

Protocols	DRG	IVG	Mobicast Routing
Mobility of destined zone	no	no	yes
Digital map needed	no	no	no
Realistic traffic flow	no	no	no
Spatial relevance	yes	yes	yes
Spatiotemporary relevance	no	no	yes
Environment	highway	highway	highway

message from vehicle V_e. At time $t+1$, V_2 and V_4 can not directly receive the mobicast messages due to temporary network fragmentation problem. At time $t+1$, V_e, V_5, and V_1 initiate $\mathrm{Z}_{t+1}^{V_e}$, $\mathrm{Z}_{t+1}^{V_5}$, and $\mathrm{Z}_{t+1}^{V_1}$ to forward the mobicast messages to vehicles V_4 and V_2. In this case, $\mathrm{ZOF}_{t+1} = \mathrm{ZOR}_{t+1} \cup \mathrm{Z}_{t+1}^{V_1} \cup \mathrm{Z}_{t+1}^{V_e} \cup \mathrm{Z}_{t+1}^{V_5}$.

3.4 Challenges and Future Perspectives

The surveyed protocols investigate how to provide multicast and geocast routing. Each of the protocols listed above in this survey has its strong point and the detailed analytical comparison is shown in Table 2. Mobility of destined zone shows that a protocol disseminates packets to a static or mobile multicast/geocast region. Existing protocols consider the static multicast/geocast region except mobicast routing protocol. Spatial relevance expresses that location of a vehicle is the factor whether this vehicle decides to receive packet. Spatiotemporary relevance expresses that both the time and location of a vehicle are the factors whether this vehicle decides to receive packet. Existing protocols investigate the single source multicast and geocast routing. However, the multi-source multicast and geocast routing are also an important issue. Therefore, some future perspectives should include the following:

1. The multi-source multicast and geocast routing are worth to develop since the multimedia services are welcome today. A multi-source multicast/geocast routing is the one that each member can be the source of message sender of the other members.

2. Reliability should be considered to design the multicast and geocast protocols. The multicast/geocast message should be delivered with high successful rate.

3. Multicast and geocast routing for comfort applications are also considered. Comfort messages are usually tolerant of delay, meanwhile, network bandwidth is generally reserved for emergency messages. It is worth to develop an efficient multicast/geocast routing protocol for comfort applications with delay-constraint and delay-tolerant capabilities with low bandwidth utilization.

4. The content of multicast/geocast message may affect driver's behavior and could change the network topology. Therefore, the relationship between the content of multicast/geocast message and network topology should be considered for protocol designing.

5. The protocol designing for multicast and geocast routing should consider the scalability. The VANET in a city environment is a potentially large-scale network.

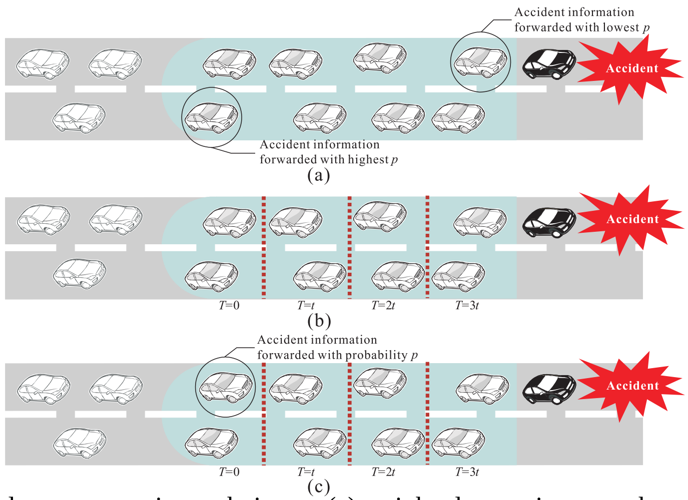


Fig. 11 Broadcast suppression techniques: (a) weighted p-persistence scheme, (b) slotted 1-persistence scheme, and (c) slotted p-persistence scheme

Therefore, future works should focus on developing a multi-source multicast and geocast routing protocol which supports applications of multiple multicast and geocast sources in a VANET. In the next section, the broadcast routing protocol is studied.

4. BROADCAST ROUTING PROTOCOL

Broadcast is the last important operation for a vehicle to disseminate a broadcast message to all the others in a VANET. This section describes existing broadcast routing protocols in VANETs as follows.

4.1 On the Broadcast Storm Problem in Ad Hoc Wireless Networks

Tonguz *et al*. [16] demonstrate that the broadcast storm problem causes serious packet collision and packet loss since too many vehicles simultaneously broadcast messages in a VANET. Tonguz *et al*. [16] thus proposed three distributed broadcast suppression techniques, weighted p-persistence, slotted 1-persistence, and slotted p-persistence schemes. In the weighted p-persistence scheme, if vehicle V_j receives a packet form vehicle V_i, vehicle V_j first checks whether the packet has been received. If vehicle V_j receives this packet at the first time, then vehicle V_j has probability p_{ij} to re-broadcast the packet. Otherwise, vehicle V_j drops this packet, under $p_{ij} = \frac{D_{ij}}{R}$, where D_{ij} is the distance between vehicle V_i and V_j, R is the transmission range. Neighbors of vehicle V_i change p_{ij} to 1 to ensure that the message must be broadcasted if they have no received the re-broadcast message after waiting a random time. In the slotted 1-persistence scheme, If vehicle V_j firstly receives this packet from vehicle V_i, then vehicle V_j waits for $T_{S_{ij}}$ time slots, vehicle V_j has probability 1 to re-broadcast the packet, where $T_{S_{ij}} = S_{ij} \times \tau$, where τ is the propagation time for one hop transmission and $S_{ij} = \left\lceil N_s(1 - \frac{D_{ij}}{R}) \right\rceil$ if $D_{ij} \leq R$; otherwise, $S_{ij} = 0$, where N_s is the default number of time-slot. The slotted p-persistence scheme combines the weighted p-persistence and slotted 1-persistence schemes. If vehicle V_j firstly receives

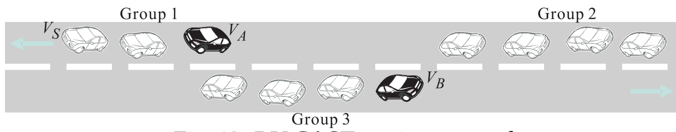


Fig. 12. DV-CAST routing protocol

the packet from V_i, then vehicle V_j waits for $T_{S_{ij}}$ time-slots. Vehicle V_j has probability p_{ij} to re-broadcast the packet. Fig. 11 gives the example of these three broadcast schemes.

4.2 DV-CAST: Broadcasting in VANET

Tonguz *et al*. [17] proposed DV-CAST for a multi-hop broadcast routing protocol in VANETs and indicate three traffic scenarios for a vehicular broadcasting; (1) dense traffic scenario, (2) sparse traffic scenario, and (3) regular traffic scenario. Tonguz *et al*. [17] integrate previously proposed routing solution in [16] to develop DV-CAST which is suitable for both of dense and sparse traffic scenarios and reduces the broadcasting overhead. In DV-CAST, each vehicle monitors the states of neighboring vehicles all the time to make the broadcasting decisions. If a vehicle V_i receives a new broadcast message, V_i firstly checks whether vehicles exist behind. If it is true, the *broadcast suppression* schemes proposed in [16] are adopted to forward the broadcast message; otherwise, V_i forwards the broadcast message via the traffic flow in the opposite direction. After V_i broadcasting message, V_i overhears for a period of time to ensure that the message is successfully broadcasted if the direction of V_i is different from the source vehicle. Fig. 12 shows that the broadcast message is initiated by V_S and it is forwarded from group 1 to group 2. Although groups 1, 2, and 3 are dense group, groups 1 and 2 encounter the temporary network fragmentation problem. Group 1 cannot directly forward packets to group 2. In this case, vehicle V_A can forward packets to group 3 which is in the opposite direction, then vehicle V_B forwards packets to group 2. Observe that, the temporary network fragmentation problem is also considered in the design of broadcasting.

4.3 Broadcast Methods for Inter-vehicle Communications System

Fukuhara *et al*. [18] proposed broadcast methods for inter-vehicle communications system to provide emergency information dissemination in VANETs. The purpose of emergency information is to announce an urgent event by broadcasting for surrounding vehicles. According to the purposes of emergency information, the proposed broadcast methods in [18] are divided into two categories, *emergency-vehicle-approach* information and *traffic accident* information. Emergency-vehicle-approach information is used to announce the urgent event to those vehicles in front of the current vehicle, so the emergency information is only disseminated ahead. Traffic accident information is used to announce the urgent event to those vehicles behind the current vehicle, the emergency information is only disseminated behind. By limiting the broadcast direction, the proposed broadcast methods [18] can provide broadcasts to a particular area and avoid mistakenly notifying other areas where the information is not needed. Fig. 13 shows that vehicle V_A broadcasts the emergency message to the restricted direction. Vehicle V_D does nothing. Vehicle V_B is located in the relay range, it re-broadcasts the emergency information. Vehicle V_C is located in notification range but not in relay range, V_C just receives the emergency information and not to re-broadcast.

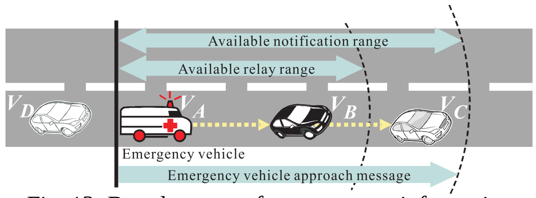


Fig. 13. Broadcast area for emergency information

4.4 Challenges and Future Perspectives

The surveyed protocols investigate how to provide an efficient broadcast routing. Each of the protocols listed above in this survey has its strong point and the detailed analytical comparison is shown in Table 3. Impacted by traffic flow demonstrates whether the protocol performance is affected by changed traffic flow. Fragmentation solution shows how a protocol overcomes the temporary network fragmentation problem. Existing broadcast routing protocols are developing for safety applications to transmit emergency messages; however, there are still some comfort applications which require an efficient broadcast routing protocol, such as public information, advertisements, and navigation information. Therefore, some future works should include the following:

1. A possible future work is how to design an efficient broadcast routing protocol for comfort applications with delay-constraint and delay-tolerant capabilities and low bandwidth utilization. That is, comfort messages are usually not urgent, so it can be delivered under a constrained delay time. Besides, bandwidth should be reserved for safety applications, so an efficient broadcast routing protocol for comfort applications should keep low bandwidth utilization.
2. Broadcast routing protocols for comfort applications should be able to integrate multiple partial comfort messages into a complete message since mass information cannot be completely delivered at once under low bandwidth utilization.
3. A major challenge in protocol design is how to develop reliable broadcast routing protocols for comfort applications to ensure that broadcast messages are successfully disseminated to all the other vehicles in a VANET.
4. The broadcast message should be able to disseminate under low network density. The network density is usually low in off-peak hour; however, the broadcast message is still necessarily disseminated to all vehicles in a network.
5. To provide an efficient broadcast routing, scalability should be considered since a VANET is large and metropolitan-scale network.

Hence, how to design a reliable broadcast routing protocol for comfort applications with a delay-tolerant capability and low bandwidth utilization will possibly be the future works in VANETs.

5. CONCLUSION

Unicast, multicast, and broadcast routing operations are key issues in the network layer for VANETs. This work surveys existing unicast, multicast, and broadcast protocols for

Table 3. Comparison of broadcast routing protocols

Protocols	On The Broadcast Storm Problem	DV-CAST	Broadcast Methods
Impacted by traffic flow	yes	yes	no
Digital map needed	no	no	no
Environment	highway	highway	highway
Road direction	single direction	dual direction	single direction
Fragmentation solution	probability re-broadcast	reverse traffic flow	range restriction

VANETs. The unicast routing protocols are split into min-delay and delay-bound approaches. The min-delay unicast routing protocols construct a minimum-delay routing path as soon as possible. The delay-bound routing protocol utilizes the carry-and-forward technique to minimize the channel utilization within a constrained delay time. This work also surveys important multicast and geocast protocols for VANETs. The multicast in VANETs is defined by delivering multicast packets from a mobile vehicle to all multicast-member vehicles. The geocast in VANETs is defined by delivering geocast packets from a source vehicle to vehicles located in a specific geographic region. A mobicast routing protocol in VANETs is also described. Finally, broadcast protocols in VANETs are also introduced. We predict the tendency of the design of routing protocols for VANETs must be the low communication overhead, the low time cost, and high adjustability for the city, highway, and rural environments.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers and the editors for the valuable suggestions to improve this paper. This research was supported by the National Science Council of the R.O.C. under grant NSC-97-2221-E-305-003-MY3.

REFERENCES

1. A. K. Saha and D. B. Johnson, "Modeling mobility for vehicular ad hoc networks," *ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, Oct. 2004, pp.91-92.

2. H. Safa, H. Artail, and R. Shibli, "An Interoperability Model for Supporting Reliability and Power-Efficient Routing in MANETs," *International Journal of Ad Hoc and Ubiquitous Computing (IJAHUC)*, vol. 4, no. 2, 2009, pp.71-83.

3. T. Sawamura, K. Tanaka, M. Atajanov, N. Matsumoto, and N. Yoshida, "Adaptive Router Promotion and Group Forming in Ad-Hoc Networks," *International Journal of Ad Hoc and Ubiquitous Computing (IJAHUC)*, vol. 3, no. 4, 2008, pp.217-223.

4. R. Manoharan and S. L. P. P. Thambidurai, "Energy Efficient Robust On-Demand Multicast Routing Protocol for MANETs," *International Journal of Ad Hoc and Ubiquitous Computing (IJAHUC)*, vol. 3, no. 2, 2008, pp. 90-98.

5. C. Lochert, M. Mauve, H. Fera, and H. Hartenstein, "Geographic routing in city scenarios," *ACM SIGMOBILE Mobile Computing and Communications*, vol. 9, no. 1, 2005, pp.69-72.

6. J. Zhao and G.Cao, "VADD: Vehicle-Assisted Data Delivery in Vehicular Ad Hoc Networks," *IEEE Computer Communications*, Apr. 2006, pp.1-12.

7. V. Naumov and T. Gross, "Connectivity-Aware Routing (CAR) in Vehicular Ad Hoc Networks," *IEEE International Conference on Computer Communications (INFOCOM)*, May 2007, pp.1919-1927.

8. Y. S. Chen, Y. W. Lin, and C. Y. Pan, "A Diagonal-Intersection-Based Routing Protocol for Urban Vehicular Ad Hoc Networks," *Telecommunication System*, 2009 (under revision).

9. T. Taleb, E. Sakhaee, A. Jamalipour, K. Hashimoto, N. Kato, and Y. Nemoto, "A Stable Routing Protocol to Support ITS Services in VANET Networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, 2007, pp. 3337-33347.

10. W. Sun, H. Yamaguchi, K. Yukimasa, and S. Kusumoto, "GVGrid: A QoS Routing Protocol for Vehicular Ad Hoc Networks," *IEEE International Workshop on Quality of Service (IWQoS)*, June 2006, pp. 130-139.

11. S. Wan, J. Tang, and R. S. Wolff, "Reliable Routing for Roadside to Vehicle Communications in Rural Areas," *IEEE International Conference on Communications (ICC)*, May 2008, pp.3017-3021.

12. A. Skordylis and N. Trigoni, "Delay-Bounded Routing in Vehicular Ad-Hoc Networks," *ACM International Symposium on Mobile Ad hoc Networking and Computing*, Nov. 2008, pp.3017-3021.

13. H. P. Joshi, M. Sichitiu, and M. Kihl, "Distributed Robust Geocast Multicast Routing for Inter-Vehicle Communication," *WEIRD Workshop on WiMax, Wireless and Mobility*, May 2007.

14. A. Bachir and A. Benslimane, "A Multicast Protocol in Ad hoc Networks Inter-Vehicle Geocast," *IEEE Semiannual Vehicular Technology Conference*, vol. 4, Apr. 2003, pp.2456-2460.

15. Y. Chen, Y. W. Lina, and S. L. Lee, "A Mobicast Routing Protocol for Vehicular Ad Hoc Networks," *ACM/Springer Mobile Networks and Applications (MONET)*, 2009.

16. O. K. Tonguz, N. Wisitpongphan, J. S. Parikh, F. Bai, P. Mudalige, and V. K. Sadekar, "On the Broadcast Storm Problem in Ad hoc Wireless Networks," *International Conference on Broadband Communications, Networks and Systems*, Oct. 2006, pp.1-11.

17. O. Tonguz, N. Wisitpongphan, F. Bai, P. Mudalige, and V. Sadekar, "Broadcasting in VANET," *Mobile Networking for Vehicular Environments*, Mar. 2007, pp.7-12.

18. T. Fukuhara, T. Warabino, T. Ohseki, K. Saito, K. Sugiyama, T. Nishida, and K. Eguchi, "Broadcast Methods for Inter-Vehicle Communications System," *IEEE Wireless Communications and Networking Conference*, vol. 4, Mar. 2005, pp.2252-2257.

19. A. Goel, K. G. Ramakrishnan, D. Kataria, and D. Logothetis, "Efficient Computation of Delay-Sensitive Routes from One Source to All Destinations," *IEEE Conference on Computer Communications (INFOCOM)*, Apr. 2001, pp.854-858.