

Review Article

Road-Aware Routing Strategies for Vehicular Ad Hoc Networks: Characteristics and Comparisons

Kashif Naseer Qureshi,¹ Abdul Hanan Abdullah,¹ Jaime Lloret,² and Ayman Altameem³

¹*Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Malaysia*

²*Universidad Politecnica de Valencia, Spain*

³*College of Applied Studies and Community Services, King Saud University (KSU), Riyadh, Saudi Arabia*

Correspondence should be addressed to Kashif Naseer Qureshi; kashifnq@gmail.com

Received 16 November 2015; Accepted 1 February 2016

Academic Editor: Khalil El-Khatib

Copyright © 2016 Kashif Naseer Qureshi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The tremendous advancement in new wireless technologies has led to a renewed interest in vehicular ad hoc networks. However, due to the high mobility and dynamic nature of vehicular networks, great challenges exist for data delivery. In order to design efficient and smart routing strategies for stable communication, several different types of routing protocols have been integrated. One efficient type is the road-aware routing protocol which is based on road statistics (e.g., traffic density, intersections, and road segments). The main objective of the survey conducted in this paper is to discuss the recent road-aware routing protocols and categorize them according to different aspects. In addition, this review investigates the protocols in terms of their forwarding processes, routing metrics, recovery mechanisms, and performance. Moreover, routing challenges and comparisons are discussed.

1. Introduction

Vehicular ad hoc networks (VANETs) are an emerging class of mobile ad hoc networks (MANETs) for pervasive wireless communication among vehicles. In VANETs, the vehicles communicate and coordinate among themselves through pure ad hoc (vehicle-to-vehicle) and fixed infrastructure modes. Vehicular networks have received intensive attention due to the wide range of potential safety and comfort applications and services in the transportation sector [1]. These applications assist drivers to be alert to any critical situation (such as a car accident) and provide precrash sensing, unseen obstacles, and collision warnings for improved passenger safety and the alleviation of traffic flow. Besides safety applications, there are a plethora of comfort applications that provide entertainment to passengers such as Internet access, e-commerce, and weather information [2, 3].

The growing interest in this field is an incentive for researchers, car industry stakeholders, engineers, and governments to create effective solutions and platforms for vehicular communication. Being a subclass of MANETs, VANETs have some unique characteristics such as mobility

patterns, high vehicle velocity, dynamic topologies, and frequent link disruptions [4]. These characteristics generate particular challenges and research issues for handling multihop communication in a network. The networks suffer from frequently broken routes, frequent packet dropping, and high overhead due to failure notifications and route repairs, leading to high transmission delays and low data delivery ratios [5, 6].

To handle multihop routing in VANETs, a network requires a robust and efficient routing solution to handle the special characteristics of the network. Routing strategies for multihop communication have gained researchers' attention and are classified into different types. The most effective type is geographical routing which can handle restricted and bidirectional vehicle movements along roads. Geographical routing protocols use location information for data delivery and are considered to be a promising type of routing to handle scalability and robustness when dealing with the dynamic topologies in a network [7–9]. In these protocols, digital maps, global positioning system (GPS) receivers, and navigation systems are used to find the geographical locations of

vehicle nodes [10]. However, with frequent network partitions, limited link lifetimes, and high vehicle velocity, providing accurate traffic information remains a challenge [6, 11].

Many geographical routing protocols have been proposed in recent research with different types of routing metrics. Most of the metrics are related to the road and vehicle information in a network such as the traffic density, velocity, road map, vehicle direction, and road segments. These types of protocols are called road-aware routing (RAR) protocols due to their network and road awareness characteristics. These protocols have gained attention due to their enhanced performance and efficient routing decisions in a network. Basically, RAR protocols use intervehicle connectivity, traffic density, and junctions for their routing decisions.

The RAR schemes are viable as they can work with traffic awareness and network variables. These protocols determine the network links' reliability, analyze the road and vehicle information, make routing decisions among vehicles, and avoid poor communication quality and less dense networks. However, despite many advantages, the RAR protocols present several open research challenges and issues. One of the significant challenges is the network and traffic status evaluation in terms of accurate and up-to-date information about the vehicles and networks. The review conducted in this paper discusses the RAR protocols and highlights the limitations, challenges, and features of recent RAR protocols. To the best of the authors' knowledge, this survey is the first to review RAR protocols in the VANET environment. The main objectives of this paper are as follows:

- (1) To discuss the most recent RAR protocols in the field of VANET and explore their operations and processes with a critical analysis of the routing metrics.
- (2) To highlight the challenges and limitations of the protocols.
- (3) To summarize the protocols' critical issues that should be taken into account in the design of robust routing schemes for networks and to recommend future directions for further research.

The remainder of the paper is structured as follows: Section 2 provides an overview of vehicular routing protocols and discusses the most effective type of routing. Section 3 presents a classification of RAR protocols and discusses each of the categories in detail. Section 4 discusses the challenges and issues in RAR protocols through a comparison and an investigation of their limitations. Section 5 identifies open research issues and future research directions. Section 6 concludes the work.

2. Overview of Vehicular Ad Hoc Network Routing Protocols

VANETs require efficient routing in the presence of environmental obstacles and dynamic and unpredictable topologies with high network mobility to maximize throughput, control overhead, and minimize packet loss. The vehicle nodes are dynamic in nature, and finding an optimal route to the

destination is still a challenge. Routing strategies are defined based on the application requirements and are distinguished based on architecture. Basically, vehicular architecture is categorized into three types: pure ad hoc, infrastructure-based, and hybrid. The traditional mobile ad hoc routing protocols are less able to handle vehicular characteristics [12]. Routing protocols are categorized into different types such as topology, geographical, cluster, geo-cast, and multicast routing protocols. Most routing protocols belong to the two main categories of topology and geographical routing. In topology-based routing, the vehicle nodes store the information about the network topology and depend on static end-to-end routes. Thus, these protocols suffer from high maintenance (in terms of communication and memory) and unnecessary flooding issues. Due to the high communication overhead, high mobility, and high resource cost, the updates and routing requests are outdated [13]. In order to overcome topology-based routing constraints, geographical routing protocols were introduced, whereby the protocols only store information about neighbors.

2.1. Geographical Routing Protocols. Geographical routing protocols eliminate the need for topology information as the nodes only store information about the directly accessible neighbors within transmission range. These protocols rely on the position information of vehicle nodes to forward the data toward the destination. The position information is known through GPS or periodic beaconing messages. Recently, vehicle manufacturing companies have begun to offer GPS services with digital maps in their cars, which are important for geographical positions for routing. Geographical routing has been identified as a more reliable and efficient approach for VANETs due to the development of location services [14]. In order to find the neighbor nodes in a network, each vehicle maintains and updates a neighbor table via the exchange of beacons or hello messages with its neighbors.

Greedy perimeter stateless routing (GPSR) [15] is a basic geographical routing protocol. The protocol uses two working modes, namely, the greedy and perimeter modes, to forward the data toward the destination. The greedy mode sends data to the node that is closest to the destination but may encounter local optimum or maximum issues. Local optimum or maximum issues occur when the forwarder vehicle node is near the destination node but its neighbors and the destination node are not accessible through one-hop communication. In such scenarios, the protocol uses the perimeter or recovery mode which routes the data packets through the right-hand rule and selects the next forwarder neighbor. According to the right-hand rule, whenever a source node receives a data packet from an edge node, it sends the data packet to its next edge counterclockwise. Afterward, the perimeter mode is switched to greedy mode; otherwise, the packet forwarding will continue with the perimeter mode to forward the data towards the destination. The protocol may not perform well in an uneven traffic distribution environment with high mobility and different road segments in the network. This is because the protocol activates face routing and the data are forwarded through a number of faces toward the destination.

In order to address the limitations of geographical routing, researchers have proposed new routing metrics which involve road awareness in routing decisions. The road awareness metrics are integrated with geographical routing protocols for efficient data transmission in a network by adopting information such as road density, vehicle direction, distance, intersections, digital map, and road segments. The next section discusses the concept and classification of road awareness-based protocols in detail.

3. Concept of Road Awareness and Classification of Road-Aware Routing Protocols

This section discusses the most recent road-aware geographical routing protocols and their performance and critical issues in VANETs. Geographical routing protocols have been recognized as the most reliable and effective approach for vehicular networks [16, 17]. The performance and capabilities of simple geographical routing protocols are limited in VANETs due to the special characteristics of vehicular networks. In vehicular networks, the moving nodes are restricted by roads and high and low traffic density affects the transmission lifetime and lead to disconnectivity and network overhead [18]. Real-time road information is very significant in order to check the traffic congestion for routing decisions. The RAR protocols utilize the road characteristics and network status such as intersections, traffic density, vehicle distribution, communication load, link qualities, distance, and direction. Therefore, RAR provides more feasible geographical information for routing decisions in a network based on road awareness metrics.

3.1. Overview of Road-Aware Routing Protocols. Road-aware geographical routing protocols utilize different road measurements for packet forwarding and routing decisions. These protocols utilize the intersections, traffic density, link quality, distance, and direction for data forwarding and, in the case of a link disconnection, a recovery strategy is adopted. These protocols have three main processes:

- (1) Calculate the routing metrics and recalculate or update the routes in new conditions.
- (2) Measure the routing metrics for realistic indications about road conditions and efficient routing.
- (3) Adopt greedy forwarding to forward the data packet and, in the case where there is no neighbor node close to the destination, adopt a recovery strategy.

In the first process, these protocols define the routing metrics in order to calculate the routes with the help of network conditions. The information about the routing metrics can be recalculated and updated. The routing metrics are used to check the traffic status and network conditions. In the second process, the protocols measure the routing metrics for realistic traffic conditions. Afterwards, the packets are forwarded to the next intersection or hop by a forwarding mechanism. The simplest forwarding method is the greedy

approach, which sends the packets to the neighbor node that is closest to the destination. However, the simple greedy approach has some drawbacks. In order to address the greedy approach issues, new forwarding mechanisms have been proposed to add distance, direction, and velocity metrics [19]. These greedy protocol enhancements still face packet forwarding problems but with lower probability. In the third process, if a neighbor node near the destination is not available, these protocols switch to a recovery strategy.

Different types of recovery strategies have been adopted in geographical routing protocols, such as the right-hand rule, carry-and-forward, and recalculation of the forwarding path. The right-hand rule states that, after receiving a packet from the edge node, the node will forward the packet to the next counterclockwise edge node. Whenever a forwarding node close to the destination is available, the protocol switches back to forwarding mode. This approach encounters the issue of looping due to the rapid exchange of constructed graphs in the network. In the case of the local maximum issue, the carry-and-forward approach is adopted whereby the node carries the packet until it finds an optimal candidate node in the network. The carry-and-forward approach still suffers long delay issues in a network [20]. The recovery strategy of recalculating the forwarding paths also leads to network delay and high overhead.

3.2. Routing Strategies. This section discusses the routing process of RAR protocols. RAR protocols are designed with different routing strategies which are mainly classified into five categories: full RAR, intersection-aware routing, twofold decision routing, traffic density-based routing, and source-based routing. A taxonomy of road-aware protocols, including the relevant routing strategy and year of publication, is shown in Figure 1. The features of each routing strategy are discussed in the following subsections.

3.2.1. Full Road-Aware Routing. In the full RAR strategy, the protocol takes full road awareness in order to traverse data packets toward the destination. These protocols have two main mechanisms for data forwarding: find the shortest path and broadcast the data packets to find the best path toward the destination. The road map is converted through the Dijkstra algorithm into a weighted graph, wherein the junctions and road edges are graph vertices. The road edges are weighted based on road lengths and traffic density. Some other RAR metrics also determine the data forwarding such as the distance of destination, direction of node, link quality, and transmission delay. After checking the road awareness through the weighted graph, positions, and intersections, the road awareness information is added in each packet header before forwarding. These protocols are efficient due to multiple levels of road awareness and are suitable for the vehicular environment. On the other hand, these protocols have some limitations and challenges. One of the main limitations relates to saving the road information in each packet such as the information about the intersections and the positions of anchors. In a dynamic vehicular environment, the lifetime of the path is inversely proportional to its length [8]. When each vehicular node takes the decision, it is static

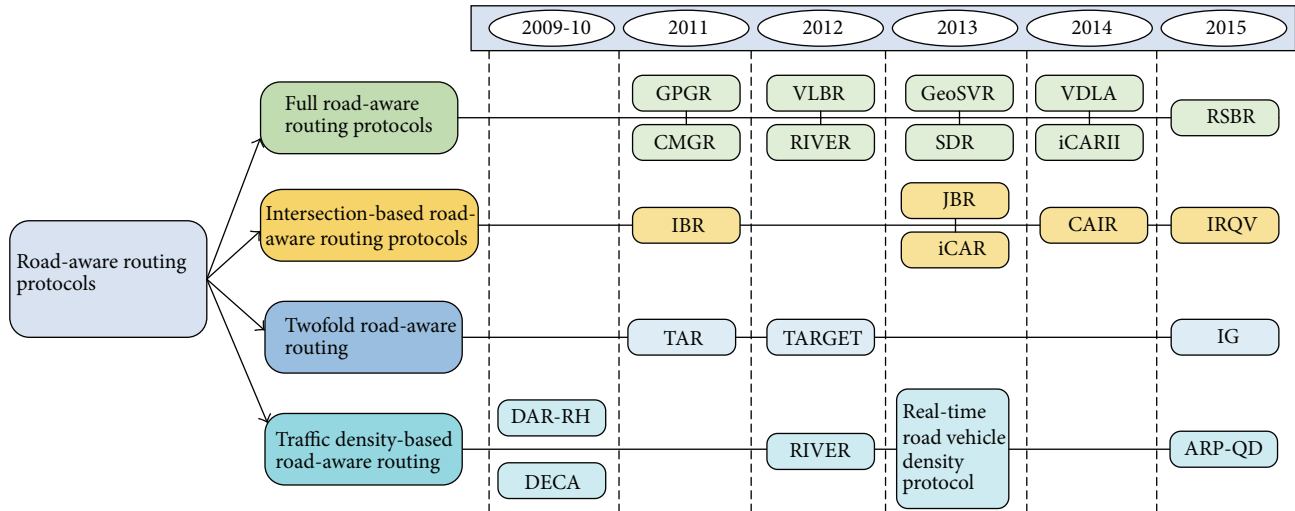


FIGURE 1: Taxonomy of road-aware routing. Notes: ARP-QD = adoptive routing protocol based on QoS and traffic density; CMGR = connectivity-aware minimum delay geographic routing; CAIR = connectivity-aware intersection-based routing; DECA = density-aware reliable broadcasting protocol; DAR-RH = density-aware routing using road hierarchy; GPGR = grid-based predictive geographical routing; GeoSVR = Geographic stateless VANET routing; IBR = Intersection-based routing; IG = improved geographical; iCARII = connectivity-aware routing protocol; IRQV = intersection-based routing with quality of services; JBR = junction-based routing; RIVER = reliable intervehicular routing; RSBR = road selection-based routing protocol; SDR = stable direction-based routing; TAR = traffic-aware routing; TARGET = traffic-aware geographical routing; VLBR = VANET load balancing routing; VDLA = traffic density and load-aware routing.

for each packet and cannot be adjusted in line with the traffic conditions. To determine the complete road awareness metrics for establishing the route, the full RAR protocols face network overhead issues especially for recalculating the path. By applying the Dijkstra algorithm on a large city map, these protocols have computational complexity issues.

3.2.2. Intersection-Aware Routing. In intersection-aware routing, the candidate node dynamically selects the next intersection and uses other feasible routing metrics. Thus, the candidate node selects the next intersection for adjacent roads and selects the optimal route for packet forwarding. The protocol determines the metrics at each intersection based on road awareness and traffic situations. These protocols have less network overhead compared to full RAR protocols, as the data packets only carry the next intersection and destination positions rather than full road information. These protocols adopt the unique characteristics of vehicular networks for routing decisions. Despite the low overhead, these protocols still face the delay time issue due to the selection of the intersections one by one [20]. Further, keeping and updating the routing metrics for the routing decisions are difficult. In addition, these protocols initiate the measurement process frequently and this is the cause of low data delivery and high network overhead.

3.2.3. Twofold Decision Routing. Twofold decision routing addresses the shortcomings of the full RAR and intersection-aware routing by determining different metrics between intersections and at intersections. This type of routing is suitable for dense and spare urban environments, where the traffic situation is rapidly changing. Thus, these routing

strategies have less network delay due to the processing of different metrics and decisions. The selection of intersections for routing may lead to the disconnectivity issue due to partitioning in the network and also faces network overhead issues. With twofold strategies, the network overhead increases to add simple direction and forward progress metrics between intersections. The network overhead also increases when the candidate node is at an intersection and the node determines different road metrics. To design twofold routing, it is important that protocol adopts more realistic routing metrics for routing.

3.2.4. Traffic Density-Based Routing. Traffic density estimation for routing is an important metric for real-time applications in vehicular networks. Traffic density-based routing protocols determine traffic density by attaining density feedback from the network and forwarding the data packets with low and high density routes. These protocols are suitable for an urban environment where the movement of vehicles is affected by driving instructions and obstacles that cause disconnection in wireless signals. The communication between vehicles must be line of sight, and movements are limited to restrictions. To address these issues, these protocols use the density metric to find the most connected path between the source and destination.

4. Full Road-Aware Routing

This section discusses recent full road awareness-based routing protocols. The protocols' metrics, functions, and forwarding mechanisms are highlighted, along with the open research issues and challenges. Table 1 presents a comparison

TABLE 1: Full road-aware routing protocols.

Protocol, year	GPGR, 2011 [21]	CMGR, 2011 [22]	VLBR, 2012 [24]	RIVER, 2012 [25]	GeoSVR 2013 [26]	SDR, 2013 [27]	VDLA, 2014 [16]	iCARII, 2014 [28]	RSBR, 2015 [29]
Forwarding approach	Selecting the neighbor closest to the destination within its transmission range	Greedy forwarding	Using sparse network for furthest neighbor and using density with least network load	Greedy forwarding up to near anchor point	Restricted forwarding approach	Directional forwarding along the road segment	Greedy forwarding	Greedy forwarding	Greedy forwarding
Routing metrics	Road topology, position, direction, and velocity	Vehicular density and route trip time	Network load, traffic density, and distance	Road topology, network connectivity, length of path, and distance	Vehicle density and distance	Direction, relative velocity and transmission range	Traffic density, traffic load, and distance	Network connectivity, expected lifetime, and delivery delay	Distance, velocity, and transmission range
Routing metrics measurement	Using beacon messages containing the grid coordinates	Measures through route discovery and replying to messages	Determining traffic density and network load in header as an indicator	Using weighting values, where a small value shows the high edge reliability, a large value shows the unreliable edge, and maximum value shows the untraversable path	Measuring higher density roads	Determining the direction, relative velocity, and transmission range	Determining traffic density, load, and distance metrics through NICP in hello packets	Performing road segment evaluation by sending unicast packet	Vehicle position estimation
Load balancing	No	No	Yes	No	No	No	Yes	No	No
Recovery mechanism	No	Carry-and-forward mechanism switching the path for load balancing	Carry-and-forward mechanism and switching the path for load balancing	Through route recalculation	N/A	Recovery through RERR packets	Recovery through recalculating the traffic density	Sending new path request to location centers for recovery	Recovery through speed adjustment method
Route selection	Source-based	Source-based	Source-based	Source-based	Source-based	Source-based	Source-based	Source-based	Source-based
Traffic awareness	Shortest path	Relying on broadcast area	Using k -shortest paths	Using Dijkstra algorithm for most reliable path	Subgraph between source and destination	Relying on broadcast area	Minimum hop and shortest path	Shortest path	Shortest path
Parameters for shortest path selection	Position, direction, and velocity	N/A	Distance and statistical density	Network connectivity and road topology	Road length and width	Transmission range	Traffic density, load, and distance	Network connectivity	Minimum road rating
Protocols used for comparison	GPSR, GPCR, and GPUR	VADD and A-STAR	GPSR	STAR and GPSR	AODV and GPSR	ROMSGP, AODV, EARP, and D-LAR	Gpsr+ and GPCR	GyTAR, GPSR	P-GEDIR, GyTAR, A-STAR, and GSR

TABLE I: Continued.

Protocol, year	GPGR, 2011 [21]	CMGR, 2011 [22]	VLBR, 2012 [24]	RIVER, 2012 [25]	GeoSVR 2013 [26]	SDR, 2013 [27]	VDLA, 2014 [16]	iCARII, 2014 [28]	RSBR, 2015 [29]
Route update	Recalculating the grid coordinates calculation when road topology changes	Route update by route discovery messages	Update with k -shortest paths for balancing network load	Using beacon messages and recalculating the forwarding path at anchor points	No route maintenance	Through control messages	Through enhancing hello messages	Through road side units broadcasting	Route update by beacon messages
Packet delivery ratio	Higher	Higher	Higher	Higher	Higher	Higher	Higher	Higher	Not measured
End-to-end delay	Not measured	Higher with low traffic density	Lower	Higher	Lower with AODV and higher with GPSR	Lower	Lower	Lower	Lower
Network overhead	Not measured	Not measured	Not measured	Higher	Not measured	Smaller	Not measured	Not measured	Not measured

of the protocols in terms of their routing metrics, forwarding and recovery methods, and performance.

4.1. Routing Process of Full Road-Aware Routing. The concept behind full road awareness is to determine the road characteristics for optimal routing in a network. The first protocol in this category is grid-based predictive geographical routing (GPGR) [21] which is based on the prediction mechanism of moving vehicles' positions and mapping the data to generate the road grids. The design assumes that vehicles move only along the road grids and are equipped with location services (i.e., GPS) and a preloaded digital map for road information. The location services also provide traffic statistics such as direction, velocity, and road topology. The protocol partitions a geographical area into two-dimensional logical grids, where a located vehicle at a position in a grid is capable of transmitting the data to its eight neighbors' grids. The beacon messages contain grid coordinates rather than the vehicles' positions. By the floor function, every vehicle computes its current grid coordinates. The sender node selects the nearer node with destination and predicts the future position by road grids. To check the distance between two grids, the protocol uses the Euclidean distance formula. Afterwards, the position of the next relay candidate is set with the velocity and direction of the relay candidate.

The connectivity-aware minimum delay geographic routing (CMGR) protocol [22] requires high connectivity for route selection in a sparse network. In dense situations, the protocol regulates the routes by suitable connectivity and selects the route with minimum delay. The protocol assumes the availability of gateways which are randomly distributed along the roadside and connected to the IPv6 network. The route discovery messages are used to construct full paths toward the gateways and intermediate vehicles attach their locations and rebroadcast the discovery messages. The gateway receives several messages from the same vehicle node via different routes and then selects the most appropriate route after the connectivity evaluation. In addition, for every route discovery message, the gateways send the receiving messages to the source vehicle and then the decision is taken by the source node about the suitable route based on the connectivity evaluation. Further, the CMGR protocol continuously checks the routes' quality and most of the data are forwarded from the favorable routes. As a result, there are higher delays, congestion, and packet dropping issues in such routes. To eliminate the network bottleneck and enhance the network throughput, a load balancing mechanism is needed [23].

VANET load balancing routing (VLBR), as proposed in [24], aims to balance the traffic between all possible connected paths by achieving congestion feedback from the network and switching to less congested routes by applying the k -shortest path algorithms. The protocol performance was tested with some assumptions, where vehicles are equipped with positioning services and a digital roadmap and roads are dense with traffic. A city map is converted into a directed graph and then the intersection sequence is calculated by the source with each road weighted by its length and vehicular density. The protocol uses three phases. In the first phase,

each source node determines the list of intersections and stores the information in the packet header. In the second phase, the protocol takes the forward decision and is selected as a forwarder node between intersections. In the third phase, the protocol achieves feedback from the network and switches the path in the case of pre-congestion. The frequent path switching approach causes network delay and packet drop issues in the network. Furthermore, the k -shortest path algorithm is not feasible due to its heavy computational process leading to extra load in the network.

Reliable intervehicular routing (RIVER), as proposed in [25], is based on the utilization of an undirected graph that presents the street layout and selects the route by assigning reliable ratings. RIVER determines the neighbors and destination positions by beaconing plus a preloaded street map. The protocol identifies a path and forwards the message along the path, which is connected with a number of geographical locations. The protocol uses the piggybacking approach to distribute the traffic information rather than broadcasting or flooding. Afterward, the protocol uses the real-time information and recalculates the forwarding path at anchor points. The protocol has two types of monitoring, namely, active and passive, whereby the probe messages belong to active traffic monitoring and the receiving messages belong to passive traffic monitoring. The probe messages are used to find out whether or not a message can be delivered to a specific route. The vehicle node near a street vertex sends a probe message to the next hop which is projected on the same vertex. The probe message is received by the node that is closest to another vertex and at the same time is returned to its original source which becomes aware of the connected path. The weak point of this protocol is reliability, as every node assigns edge weight to every known edge by active and passive monitoring messages. The Dijkstra least weight algorithm is used to compute the consistent path. The protocol has computational complexity in terms of the weighting calculations.

The geographic stateless VANET routing (GeoSVR) protocol, as presented in [26], aims to improve packet forwarding and overcome the local maximum and disconnectivity issues in a network. The protocol presents two algorithms, namely, restricted and optimal forwarding algorithms. The restricted forwarding algorithm aims to solve unreliable wireless channel issues, and the optimal forwarding algorithm aims to overcome the local maximum issues. Optimal forwarding determines the path and assigns a weighting value to every road and considers the vehicle density with destination distance in order to avoid sparse connectivity. The protocol selects the shortest path with the minimum weight value through the Dijkstra algorithm and inserts the value into the packet header. After path calculation, the restricted forwarding algorithm is used to select a neighbor node in a restricted range for the next hop selection. This is because interference, communication distance, and weak signal strength issues lead to packet losses. Figure 2 illustrates examples of the optimal and restricted forwarding algorithms. The Dijkstra algorithm requires high computational processes to navigate the complete weighted graph to find the connected path. Therefore, to address this issue, the protocol uses the subtraction method,

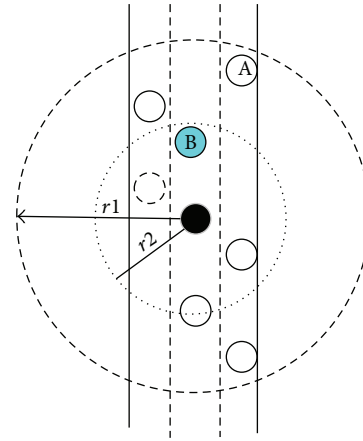
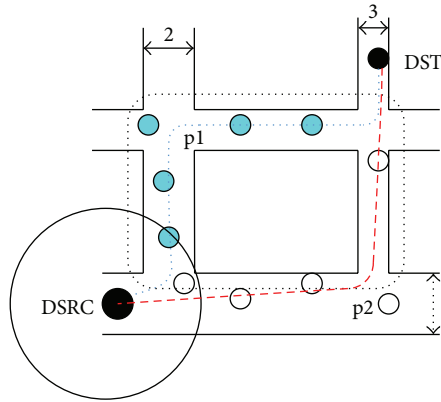


FIGURE 2: Examples of optimal and restricted forwarding in GeoSVR.

where it subtracts a connected subgraph from the entire weighted graph by using the source and destination positions. The protocol determines the junctions that are close to the destination and constructs a rectangular subgraph instead of a complete city map.

In [27], the stable direction-based routing (SDR) protocol based on direction and path duration prediction was proposed. In SDR, a node rebroadcasts the request message based on the neighbors' and destination node's positions. Next, the protocol labels each link with link expiration time and considers the link stability. Finally, the path with the longest expiration time is preferred as the most suitable path. SDR uses an improved flooding approach based on direction angle to decrease the effects of a broadcast storm in the network. After the direction angle is attained, it is compared with the threshold value. If the value is smaller than the threshold, the node will add its own address and rebroadcast the packet. In this way, the protocol reduces the route discovery messages due to broadcasting within direction angle nodes. The protocol also provides route maintenance: when a primary path that is used for routing breaks, the node that notices that this link is broken sends the route error (RERR) packet to the source. Afterwards, the source node selects the next optimal path. SDR reduces the time for finding a new path and avoids the broken paths.

The traffic density and load-aware routing (VDLA) protocol, as proposed in [16], is based on sequential intersection selection to construct a route. The protocol determines the real-time traffic load and density with destination distance and collects all the information through a centralized mechanism. The protocol computes these routing metrics, and the packets avoid the disconnected roads with load balancing in order to mitigate network congestion. The protocol has the ability to reduce unnecessary hops by selecting the intermediate junctions before a packet reaches a junction. For routing, VDLA prefers roads that are well-connected with lower traffic loads. However, the protocol adds the network information collection packet (NICP) which includes a number of nodes and the total length of the buffer queue and total neighbors and transmits it from one end to the other. Figure 3 shows

the process of the VDLA protocol, where Vehicle A activates the NICP. In Road RS 2, the farthest neighbor Node B adds the information and the updated packet is forwarded to Node C. Node C repeats the same method until the packet reaches the destination. The protocol suffers from network overhead due to the addition of the NICP and the exchange of the complete road information in the hello messages.

The connectivity-aware routing protocol (iCARI), as proposed in [28], aims to improve routing performance by increasing delivery delay. The protocol uses the infrastructure facilities and broadcasts one-hop beacon messages to update the driving conditions periodically, with vehicles equipped with GPS-based on-board units and roadside units for Internet access. iCAR has four main components: road segment evaluation, path lifetime calculation, next junction selection, and next hop selection. In road segment evaluation, the protocol dynamically senses the different parts of the network in order to determine the real-time information about the network. For this phase, the protocol broadcasts lightweight control packets to check the traversing roads and intersections. Next, the protocol calculates the lifetime in order to determine the path lifetime and taken decision by the one-hop neighbors' vehicles. After calculating the lifetime, the protocol selects the next junction and determines the existence of the path toward the destination. If the packet reaches a disconnected road due to an unexpected delay, then the current candidate node sends a new path request to the location centers for recovery. In the last step, the protocol selects the next hop by the greedy forwarding approach.

The road selection-based routing protocol (RSBR), as proposed in [29], aims to select the best road by calculating the rating of the roads connected with a junction. The protocol assumes that every vehicle maintains a neighbor table by updating beacon messages. The vehicles are moving in the same direction and maintain their location positions by a GPS service. The protocol solves the network gap problem by selecting the best road by predicting the gaps at an early stage. If vehicles encounter the gaps while forwarding the data, they switch to recovery mode and speed is adjusted according to backward and forward positions. The protocol

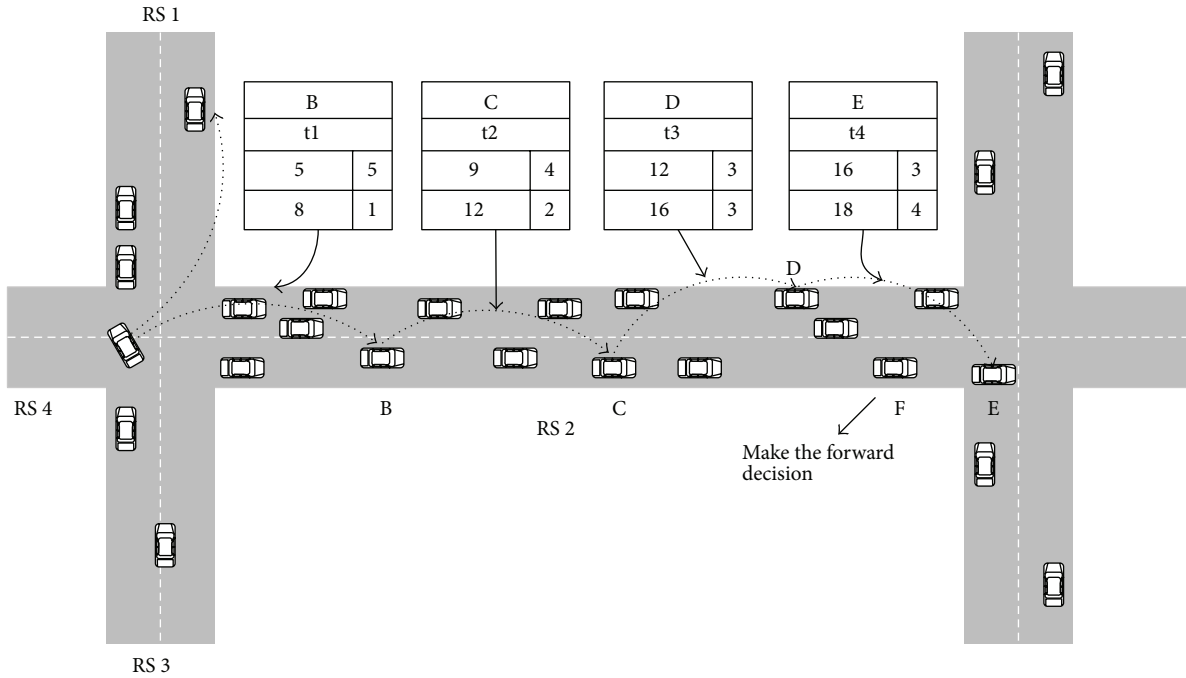


FIGURE 3: Routing process of VDLA with traffic density and traffic load collection.

is evaluated by considering three performance metrics: end-to-end delay, number of hops, and network gap encounters. The protocol performs better in terms of delay compared to other state-of-the-art routing protocols and encounters fewer gaps in the network. The vehicular environment is dynamic and, due to high velocity, the topology changes frequently. The protocol suffers from the network overhead issue due to its network gap detection mechanism.

4.2. Routing Metrics, Forwarding, and Recovery Mechanisms. GPGR utilizes four routing metrics: road topology, position, direction, and velocity. The geographical routing protocols select the node nearest to the destination as a relay node, and this leads to local maximum issues in the network. To solve this issue, GPGR selects the relay vehicle based on a location prediction mechanism for the relay vehicle selection process. The protocol selects the relay vehicle based on vehicle movement information (including velocity, direction, position, and road topology) to forward the data packets toward the destination. The protocol then uses map data to generate a road grid and predict the exact movement positions of vehicles for the selection process. The use of the exact moving positions and the prediction mechanism reduces the link breakage issue in VANETs. The information on the number of nodes and the vehicle velocity is used to evaluate the packet delivery ratio. However, the protocol does not define a mechanism of recovery.

On the other hand, CMGR uses the changing network status to forward the data packets and applies greedy forwarding to forward the packets in the network. The protocol measures the maximum and minimum vehicular density along the road and the connectivity of the routes in order to select an optimal route. The number of neighbors (local density),

maximum and minimum average density rates, and the trip time are the metrics used for evaluation. Furthermore, these metrics differentiate the routes based on information about their density through route discovery packets from the source node to the gateway. The route with the least trip time and highest density is taken into account as a high connectivity route. The changing value of the vehicle density is attached to the route discovery packets before broadcasting. The protocol calculates the routing metrics in two steps as follows:

- (1) For every vehicle in the network, the expected value of the vehicle density change rate is calculated by beacon messages.
- (2) After exchanging the beacon messages, the average expected values of the density are calculated.

After calculating the trip time from the route discovery messages, the gateway computes the difference between the time when the route discovery messages were received and the time when the packets were generated. The protocol adopts the carry-and-forward approach to overcome the network partition issues.

The VLBR aims to balance the traffic by attaining the network load, traffic density, and distance between potential connected paths. The protocol utilizes two forwarding methods, which are based on network conditions. When a road is sparse, the protocol selects the neighbor closest to the next intersection and when a road is dense then the protocol follows the same procedure with less network load compared to other neighbors. The protocol computes the congestion feedback from the network and switches to lower congestion routes through the k -shortest path algorithm. The protocol has the ability to balance the traffic with higher density and least distance cost instead of only choosing the shortest path.

The network load is measured by the contention window size, which increases with each transmission trial. In addition, the collision probability is computed with the help of the current contention window size and previous computed probability. Then, each vehicle node exchanges the computed probability with its neighbors. If all the neighbor nodes have a higher collision probability compared with the specified threshold then it is considered to be a heavy loaded area. However, the authors did not discuss the threshold value calculation or the measurement of the collision probability for all nodes in the case of highly dense roads. The protocol adopts the carry-and-forward approach for recovering the route and also the path switching method for load balancing in the network.

RIVER utilizes greedy forwarding up to the near anchor point and uses a path that connects with a number of geographical locations and the distance to the destination to forward the data packet. The protocol identifies neighbor nodes by reference to information about the road topology, network connectivity, distance, and length of path towards the destination. The protocol determines the estimated network reliability and real-time traffic monitoring based on the active and passive monitoring methods. Active traffic monitoring refers to probe messages that move along a particular street edge to compute the connectivity of the street edge. The real-time active traffic monitoring allows the protocol to avoid routes with coverage area gaps. After active information, the passive information is distributing to other nodes. In medium vehicle density, the protocol's performance is better compared to high density networks due to the computational complexity of calculating the street segments. The protocol uses an alternative route in the case of route failure and performs recovery through route recalculation.

GeoSVR uses the restricted forwarding approach and utilizes vehicle density and distance to destination as routing metrics. However, its authors assumed that the wider road implies higher probability of high density. The optimal forwarding path measures the weighting values of the road edge by the width and length of the road. The vehicular environment is changing rapidly and is not constant without any guarantee of high density in wide roads. By using this protocol, the number of hops in the network is increased and this leads to delay in the network. Moreover, the protocol lacks a recovery mechanism in the case of route failure in the network.

In order to find the best path in the network, SDR uses directional forwarding along the road segment and utilizes three routing metrics: direction, relative velocity, and transmission range. The protocol uses an improved flooding mechanism by combining the position information with the direction angle to reduce the effects of the broadcast storm issue in the network. The protocol then uses mobility information to estimate the link expiration time by reference to the transmission range, position, and movement of vehicles. In the case of the local maximum issue in a network, the protocol uses a recovery approach through sending RERR packets to the source node about the suspected link and selecting the next best path.

The VDLA uses greedy forwarding and utilizes the real-time traffic density, load, and distance for routing decisions.

Information on the vehicle density and traffic load is incorporated by GPS and preloaded digital maps. The traffic load is defined by the centralized method, whereby an information collection packet is initiated by nodes including the number of nodes and total neighbors. After the information is obtained, a weight is calculated by distance and network load to find the shortest route in the network. In the case of local maximum, the protocol utilizes a recovery approach and recalculates the traffic density for the optimal route in the network.

The iCARII protocol uses beacon messages that include current location and mobility vectors to report driving and traffic conditions periodically. The protocol greedily forwards the packets toward the destination. The vehicle node measures road segments by sending unicast packets to take the network connectivity into account and sends the report to the location centers. The collected report is based on the delivery delay of the packet and minimum expected lifetime of the local network connectivity at the road. The protocol has computational complexity due to its delivery delay calculation. In the case of route failure, the protocol has a recovery mechanism, whereby the node sends a new path request to the location centers to recover the route or find another best route for data forwarding.

RSBR utilizes three routing metrics: distance, velocity, and transmission range. The protocol is similar to other greedy protocols with some extra features. Every vehicle updates and maintains its routing tables by beacon messages. The protocol determines the road condition before sending the data until the last junction is reached. Furthermore, the protocol is based on the Dijkstra algorithm and calculates the distance between the source and destination. The vehicular network topologies are changing rapidly and are the cause of outdated data and network delay issues in the network. In a network gap situation, the protocol recovers the route through the speed adjustment method to adjust the speed of vehicles.

4.3. Discussion and Comparison. This section presents a discussion and comparison of the routing protocols explained above. The challenges and open issues are also discussed.

Most of the discussed RAR protocols consider the routes with high connectivity and traffic density. These types of roads are highly congested and create a bottleneck problem in the network. The load balancing property overcomes this issue during the selection of a route toward the destination. The VLBR has load balancing capabilities but still has some limitations. The protocol does not avoid network congestion on the used route and does not switch to another route when all the nodes reach the congestion threshold in the first route. Therefore, some routes might be congested and have bottleneck issues. In addition, the source node does not get any feedback about calculating and classifying the k -shortest paths to update the first path. These limitations may result in packet loss and delay issues in the network.

Most of the VANET routing protocols adopt the Dijkstra algorithm to find the shortest path in the network. This approach may not be feasible without determining the network and traffic status as input. In large city areas, these

protocols need to traverse the complete weighted graph and this leads to high computational complexity. To overcome this issue, GeoSVR uses the subgraph mechanism. On the other hand, only considering the shortest routes in the network and avoiding better network or traffic condition routes are not a better approach. To address this issue, most of the RAR protocols measure the routing metrics first and then look for the shortest paths in the network.

As the RAR protocols use different routing metrics to find the optimal route in the network, assigning equal weighting factors to these routing metrics might not be a suitable method. In addition, further investigation is needed to set the suitable weighting ratio based on different traffic and road conditions. The reliance of some RAR protocols on the traffic density measurement can lead to wrong indications such as some areas being more congested especially near intersections and traffic lights. Therefore, only using the vehicular density measurement is not suitable without vehicular spatial distribution in the network. Taking vehicle density with direction metrics into account will lead to better performance and lower connection breakage probability compared to routes with the same density, as stated in [30]. Some RAR protocols, such as VLBR and iCARI, utilize statistical density. However, statistical density is not suitable due to the highly dynamic environment as stated in [31]. The periodic route discovery in CMGR and frequent route recalculation in VDLA lead to high network overhead and affect the packet delivery ratio in the network. Therefore, the timer or route lifetime is sufficient to overcome unnecessary recalculation.

Many of the full RAR protocols discussed above adopt the carry-and-forward approach for solving the network disconnection issues. However, these mechanisms are applied when a vehicle is far away from the route destination. In CMGR, this approach applies when the vehicles are approaching junctions or turning situations.

In order to overcome RAR protocol issues, some open issues are identified as follows:

- (i) The VANET environment is dynamic in nature and route stability is still an issue. The main question is how to initiate route recalculation only when it is needed in the network and how to determine traffic conditions and overcome routing overhead.
- (ii) In the case of local maximum and network gap issues, the route discovery and rebroadcasting lead to high network overhead and take time.
- (iii) The routing protocols discussed above have been tested in a simulation environment or mobility generator environment. The simulation environment is limited in terms of the number of vehicle nodes and area. On the other hand, the real vehicular environment is denser, and the traffic is denser for the route discovery process.

These challenges open the gate to the further investigation of measures, route discovery mechanisms, and network performance in order to solve the local maximum, broadcast storm, and network gap issues.

5. Intersection-Based Road-Aware Routing

This section presents the most recent intersection-based road-aware routing protocols and discusses their routing operations and forwarding mechanisms as well as the challenges and issues they pose in a network. Table 2 summarizes the characteristics and performance of protocols in comparison with other routing protocols in terms of the end-to-end delay and packet delivery ratio.

5.1. Routing Process of Intersection-Aware Routing. Intersection-based routing (IBR) was introduced in [32] to find the shortest routing path with the shortest packet delay. IBR is based on the greedy approach for straight roads and determines the transmission range as an object of relay. When the source node is at the intersection, the protocol depends on the routing and vehicle moving direction in order to forward the data packet to the destination. Furthermore, in this protocol the vehicles only carry the same direction packets and the protocol has four possible routing conditions: one vehicle node exists in the next road segment, there is no neighbor at the intersection, the vehicle moving direction is different from the packet transfer direction, and the packet arrives at the intersection. The protocol estimates the packet delay of the road segment and the end-to-end delay of the routing path and calculates the forwarding delay with high density.

Junction-based routing (JBR), as proposed in [33], utilizes the selective greedy forwarding approach up to the node that is located at the intersection and is close to the destination. The protocol divides the vehicle nodes into simple and coordinator nodes. In the first phase, when the simple node wants to send the packet, the protocol checks the nodes that are closer to the destination. If the protocol finds the coordinator node, it is chosen as the next hop and, if not, a simple node is selected. When a packet arrives at an intersection, the protocol checks the qualified coordinator nodes and selects the coordinator node that is closest to the destination.

The first version of iCAR was presented in [34] based on the delivery delay for each road and real-time vehicular traffic density. The protocol is unicast and designed for multihop vehicular applications such as file sharing, online advertisements, and chatting applications. In this protocol, when a node is reaching an intersection, the next road is evaluated and selected based on the highest density with least delay. The protocol is efficient for dense urban areas, where it selects the next junction to forward the data packets toward the destination.

In [4], a connectivity-aware intersection-based routing (CAIR) protocol was proposed as an intersection-based protocol that depends on the higher probability of connectivity and lower delay. CAIR selects intersections dynamically by the prediction-based greedy approach to forward data toward the destination. The intersection selection is based on the rectangle area searching method, where the position of source and destination nodes and restricted searching area are bounded by ellipse. The protocol has four main functions: checking the vehicular density to select the road

TABLE 2: Intersection-based road-aware routing protocols.

Protocol, year	IBR, 2011 [32]	JBR, 2013 [33]	iCAR, 2013 [34]	CAIR, 2014 [4]	IRQV, 2015 [35]
Routing metrics	Packet transfer and moving direction of vehicles, transmission range, and traffic density	Transmission range, distance, direction	Distance, vehicular density, network load, and vehicular distribution	Direction, connectivity, and traffic density	Connectivity, transmission delay, direction, and distance
Routing metrics measurement	Commences when vehicle forwarder node leaves road	Commences when vehicle forwarder node leaves road	Initiated when validity period expires	Initiated when validity period expires	Initiated when validity period expires
Load balancing	No	No	No	No	No
Recovery mechanism	Carry-and-forward strategy	Recovery through smaller minimum angle method	Store-carry-and-forward	Store-carry-and-forward	Carry-and-forward mechanism
Route selection	Based on traffic conditions at the intersections	Based on coordinator and simple node	Next-intersection selection strategy	Next-intersection selection through improved greedy forwarding strategy	Next-intersection selection through QoS parameters
Protocols used for comparison	VADD and GyTAR	GPCR	GyTAR and GPSR	GPSR, CAR, and JBR	GSR and CAR
Packet delivery ratio	High	High	High	High	High
End-to-end delay	Low	Low	Low	Low	Low
Network overhead	Low	Not measured	Higher	Higher	Not measured

segment by taking connectivity and delay into account, using directional forwarding to decrease the average transmission delay, performing on-demand position forecasting for the next hop relay selection, and executing a junction-based forwarding strategy.

An intersection-based routing with quality of services (IRQV) was proposed in [35], based on ant colony optimization. The protocol is based on three processes: intersection selection, network exploration with quality of service (QoS), and path selection. The protocol is dynamic in nature and selects the next intersection by adopting the greedy forwarding mechanism to forward the data toward the destination. Furthermore, the protocol does not rely on complete routing paths; it only explores the terminal at an intersection. The author claimed that, by adopting this mechanism, the protocol increases the routing path stability, reduces redundant routing, and relieves channel congestion in highly dense scenarios. In addition, the forward ants (vehicle nodes) of IRQV shorten the network exploration time and decrease the instant traffic information effects.

5.2. Routing Metrics, Forwarding, and Recovery Mechanisms. The IBR protocol is based on the carry-and-forward approach and avoids reverse moving direction packets and checks the direction of vehicles and routing path with traffic and routing delays for each road segment. For forwarding the data in the network, each vehicle maintains its road segment table including the road segment ID, latest update time, and number of vehicles in the segment. The IBR then follows the two road models (straight road and intersection) and adopts greedy forwarding, transmission range, routing direction of packet, and vehicle moving direction in order to forward the data packets toward the destination. The complete forwarding process of IBR is shown in Figure 4. The protocol adopts the carry-and-forward strategy to deal with uneven distribution scenarios.

JBR adopts the direction and transmission range, as in IBR, with the selective greedy forwarding approach. The protocol divides the vehicle nodes into simple and coordinator nodes, checks the node closest to the destination, and forwards the packet accordingly. If local optimum is

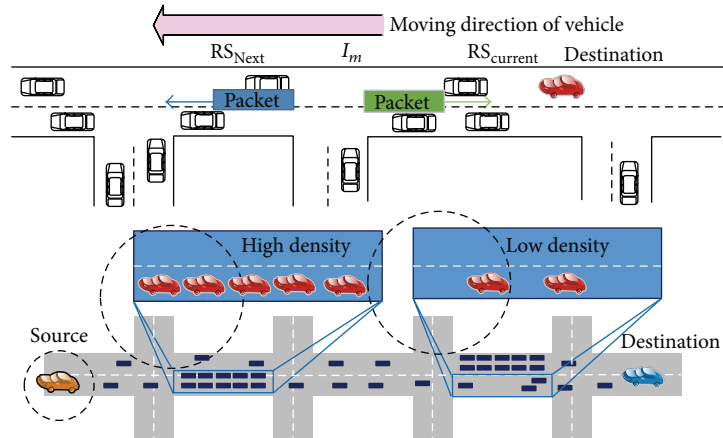


FIGURE 4: Moving direction and high and low traffic density estimations in IBR.

reached, the protocol uses a recovery mechanism based on the minimum angle method until the packet is forwarded to a node that is closer to the destination.

iCAR uses a control packet (CP). It does not form a fixed cell and forwards the CP along the street to forward the packets. In addition, based on the connectivity of the road and number of vehicles, the CP indicates the next intersection. The routing metrics are measured by the validity period of given road segment and predict the time of disconnection. When the CP is received at an intersection, the closer vehicle generates the updated score.

The updated score is then broadcast using the beaoning approach and sent back to the location of the CP; the road score is set with zero value and considered to be disconnected. iCAR adopts an improved greedy forwarding approach to forward the packets between intersections and also looks at the highest received signal strength indication.

CAIR uses the direction, connectivity, and traffic density to find a robust route in the urban environment. The protocol determines the rectangle restricted area, where the source and destination nodes are in line with the axis and an ellipse shape could be formed. In this context, each intersection is involved in the routing process and the routing path could be determined. In addition, if the protocol does not find the path, the route discovery process is initiated and the source node sends the request packets (RREQ) within the search area. The RREQ packets contain traffic density, road length, number of lanes, and intersection information and set up a timer and store this route in a cache. In addition, when the timer expires, the destination node stops receiving the RREQ messages and determines the route probability based on the connectivity and packet delay by timestamps.

The IRQV uses connectivity, transmission delay, direction, and distance to find the best route in the network. The source node checks the direction of the communication terminal (CT) and the distance between the CT and its neighboring intersection to determine the terminal intersection. The highest grade terminal intersection is selected, whereby the source node sends RREQ messages to check the available

route by forward and backward ants. The data packets are then dynamically forwarded to the next intersection and the greedy carry-and-forward algorithm is used to recover and relay the data between two intersections.

5.3. Discussion and Comparison. The IBR protocol depends on the weighting adjustment of roads to find the optimal route in the network. Weighting factors are assigned to routing metrics in order to create the weighting score for roads. These weighting factors have positive and negative effects on routing decisions. In addition, the weighting value of routing metrics (density, connectivity, and distance) depends on the protocol design and evaluation metrics. However, the weighting calculation is not carried out in the aforementioned protocols. The measurement of routing metrics is another significant process for routing. However, this process leads to network overhead and computational complexity issues. Therefore, this process needs to be done when it is essential and when a measurement is requested such as in iCAR, CAIR, and IRQV. Thus, the measurement results must be updated for better routing decisions at intersections with the lowest network overhead. In the iCAR protocol, the time validity function for routing metrics is very effective to deal with outdated measurements and is better than fixed validity intervals. A few IBR protocols consider road transmission delay as an indicator of network connectivity. However, the transmission delay is not only caused by network disconnections, it can also be caused by network load. Thus, in order to measure real-time road connectivity, the link lifetime is also taken into account.

Simple greedy approaches in routing protocols may result in packet dropping due to the high mobility of vehicles. The improved and selective greedy methods overcome this issue (as in JBR). Another important point relates to the use of prediction-based strategies for neighbor positions, which is not always accurate compared to updated routing tables. Most intersection-based RAR protocols depend on the carry-and-forward approach for the recovery process. However, this approach leads to network delay issues. Therefore, some

TABLE 3: Twofold road-aware routing protocols.

Protocol, year	TAR, 2011 [36]	TARGET, 2012 [37]	IG, 2015 [38]
Routing metrics on straight road	Traffic density, velocity, and direction	Distance	Direction and forwarding progress
Routing metrics at the intersection	Large volume of vehicle average neighbor	Node density, link quality, and distance	Traffic density and distance
Routing metrics measurement	Initiated when fixed interval period expires	Commences when vehicle forwarder node leaves road	Initiated when fixed interval period expires
Recovery mechanism	Carry-and-forward strategy	Checks the link quality and returns data to intersection node	
Route selection	Greedy forwarding strategy	Greedy forwarding strategy	Greedy forwarding strategy
Protocols used for comparison	GPSR	GPSR	RBVT-R, AODV, and GPCR
Packet delivery ratio	High	High	High
End-to-end delay	Low	Low	Low
Network overhead	Not measured	Not measured	Not measured

other strategies must be considered in future research. The network load is another important criterion to avoid congestion issues in the network. The network load is not included in any of the aforementioned intersection protocols.

6. Twofold Decision Routing

The most recent twofold RAR protocols are discussed in this section. The discussion is based on an overview of their routing processes, forwarding strategies, routing metrics, and recovery techniques, followed by a comparison of their characteristics and performance. Table 3 presents a summary of the twofold RAR protocols.

6.1. Routing Process of Twofold Decision Routing. Among the recent twofold RAR protocols, the traffic-aware routing (TAR) protocol [36] is based on the selection of roads using the vehicle average neighbor volume for forwarding the data packets. The protocol assumes that the vehicle nodes are equipped with a simple electronic map and are able to collect the road topology information. Vehicles are aware of their own location and the destination location through location-based services. The TAR protocol computes the road traffic volume and selects the greatest value path for the data forwarding process. The protocol has two modes of operation, namely, the inside road (or between intersection) and at the intersection. When the forwarding node is at the intersection, it can check the high average neighbor nodes and select a forwarding path. When the forwarding node is in between intersections, it uses the greedy strategy.

A similar routing protocol known as traffic-aware geographical routing (TARGET) was proposed in [37]. The protocol consists of three basic modules: intersection selection, data delivery, and recovery strategies. The protocol makes its decision with high density areas instead of low density areas. In this protocol, the nodes are divided into ordinary and junction nodes. One monitor is assigned for each junction.

The junction node, which is closest to the centre, is set as a monitor for communication with other neighbors. It monitors the traffic information periodically and broadcasts to other junction nodes. Furthermore, every monitor node maintains the neighbor table and analyzes the link quality and node density.

As proposed in [38], improved geographical (IG) routing establishes communication between vehicle nodes. In the testing of the protocol, the author assumed that all vehicles were equipped with GPS and digital road maps and that the vehicles were aware of the position and direction information by beacon messages. In this protocol, the routing is based on multiple metrics to ensure the wireless link stability, reliability, and packet forwarding progress. The IG protocol has two modes of routing, namely, packet forwarding between intersections and at the intersection. Through an improved greedy mode, the protocol forwards data packets in the urban scenario.

6.2. Routing Metrics Measurement, Forwarding, and Recovery Mechanisms. The TAR protocol uses the vehicle average neighbor volume to perceive information at the intersection and selects the greatest value path for data forwarding. Traffic density is a very significant routing metric for collecting the information. The traffic density is calculated through gathering the number of vehicle neighbors and computing the density average as adopted in TAR, IG, and TARGET. When the node is in between intersections, the protocol measures vehicle velocity, current road ID, and position information. The greedy forwarding strategy is adopted to select the closest node to the next junction as a next hop. If the node is not found, then the carry-and-forward strategy is applied to move forward until it finds a node. The protocol has some conditions to calculate the next road characteristics such as neighbor node estimation and last neighbor estimation.

On the other hand, TARGET depends on node density with the link quality of each road as the routing metrics.

The protocol maintains the routing table with information about the straight distance to every neighbor intersection, the time the last defective packet arrived for each road, and the node density of each road. TARGET uses the Dijkstra algorithm to select the next shorter path intersection when the node is located between intersections. Furthermore, when a node is at the intersection, the node excludes the uncounted roads and selects the next forwarder. In the recovery strategy, if data are not delivered due to link breakage, it returns to the last junction. That junction is considered to be an unconnected road and another intersection is selected for data forwarding.

In the IG protocol, the routing metrics are different in between intersections and at the intersection. Direction and forwarding progress are taken into account when the node is between intersections. When not at the intersection, the protocol uses traffic density and distance toward destination. The protocol is based on multimetric type for the harsh vehicular environment, where a single metric is not as effective for data forwarding. The protocol has a score function mechanism for determining the routing metrics and a beacon reception rate mechanism for measuring the channel quality.

6.3. Discussion and Comparison. This section presents a comparison of twofold decision routing protocols and a discussion of their limitations and issues. The main discussion points are as follows:

- (1) In the aforementioned twofold decision protocols, the routing decisions are different in between intersections and at the intersection. This strategy is adopted due to the harsh vehicular environment, where one routing metric is not as effective. The different routing metrics should have an effective strategy for measurement. In addition, the selection of routing metrics is another significant criterion during the development of these types of strategies.
- (2) This type of strategy lead to network overhead and computational complexities but it depends on the routing metrics selection and processes.
- (3) These protocols have the same strategies as junction-aware routing but with different routing metrics.
- (4) Another issue relates to the calculation of the routing metrics: this process leads to network overhead and has not been investigated in the aforementioned protocols.
- (5) The real road environments are dynamic in nature, where the roads are often short between intersections. In these cases, these protocols suffer from computational complexity to divert their operational processes. However, there is a need to test these protocols in a real test environment.

After the brief comparison of the twofold decision routing protocols, the following points can be concluded:

- (1) Twofold decision routing protocols are more effective compared with intersection-aware routing protocols, due to different routing metric strategies.

- (2) In some protocols, the network overhead and computational complexity are not taken into account. However, these measurements are very significant in these protocols.

7. Traffic Density-Based Routing

This section presents the most recent traffic density-based routing protocols. The routing metrics, forwarding, and recovery strategies are discussed, followed by a comparison. Table 4 presents a comparison of the protocols and their performance.

7.1. Density-Based Routing Protocol Processes. Road or traffic density is one of the important metrics to evaluate for routing in a network. To this end, the density-aware routing using road hierarchy (DAR-RH) was proposed in [39]. The DAR-RH protocol exploits the hierarchical road information for forwarding the data towards the destination. Further, the protocol uses hierarchies to classify the city roads and greedily forward the packets to the destination. The road hierarchy is categorized into freeway, trunk, secondary road, and township roads. The protocol computes the shortest path information in every hierarchy and calculates the shortest path.

The density-aware reliable broadcasting protocol (DECA), as proposed in [40], is based on the utilization of one-hop local density information to forward the data toward the destination. The protocol selects the neighbor node with the highest density, rebroadcasts the message, and maximizes the number of neighbor nodes. Further, if the selected node does not have higher density, then another node with higher density is selected thus avoiding the loop problem.

In [41], a real-time traffic density protocol was proposed for reliable and fast communication in urban scenarios. In this protocol, every vehicle node calculates the traffic density of the road and establishes a reliable route towards the destination. Each vehicle periodically transmits the beacon messages containing the direction and the total number of reverse cars (TRC) with the position. The TRC estimates the road vehicle density and sends a beacon message to its one-hop neighbor. Figure 5 presents an example of a road layout with various road densities in the real-time traffic density protocol.

An adoptive routing protocol based on QoS and traffic density (ARP-QD) was proposed in [42] to find the best path by determining the hop count and link duration in a network. The protocol also uses an adaptive neighbor discovery protocol to determine the local traffic density. The protocol assumes that destinations are always located at intersections and vehicles are equipped with GPS services. The protocol uses connectivity and distance (CDP) metrics. When a node is at the road segment, it selects the next hop with parallel direction; if a node is at the intersection, it selects the next segment first and then decides the next hop. The candidate intersections are defined as the adjacent intersection with the shortest path lengths, where the vehicle node is formularized along the road to find the best routing

TABLE 4: Comparison of traffic density-based road-aware routing protocols.

Protocol, year	DAR-RH, 2009 [39]	DECA, 2010 [40]	Real-time road vehicle density, 2013 [41]	ARP-QD, 2015 [42]
Routing metrics	Traffic density and distance	Traffic density	Traffic density and direction	Traffic density, direction, and network connectivity
Routing metrics measurement	Commences when vehicle forwarder node leaves road	Initiated when fixed interval period expires	Commences when vehicle forwarder node leaves road	Through segment, connectivity, and distance-based weighting function
Recovery mechanism	The test packet is sent via a higher hierarchy	Buffers the message and randomly sets a waiting timeout	N/A	Carry-and-forward mechanism
Route selection	Source-based	Source-based	Source-based	Source-based
Protocols used for comparison	GPSR	AckPBSM and DECA + ABI	GPSR	GPSR and ROMSGP
Packet delivery ratio	Higher	Not measured	High	Higher
End-to-end delay	Low	Not measured	Not measured	Average
Network overhead	Not measured	Low	Low	Not measured

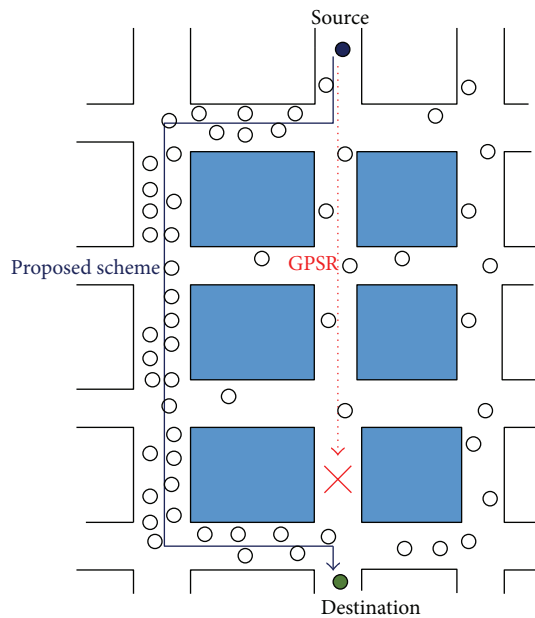


FIGURE 5: Road layout example with different traffic densities in real-time traffic density protocol.

path. The protocol selects the most connected path in order to avoid any disconnectivity. Once the next segment is selected, the direction of the packet delivery is determined. To find the best route in a network, the protocol uses a hop-by-hop strategy. The protocol has a higher data delivery ratio and low end-to-end delay compared with GPSR and ROMSGP. However, its assumption that the destination is always at an intersection is not realistic.

7.2. Routing Metrics Measurement, Forwarding, and Recovery Mechanisms. The routing metrics of the DAR-RH, distributed efficient clustering approach (DECA), and real-time road density protocols are traffic density measurement and estimation, where the number of vehicles is measured with the road per unit length. In some studies, researchers have measured traffic density with the number of vehicles in a certain area of a city [43, 44]. However, vehicle density is a significant traffic condition metric for routing in a network.

DAR-RH uses road hierarchies and discovers the route by calculating the shortest path in a spatial model. For the shortest path, the protocol uses distance and real-time traffic density by testing the packets in a unicast manner. These packets gather the traffic density information at every intermediate node until reaching the destination. In the case of the local maximum issue, the protocol transfers packets along another route. The protocol adopts greedy forwarding to check each intermediate waypoint as a temporary destination.

The density-aware reliable broadcasting protocol uses local density information to select the next rebroadcasting node with the store-and-forward strategy. The neighbor with the highest density rebroadcasts the messages and the other neighbor nodes store the messages and set a waiting timeout. If a forwarder node faces the channel error or collision issues, the other neighbor nodes rebroadcast the messages. DECA has two neighbor lists, known as the broadcast list and neighbor list, and maintains the identifiers for all one-hop neighbors and their local density. DECA overcomes the broadcast storm problem, where a large number of data messages are dropped and the protocol needs to retransmit messages for reliable status. Further, in this context, DECA uses the highest beacon interval for fast dissemination by dynamically calculating the adaptive beacon interval.

The real-time road vehicle density protocol provides stable routes by using high vehicle density as a routing metric. Each vehicle node periodically transmits beacon messages containing the direction, TRC, movement, and own location. The TRC value is computed to estimate the road vehicle density and sends its value to the one-hop neighbors. Whenever this value is received the “cars” field value is increased by one. For searching the routes, the protocol uses request and reply messages.

The ARP-QD uses a novel metric for selecting the best next hop in a network such as direction, relative speed, and distance between the candidate and neighbor node and sets the connectivity and distance (CDP) value. The node with the largest CDP is selected as the next hop. In the case of routing path failures, the protocol adopts the carry-and-forward recovery strategy. In the carry-and-forward strategy, the protocol carries the packet until another node moves into transmission range to transfer the packet. In order to complete the routing process, ARP-QD uses the shortest length road segment and intersection and determines three routing metrics (traffic density, direction, and network connectivity). These routing metrics are combined in two methods: connectivity and distance and segment selection weight. With the help of these two methods, the protocol obtains a qualified path to fulfill the QoS requirements and balance the path efficiency and stability. The protocol utilizes a recovery strategy with the carry-and-forward approach based on the local vehicle density information.

7.3. Discussion and Comparison. The vehicle density estimation protocols have some limitations in the field of vehicular networks due to high variable density and high mobility. These protocols are based on VANET traffic flows and communication properties. However, the traffic density strategies have the following issues:

- (1) These traffic density-based routing protocols use beacon messages to update the neighbor tables. The frequent broadcasting of these messages leads to communication overhead.
- (2) The vehicle density is more condensed near traffic lights and intersections compared to other roads. These situations are the cause of wrong indications of high density for the protocols.
- (3) Some density-based routing protocols do not include the vehicle direction in the calculation of the vehicle density. Basically, the vehicle direction toward the destination with high vehicle density is projected for routing and is considered to be a favorable metrics. However, without vehicle direction, the route with the same vehicle density but in two directions has higher breakage probability [30]. DECA only uses vehicle density and suffers more from this issue compared to the DAR-RH, ARP-QD, and real-time protocols.

As shown in the summary of the traffic density estimation-based routing protocols (Table 4), these protocols suffer from network overhead. To solve this issue, the protocols should have feasible routing metrics to deal with a dynamic

environment. Most of the protocols in this category are based on the greedy forwarding approach and face local maximum issues in the network, where the vehicle nodes do not find the nodes that are close to the destination.

8. Conclusion

This review presented an overview of RAR protocols for VANETs. The main objective was to present a comparison of recent protocols of this type and categorize them, according to different aspects, into full RAR, intersection-aware, twofold routing, and density-based routing protocols. Further, for every category, the routing metrics, processes, and forwarding mechanism and recovery approaches were inspected. RAR protocols belong to the category of geographical routing and are designed to address the geographical routing issues. The review also presented a comparison of the RAR protocols and a discussion of their challenges.

RAR protocols are more efficient for the dynamic and highly mobile vehicular environment. However, the main challenge in these protocols is how to achieve sufficient routing with lower network overhead. Thus, the routing metrics and their measurements play a significant role in this regard and need more improvement. Moreover, the RAR protocols must be more aware of road conditions.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for funding this research. The research is supported by Ministry of Education Malaysia (MOE) and conducted in collaboration with Research Management Center (RMC) at Universiti Teknologi Malaysia (UTM) under VOT no. QJ130000.2528.06H00.

References

- [1] M. J. Piran, G. R. Murthy, and G. P. Babu, “Vehicular ad hoc and sensor networks; principles and challenges,” *International Journal of Ad Hoc, Sensor & Ubiquitous Computing*, vol. 2, no. 2, pp. 38–49, 2011.
- [2] C. Tripp-Barba, L. Urquiza-Aguiar, M. A. Igartua et al., “A multimetric, map-aware routing protocol for VANETs in Urban Areas,” *Sensors*, vol. 14, no. 2, pp. 2199–2224, 2014.
- [3] M. Amadeo, C. Campolo, and A. Molinaro, “Enhancing IEEE 802.11p/WAVE to provide infotainment applications in VANETs,” *Ad Hoc Networks*, vol. 10, no. 2, pp. 253–269, 2012.
- [4] C. Chen, Y. Jin, Q. Pei, and N. Zhang, “A connectivity-aware intersection-based routing in VANETs,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, article 42, 16 pages, 2014.
- [5] S. Al-Sultan, M. M. Al-Doori, A. H. Al-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, 2014.

- [6] S. Singh and S. Agrawal, "VANET routing protocols: issues and challenges," in *Proceedings of the Recent Advances in Engineering and Computational Sciences (RAECS '14)*, pp. 1–5, Chandigarh, India, March 2014.
- [7] K. N. Qureshi, A. H. Abdullah, and R. Yusof, "Position-based routing protocols of vehicular Ad hoc networks & applicability in typical road situation," *Life Science Journal*, vol. 10, no. 4, pp. 905–913, 2013.
- [8] A. Fonseca and T. Vazão, "Applicability of position-based routing for VANET in highways and urban environment," *Journal of Network and Computer Applications*, vol. 36, no. 3, pp. 961–973, 2013.
- [9] A. Bazzi and A. Zanella, "Position based routing in crowd sensing vehicular networks," *Ad Hoc Networks*, vol. 36, pp. 409–424, 2016.
- [10] S. M. Bilal, C. J. 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, 2013.
- [11] K. N. Qureshi and A. H. Abdullah, "Localization-based system challenges in vehicular Ad Hoc networks: survey," *Smart Computing Review*, vol. 4, pp. 515–528, 2014.
- [12] L. K. Waghdhare, T. Nagrare, and K. Gudadhe, "Review on routing protocol for vehicular ad-hoc network," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, 2012.
- [13] F. Cadger, K. Curran, J. Santos, and S. Moffett, "A survey of geographical routing in wireless ad-hoc networks," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 2, pp. 621–653, 2013.
- [14] A. Dua, N. Kumar, and S. Bawa, "A systematic review on routing protocols for vehicular Ad Hoc networks," *Vehicular Communications*, vol. 1, no. 1, pp. 33–52, 2014.
- [15] B. Karp and H.-T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MOBICOM '00)*, pp. 243–254, August 2000.
- [16] C. Li, C. Zhao, L. Zhu, H. Lin, and J. Li, "Geographic routing protocol for vehicular ad hoc networks in city scenarios: a proposal and analysis," *International Journal of Communication Systems*, vol. 27, no. 12, pp. 4126–4143, 2014.
- [17] S. Parvin, M. A. Sarram, G. Mirjalily, and F. Adibnia, "A survey on void handling techniques for geographic routing in VANET network," *International Journal of Grid & Distributed Computing*, vol. 8, no. 2, pp. 101–114, 2015.
- [18] A. Pascale, M. Nicoli, and U. Spagnolini, "Cooperative bayesian estimation of vehicular traffic in large-scale networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2074–2088, 2014.
- [19] M. Al-Shugran, O. Ghazali, S. Hassan, K. Nisar, and A. S. M. Arif, "A qualitative comparison evaluation of the greedy forwarding strategies in mobile Ad Hoc network," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 887–897, 2013.
- [20] J. Liu, J. Wan, Q. Wang, P. Deng, K. Zhou, and Y. Qiao, "A survey on position-based routing for vehicular ad hoc networks," *Telecommunication Systems*, pp. 1–16, 2015.
- [21] S.-H. Cha and K.-W. Lee, "Location prediction for grid-based geographical routing in vehicular ad-hoc networks," in *Grid and Distributed Computing*, pp. 35–41, Springer, 2011.
- [22] K. Shafiee and V. C. M. Leung, "Connectivity-aware minimum-delay geographic routing with vehicle tracking in VANETs," *Ad Hoc Networks*, vol. 9, no. 2, pp. 131–141, 2011.
- [23] D. Wu, J. Luo, R. Li, and A. Regan, "Geographic load balancing routing in hybrid vehicular ad hoc networks," in *Proceedings of the 14th IEEE International Intelligent Transportation Systems Conference (ITSC '11)*, pp. 2057–2062, Washington, DC, USA, October 2011.
- [24] H. T. Hashemi and S. Khorsandi, "Load balanced vanet routing in city environments," in *Proceedings of the IEEE 75th Vehicular Technology Conference (VTC '12)*, pp. 1–6, Yokohama, Japan, May 2012.
- [25] J. Bernsen and D. Manivannan, "RIVER: a reliable inter-vehicular routing protocol for vehicular ad hoc networks," *Computer Networks*, vol. 56, no. 17, pp. 3795–3807, 2012.
- [26] Y. Xiang, Z. Liu, R. Liu, W. Sun, and W. Wang, "GeoSVR: a map-based stateless VANET routing," *Ad Hoc Networks*, vol. 11, no. 7, pp. 2125–2135, 2013.
- [27] C. Liu, Y. Shu, O. Yang, Z. Xia, and R. Xia, "SDR: a stable direction-based routing for vehicular ad hoc networks," *Wireless Personal Communications*, vol. 73, no. 3, pp. 1289–1308, 2013.
- [28] N. Alsharif and X. S. Shen, "ICARII: intersection-based connectivity aware routing in vehicular networks," in *Proceedings of the 1st IEEE International Conference on Communications (ICC '14)*, pp. 2731–2735, New South Wales, Australia, June 2014.
- [29] S. K. Bhoi and P. M. Khilar, "A road selection based routing protocol for vehicular ad hoc network," *Wireless Personal Communications*, vol. 83, no. 4, pp. 2463–2483, 2015.
- [30] R. Oliveira, M. Luisa, A. Furtado, L. Bernardoa, R. Dinisa, and P. Pinta, "Improving path duration in high mobility vehicular ad hoc networks," *Ad Hoc Networks*, vol. 11, no. 1, pp. 89–103, 2013.
- [31] K. N. Qureshi and A. H. Abdullah, "Localization-based system challenges in vehicular ad hoc networks: survey," *The Smart Computing Review*, vol. 4, no. 6, pp. 515–528, 2014.
- [32] L.-D. Chou, J.-Y. Yang, Y.-C. Hsieh, D.-C. Chang, and C.-F. Tung, "Intersection-based routing protocol for VANETs," *Wireless Personal Communications*, vol. 60, no. 1, pp. 105–124, 2011.
- [33] S. Tsiachris, G. Koltidas, and F.-N. Pavlidou, "Junction-based geographic routing algorithm for vehicular ad hoc networks," *Wireless Personal Communications*, vol. 71, no. 2, pp. 955–973, 2013.
- [34] N. Alsharif, S. Céspedes, and X. S. Shen, "iCAR: intersection-based connectivity aware routing in vehicular ad hoc networks," in *Proceedings of the IEEE International Conference on Communications (ICC '13)*, pp. 1736–1741, IEEE, Budapest, Hungary, June 2013.
- [35] G. Li, L. Boukhatem, and S. Martin, "An intersection-based QoS routing in vehicular ad hoc networks," *Mobile Networks and Applications*, vol. 20, no. 2, pp. 268–284, 2015.
- [36] H. Li, "Routing protocol of sparse urban vehicular ad hoc networks," in *Advanced Research on Electronic Commerce, Web Application, and Communication*, pp. 211–217, Springer, 2011.
- [37] X. Li, "Traffic-aware geographical routing in vehicle ad hoc networks," in *Future Wireless Networks and Information Systems*, pp. 755–759, Springer, 2012.
- [38] K. Z. Ghafoor, J. Lloret, A. S. Sadiq, and M. A. Mohammed, "Improved geographical routing in vehicular Ad Hoc networks," *Wireless Personal Communications*, vol. 80, no. 2, pp. 785–804, 2014.
- [39] J. Mouzna, S. Uppoor, M. Boussedjra, and M. M. Manohara Pai, "Density aware routing using road hierarchy for vehicular networks," in *Proceedings of the IEEE/INFORMS International*

Conference on Service Operations, Logistics and Informatics (SOLI '09), pp. 443–448, Chicago, Ill, USA, July 2009.

- [40] N. Na Nakorn and K. Rojviboonchai, “DECA: density-aware reliable broadcasting in vehicular ad hoc networks,” in *Proceedings of the 7th Annual International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON '10)*, pp. 598–602, Chaing Mai, Thailand, May 2010.
- [41] H. Yu, J. Yoo, and S. Ahn, “A VANET routing based on the real-time road vehicle density in the city environment,” in *Proceedings of the 5th International Conference on Ubiquitous and Future Networks (ICUFN '13)*, pp. 333–337, IEEE, Da Nang, Vietnam, July 2013.
- [42] 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. 2015, Article ID 631092, 13 pages, 2015.
- [43] B. S. Kerner, *Introduction to Modern Traffic Flow Theory and Control: The Long Road to Three-phase Traffic Theory*, Springer Science & Business Media, Berlin, Germany, 2009.
- [44] J. Barrachina, P. Garrido, M. Fogue et al., “I-VDE: a novel approach to estimate vehicular density by using vehicular networks,” in *Ad-Hoc, Mobile, and Wireless Network: 12th International Conference, ADHOC-NOW 2013, Wrocław, Poland, July 8–10, 2013. Proceedings*, vol. 7960 of *Lecture Notes in Computer Science*, pp. 63–74, Springer, Berlin, Germany, 2013.