CrossMark

Traffic aware routing in vehicular ad hoc networks: characteristics and challenges

Tasneem Darwish · Kamalrulnizam Abu Bakar

Abstract Vehicular ad hoc networks (VANETs) are gaining tremendous interest among researchers and industries. Although the main reason for developing VANETs is traffic safety, many applications such as traffic status monitoring, road traffic management, routing and distribution of data, have emerged. VANETs exploit multi-hop communications among vehicles to deliver data packets. However, with fast mobility and intermittent link connectivity between vehicles, efficient and reliable routing in VANETs is becoming a challenging task. In order to make routing protocols robust to frequent communication disruptions and aware of unstable traffic and network conditions, several new routing metrics have been integrated with routing protocols. Such protocols are called traffic aware routing (TAR) protocols as their routing decisions are influenced by traffic and network status. The goal of this paper is to review the most recent traffic aware routing protocols while emphasising on traffic and network conditions awareness issues. In addition, this review investigated TAR protocols capabilities and limitations in terms of routing process, routing metrics measurement, forwarding mechanisms and recovery techniques. Moreover, challenges, critical issues and open research problems were discussed in the "Challenges and issues to consider" sections.

Keywords Traffic aware · VANET · Routing · Ad hoc · Density · Connectivity · Review · Survey

T. Darwish (✉) · K. Abu Bakar
Pervasive Computing Research Group Faculty of Computing,
Universiti Teknologi Malaysia (UTM), 81310 Skudai,
Johor D. T, Malaysia
e-mail: sjdtasneem2@live.utm.my

1 Introduction

Vehicular ad hoc networks (VANETs) are special type of mobile ad hoc networks (MANETs). VANETs facilitate ubiquitous connectivity between vehicles through vehicle-to-vehicle (V2V) communications and between vehicles and fixed network infrastructures through vehicle-to-infrastructure (V2I) communications. Utilizing VANET technology enables variety of applications for improving passengers safety, alleviating traffic flow, reducing pollution and enabling in-vehicle entertainment for passengers [1–3]. In addition, vehicles could use the available network infrastructures in a city to access Internet [4].

Nowadays, VANETs are attracting researchers and engineers, as well as automobile industries and governments, as a promising solution to improve road safety and increase passengers comfort. As VANET is an active research field, many review and survey papers which highlight VANET characteristics, applications and latest open research issues have been published recently [5,6]. VANETs have unique and critical characteristics, such as: high vehicle mobility, frequent link disruptions and highly dynamic topology [7]. However, with such characteristics, handling multi-hop communications in VANETs introduces various challenges for researchers.

Delivering data between source and destination in VANET environment requires robust and efficient routing protocol designed specially for this kind of networks. Routing protocols for multi-hop communications within VANET have gained researchers attention during the last few years [8]. Vehicle movement in VANET is restricted to bidirectional movements along roads and streets. Thus, routing strategies that use geographical location information are practical and efficient for data delivery. Accordingly, geographic routing was considered a more promising routing approach

for VANETs, as it provides scalability and robustness against frequent topology changes [9,10].

Geographical routing protocols usually employ global positioning systems (GPS), digital road maps and location services [11], in order to achieve efficient routing performance. However, with the limited inter-vehicle link lifetime and frequent network partitioning, caused by vehicle mobility, it is challenging task to design an efficient geographical routing protocol for VANET. Consequently, providing vehicles with accurate network and traffic information is important to deliver data packets to destination efficiently [10].

Recently, many proposals for VANETs geographical routing protocols that consider several routing metrics have been introduced. Exploiting additional network or traffic condition metrics in routing, with the classical distance metric, has shown remarkable performance improvements [3]. Such routing protocols are called traffic aware routing (TAR) protocols as they are aware of traffic and network conditions. Basically, the main metrics which can be used to enhance routing decisions are vehicular density and inter-vehicle connectivity. A review on traffic density estimation mechanisms in VANET has been conducted in [12], which discusses their limitations and critical issues.

Traffic aware routing is more efficient as it can adapt to variable network and traffic status. In particular, as vehicles are essential entities in VANET for delivering packets from source to destination, TAR protocols make their routing decisions based on the availability of vehicles and the reliability of links between vehicles. Thus, forwarding packets through roads with low density of vehicles or poor-quality communication links can be avoided. However, traffic aware routing is an emerging field of research in VANET with many open research problems. For instance, traffic and network status evaluation is considered a critical issue in TAR protocols, as the accuracy and updateness of such evaluation affect the performance of TAR protocols. Therefore, this review paper inspects the characteristics and limitations of the most recent TAR protocols, and highlights some issues for future research in the "Challenges and issues to consider" sections.

2 Review objectives and structure

The most recent traffic aware routing protocols were investigated in this review paper, which focuses on traffic awareness critical issues and their effect on routing protocol performance. To the best of our knowledge, this is the first review paper which inspects traffic awareness issues in VANET routing protocols. The main objectives of this review paper are summarized in the following points:

1. To explore the most recent TAR protocols as well as their routing process, routing metrics measurement, forwarding mechanisms and recovery techniques.
2. To point out the limitations and shortcomings of reviewed protocols.
3. To summarize the critical issues that need to be considered, while using or developing TAR protocols, and highlight points for further research and investigation.

A general introduction and background was introduced in first section. Then, this section highlights the importance, objectives and structure of this paper. Next, is the third section which explains the concept of geographical routing in VANET. Section four introduces the main concepts and classification of VANET TAR protocols. The first category of TAR protocols, the full path TAR protocols, was investigated in Sect. 5. Section 6 explores junction-based TAR protocols, then hybrid TAR protocols were discussed in Sect. 7. Thereafter, node-based TAR protocols were investigated in Sect. 8. At Sects. 5.5, 6.5, 7.5 and 8.5 "Challenges and issues to consider" were presented, which provide comparison tables and discuss critical issues as well as open research problems. Finally, Sect. 9 concludes this work.

3 Geographic routing concept in VANET

Vehicular ad hoc networks have highly dynamic topology, as nodes move in high speeds, compared to MANET nodes. In general, there are two categories of routing protocol, topology-based and geographic routing, also called position-based. In topology-based routing protocol, all nodes have to store details about the network topology or part of it. Thus, topology-based routing suffer from high maintenance cost and relying on static end-to-end routes in highly dynamic networks. In addition, there is a high possibility that the used topology information is out of date due to nodes mobility [13]. In order to address the problem of topology-based routing, geographical routing was introduced, whereas the need for nodes to store topology information is eliminated as nodes store information about their direct neighbours only. In addition, geographical routing protocols exploit geographical positions of vehicles to route and deliver packets from source to destination instead of IP addresses, which makes them more adaptable to vehicles mobility. Therefore, geographical routing was identified as the most suitable mechanism for packet delivery among VANET nodes [9].

Due to the development of location service systems [11], more motivations are rising towards adopting geographic routing in VANETs. In addition, new generation vehicles are equipped with GPS devices that provide vehicle position information and digital maps, which are essential for geographic routing protocols. In order to be aware of neighbour
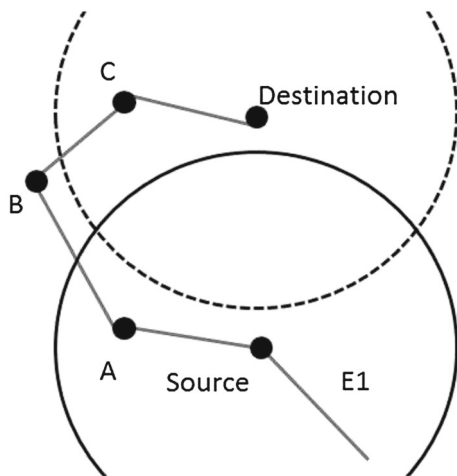
Fig. 1 Local maximum and perimeter recovery

vehicles' conditions, each vehicle maintains a neighbour table to record neighbours information. Neighbour tables are updated upon receiving neighbours' information via hello or beacon messages exchanged between neighbours.

The basic geographical routing protocol is Greedy Perimeter Stateless Routing (GPSR) protocol [14]. The concept of GPSR, which is called greedy forwarding, is to send packets to the closest node to destination. However, greedy forwarding may encounter the local maximum or local optimum problem, where the current forwarder is closer to destination than all its neighbours and the destination still not reachable by one hop communication. Therefore, a recovery mechanism is required in such a case.

The perimeter mode is the GPSR recovery mode, which route packets based on the right hand rule. The rule states that if source node received a packet from edge E1, it sends the packet to its next edge counterclockwise (i.e. to A), as shown in Fig. 1. Afterwards, the forwarding mechanism will switch to greedy forwarding if it is possible, otherwise it will continue in perimeter mode to deliver the packet to destination node. However, GPSR protocol does not consider vehicular network characteristics and has low performance in high mobility situations.

In order to overcome geographical routing shortcomings, new metrics have been introduced in routing protocols, which involve network and traffic status in routing decisions. Accordingly, integrating geographic routing with traffic awareness result in traffic aware routing protocols which adapt to variable traffic conditions. Next section explains the concept and classification of traffic aware routing protocols in detail.

4 Traffic aware routing concept and classification

Although geographical routing protocols were identified as the most efficient protocols for delivering packets in VANET

environment, simple geographical routing has no consideration for VANET special characteristics. In VANET vehicles' motion is restricted to roads and parking areas, which can be indoor or outdoor parking. Moreover, vehicular density affects inter-vehicular communication lifetime and high vehicular density may result in high network load [15]. In addition, buildings and other vehicles might become as obstacles impeding the transmission process. Nevertheless, some of these characteristics can be utilized to enhance the performance of routing protocols. In other words, considering traffic and network status as essential factors in making routing decisions, along with vehicles' positions, will result in routing packets through routes with good connectivity. Basically, this is the core idea of traffic aware routing protocols, which depend on traffic conditions and vehicles positions to deliver packets.

Traffic aware routing is a promising research field in VANET routing as it adapts to variable network conditions. For instance, empty or low vehicular density streets are avoided in packet forwarding as they are considered disconnected. In addition, forwarding a packet for a node which has no neighbours closer to destination can be eliminated by choosing roads with high density and good inter-vehicle communication quality [16]. Moreover, forwarding packets through congested routes can be prevented by considering data traffic load per street in routing decisions, which is called network status awareness. In particular, traffic aware routing utilizes some traffic and network status measurements such as: availability of vehicles in streets (i.e. vehicular density), link lifetime or quality of links between vehicles, vehicles distribution and network communication load per street. Therefore, traffic aware routing provides better performance than other geographical routing as it makes routing decisions based on traffic and network status.

4.1 TAR protocols main components

Every routing protocol apply a certain routing process, which utilizes routing metrics measurement to make routing decisions. Afterwards, packets are forwarded based on routing decision and the adopted forwarding mechanism. In case of network disconnection or routing failure a recovery technique is essential to overcome such problems. Thus, any traffic aware routing protocol consists of four main components, which are explained in the following subsections.

4.1.1 Routing process

The routing process defines how to use routing metrics to calculate routes. In addition, it defines how often routes must be recalculated or updated, to adapt to network conditions. Moreover, it controls the initiation of routing metrics measurement process.

4.1.2 Routing metrics measurement process

Various routing metrics can be used as indicators for traffic and network conditions. However, the performance of traffic aware routing depends mainly on the utilized routing metrics and the way of measuring these metrics. In particular, accurate measuring of routing metrics will give more realistic indications about traffic conditions and result in more efficient routing. In addition, providing routing metrics on time and with least network overhead, whenever they are required for routing, is critical for efficient routing.

4.1.3 Forwarding mechanisms

In all routing protocols, packets need to be forwarded to next-hop or next-junction based on a forwarding mechanism. Greedy forwarding is the simplest mechanism, which sends packets to the closest neighbour node to destination. However, greedy forwarding do not guarantee destination reachability. Therefore, several forwarding mechanisms enhanced packet forwarding by involving new parameters and metrics (e.g. speed and direction [17]) in the next forwarder selection process. However, with all the available enhancement of greedy forwarding, there is still a possibility for packet forwarding problems to occur but with lower probability [9].

4.1.4 Recovery techniques

Although TAR protocols have better performance than other protocols, disconnection and delivery failure problems can still be encountered due to VANET dynamic nature. Thus, in case there is no neighbour node closer to the destination, an effective recovery technique is required. There are three main recovery techniques utilized in geographical routing:

(i) One widely used technique is the right hand rule to traverse graphs. The concept of the right hand rule is that after receiving a packet from an edge, the node will send the packet through the next edge counter-clockwise around itself as shown in Fig. 1. The routing protocol will switch back to forwarding mode once the forwarding node is closer to destination than the node that initiated the recovery mode. However, as the constructed graph in vehicular networks can change rapidly, the right hand approach may lead to loops in forwarding [9].

(ii) The second technique is the carry and forward, sometimes called store and forward. When local maximum is encountered, the current node carries the packet until a suitable next forwarder appears or it reaches the destination. However, this technique result in long delays [18].

(iii) The third technique is based on recalculating forwarding paths, which may result in high network overhead and delay times [19].

4.2 TAR routing strategies

Routing strategies describe the routing process framework. Based on adopted routing strategy, traffic aware routing protocols can be classified into four categories: full path routing; junction-based routing; hybrid routing and node-based routing. Figure 2 shows the distribution of reviewed protocols based on publication year and routing strategy. The following subsections describe the main characteristics of each strategy.

4.2.1 Full path strategy

This strategy is based on defining a full path that a packet should traverse in order to reach its destination. There are two main mechanisms for selecting a full path between source and destination, either by using Dijkstra algorithm (i.e. shortest path algorithm) or by broadcasting packets towards destination to discover the best path. In order to use Dijkstra algorithm the roads map is converted to a weighted graph, where roads are the edges and junctions are graph vertices. Edges weights are based on length or other metrics such as vehicular density. After selecting the full path from the weighted graph, the intermediate junctions' positions, which a packet has to go through, are inserted in each packet header before forwarding. However, full path strategy has the following limitations:

– Each packet has to carry the full list of intermediate junctions or anchors positions [9].
– The path life time is inversely proportional to its length specially in dynamic VANET environment [9].
– Routing decision is static, at least on packet level, as it is done once for each packet and can not be adjusted according to traffic conditions.
– Establishing and maintaining full path introduces high overhead specially with frequent disconnections and path recalculations.
– Applying Dijkstra algorithm on large city map might generate high computational complexity.
– Utilizing broadcasting to discover routes introduces extra network overhead. Although controlled broadcast mechanisms such as opportunistic broadcasting [20] can be used, in high density situation broadcasting might cause network congestion and degrades network performance.

4.2.2 Junction-based or anchor-based strategy

Dynamic routing decision is adopted in this strategy. For each packet, the next intersection is selected dynamically based on routing metrics. Thus, an evaluation is carried out at each junction for adjacent roads, in order to select the best road

Fig. 2 Taxonomy of traffic aware routing
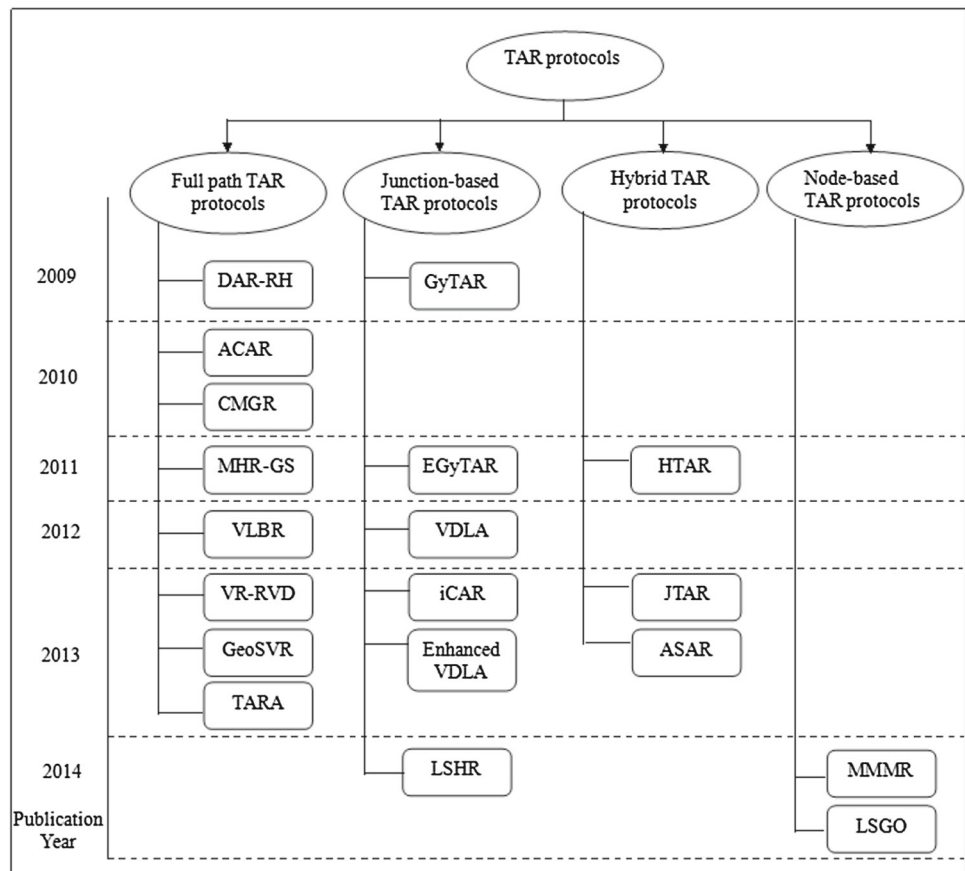
for packet forwarding. Usually the evaluation process (i.e measurement process) at each junction is based on sending information collector packets through roads, in order to collect information about road network and traffic conditions.

Junction-based routing has less overhead compared to full path routing [17], as packets only carry the position of the destination and next-intersection. In addition, it adapts more efficiently to variable VANET conditions as routing decision is done sequentially at each junction. However, the evaluation process may result in network overhead and delay time. Moreover, keeping routing metrics measurement results updated and always available for routing decisions is a critical issue. In particular, to make the measurement results always available and updated, the measurement process has to be done so frequently. However, initiating many measurement processes may result in high network overhead, which may affect data packet delivery rate.

4.2.3 Hybrid strategy

The hybrid strategy tends to overcome the limitations of full path strategy by allowing route adjustment on certain conditions. Additionally, hybrid strategy considers making routing decisions at each junction unnecessary. Thus, it tends to

decrease the delay and network overhead caused by unnecessary routing decisions at each junction.

4.2.4 Node-based strategy

In this strategy each node selects one of its neighbours as a next forwarder. The selected neighbour is the closest to destination and which satisfies the next-node selection conditions. Although node-based strategy does not result in network overhead, it does not take road conditions into consideration. Consequently, routing decision is made based on neighbours status, which is considered efficient for small scale routing only. Thus, sending packets towards disconnected roads can not be avoided in this strategy, which may cause delay and packet loss.

5 Full path traffic aware routing protocols

This section discusses the most recent full path TAR protocols in terms of routing process, routing metrics measurement, forwarding mechanism and recovery techniques. Afterwards, protocols limitations, challenges and open research issues were highlighted in the challenges and issues to consider section. Finally, Table 1 compares the

Table 1 Full path TAR protocols

Protocol	ACAR [22]	DAR-RH [23]	MHR-GS [24]	CMGR [25]	VLBR [27]	VR-RVD [28]	GeoSVR [29]	TARA [30]
Routing metrics	Connectivity (Distance and vehicular density), transmission quality	Distance, vehicular density, transmission delay	Route lifetime using: vehicle position, speed, direction	Vehicular density, route trip time	Distance, vehicular density, network load	Vehicular density	Distance, vehicular density	Distance, transmission delay, vehicular density and velocity
Routing metrics measurement	Uses statistical density. Utilizes data packets to collect real-time density	Forward test packets to Measure density and delay along the two shortest paths	Using broadcasted GA	Measured via RD packets. Result sent via RR packets	Utilizes statistical density and contention window size as network load indicator	Measured via RREQ packets	Wider road means higher density	Based on statistical data
Route selector	Source	Destination	Source	Source	Source	Destination	Source	Source
Network load balance	No	No	No	No	Yes	No	No	No
Traffic awareness scope	The shortest path	The two shortest paths	The specified geographical area for a GW	Depend on broadcast area	The k-shortest paths	All routes from source to destination	On sub-graph between source and destination	The whole city map
Route update	Recalculate path when traffic density change	After a certain period source starts new route calculation	Always update routes by receiving new GA	Always updated by broadcasting RD	Switch between k-shortest paths to balance network load (no recalculation)	No route maintenance	No route maintenance	After a time threshold
Parameters for shortest path selection	Distance and statistical density	Distance and road hierarchies	Shortest path not used	Shortest path not used	Distance and statistical density	Shortest path not used	Roads length and width	Distance, transmission delay, vehicular density and velocity

Table 1 continued

Protocol	ACAR [22]	DAR-RH [23]	MHR-GS [24]	CMGR [25]	VLBR [27]	VR-RVD [28]	GeoSVR [29]	TARA [30]
Depend on broadcasting	No	No	Yes (in specific geographic area)	Yes (Not limited)	No	Yes (Not limited)	No	No
Forwarding mechanism	To neighbour closest to destination and has lowest transmission error	Greedy forwarding	To the furthest neighbour with the longest link lifetime	Greedy forwarding	To furthest neighbour moving to destination (sparse) or furthest neighbour with least network load (dens)	Through discovered paths	Restricted forwarding algorithm	Greedy forwarding
Recovery technique	Carry and forward	Utilizes alternative route and initiate new route calculation	Switch to alternative routes	Carry and Forward (adjusted at junctions)	Carry and forward, path switching for load balancing	Not provided	Not provided	Carry and forward within allowed area
Performance								
Compared to other TAR protocols	No (VADD [43], MOVE [44], GPSR, GSR [45])	No (GPSR)	No (AODV+ [46], GPSR)	No (VADD, A-STAR [47])	No (GPSR)	No (GPSR)	No (AODV [48], GPSR)	No (GPSR)
Network overhead	Lower	Not measured	Not measured	Not measured	Not measured	Lower	Not measured	Not measured
End-to-end delay	Lower	Lower	Lower	Higher delay at low density	Lower	Not measured	Lower than AODV and higher than GPSR	Lower specially for large distance
Packet delivery ratio	Higher	Higher	Higher	Higher	Higher	Higher	Higher	Higher

characteristics and performance of the following reviewed protocols. In addition, for each full path TAR protocol, Table 1 shows to which protocol it was compared and what was its evaluation in terms of network overhead, end-to-end delay and packet delivery ratio.

5.1 Routing process of full path TAR protocols

First a connectivity aware routing was introduced in [21], then the author introduced an enhanced adaptive connectivity aware (ACAR) routing protocol in [22]. ACAR selects a route based on network connectivity and estimated transmission quality of road segments. The author assumed that vehicles installed a pre-loaded digital map, which not only describes the land attributes such as road topology and traffic light period but also is accompanied by traffic statistics such as traffic density and average velocity at certain times of the day.

ACAR computes the shortest and best path towards destination based on the digital map and density statistical data, then attaches path information to packets and transmits packets along selected path. Forwarded packets to destination will collect information about vehicular density along the path in its header. After a pre-defined number of packet transmissions, the density information of road segments in the route can be obtained at the destination. If the difference between measured and statistical density values for a certain road segment exceeds a threshold, then a notification is sent to the source about the updated density of that road. Next, the source node re-computes a new route based on the recently acquired and more accurate density data. Although, ACAR use statistical density measurements, it also adapts to any density variation caused by variable network conditions.

A Density Aware Routing using Road Hierarchy (DAR-RH) was proposed in [23], which takes advantage of roads hierarchy information to transfer data from source to destination. The author classified roads to three different hierarchies according to roads type and size. DAR-RH maintains two routes; in case of failure of the primary route, an alternative route is utilized at a higher hierarchy under sparse network conditions. The two routes are defined by calculating two shortest paths based on road hierarchy using Dijkstra algorithm. Then, real time traffic density on the calculated routes is measured. Afterwards, destination chooses the best connected and minimum delay route based on measured density and delay. However, the two calculated paths may not be the best as they were selected based on distance only, then density was used to identify which one is the primary path of them.

In order to keep routes updated, DAR-RH apply a timeout period, after which the source will initiate a new route calculation process. However, the calculation of the timeout period was not explained by the author. In addition, the route maintenance process generates about 30 % of the total network load, which may affect the performance of the routing protocol. Moreover, as the two routes calculation is based on distance, there is a high probability to select the same routes again unless source or destination changes its position. Thus, the route updating process does not consider road traffic variations.

For connecting vehicles to the internet, a multi-hop routing protocol and gateway selection scheme (MHR-GS) was proposed in [24], where the route with the highest link stability is selected. MHR-GS adopts the concept of proactive routing, where routes are discovered and constructed before it is actually required. Vehicle mobility parameters were utilized to select routes with the longest lifetime to connect to an IP network through gateways. In MHR-GS Gateway advertisements (GA), which are broadcasted within a specific geographical area, are used to discover the most stable routes towards a GW. Each GA message includes the gateway related information and the route stability parameter. By receiving a GA a vehicle will have information about the originator GW ID, the path to that GW and the traversed path stability. Vehicles always keep a list of alternative routes as they record the routes of received GA messages in a routing table. The route with the highest stability is selected as the current forwarding route.

A connectivity-aware with minimum-delay geographic routing (CMGR) protocol was proposed in [25]. The main concept of CMGR is to select routes with high connectivity in sparse situations; however, in dense situations CMGR determines the routes with adequate connectivity and selects the route with the minimum delay among them. The author assumed the availability of Gateways (GWs), arbitrarily distributed along roadsides, in order to connect to a fixed IPv6 network.

CMGR constructs full paths towards gateways by broadcasting route discovery (RD) messages. Intermediate vehicles attach their location then rebroadcast the RD message. As a GW might receive several RDs from the same vehicle via different routes, the GW will select the most efficient route based on connectivity evaluation. Similarly, If the RD is received by several GWs, each one sends back a route reply (RR). Upon receiving RR messages at the source vehicle, the most suitable route is selected based on connectivity evaluation. The CMGR protocol continuously inspects the quality of routes and the possibility of connecting to newly emerged GWs, by broadcasting new RDs upon receiving a RR for the second route.

As many vehicles select the highest quality roads (i.e. roads with high vehicular density or connectivity), great amount of data packets will be forwarded through these favourable routes. As a result, higher congestion, delays and drop ratios will occur on such routes. Therefore, in order to

eliminate network bottlenecks and increase network throughput, a load balancing mechanism is essential [26].

In [27], the VANET Load Balanced Routing (VLBR) protocol was proposed. The main aim of VLBR is to balance network load on favourable routes by employing congestion feedback, in order to switch to less congested routes. In VLBR, the availability of a digital city map, which provides vehicular densities for roads at certain times was assumed by the author. First, city map is converted to a directed graph, then the intersections sequences for the k-shortest full paths towards the destination is calculated by the source vehicle. Each road in the map has a weight based on its length and vehicular density. Second, the intersections sequence of the best path is placed in the header of the packet and the packet is forwarded to the first intersection. Although, the k-shortest path algorithm does not impose heavy computational complexity, as stated by the author, applying such algorithm so often generates extra computational load specially in VANET dynamic environment.

The network load balancing of VLBR is done among the k-shortest paths by sending a congestion warning packet to the source vehicle. Upon receiving such a packet the source will switch to the second best path for a certain time (i.e. timer). The congestion warning is sent when the load status for all nodes on that road reaches a congestion threshold. After using the second best path for a certain time the source node will resume using the first best path. However, as the source has no information about the status of the first path, it might switch back to the first path while it is still under congestion. Accordingly, frequent path switching may happen, which cause high packet loss ratio and delay. In addition, VLBR considers load balancing among nodes on a single path while selecting the next forwarder, which will be explained in Sect. 5.3.

In [28] a VANET routing protocol based on real-time road vehicle density (VR-RVD) was proposed. The aim of this protocol is to establish routes with the highest vehicular density based on real-time density measurement. Routes from source to destination are established using route request (RREQs) and route reply (RREPs) messages. The source node generates a RREQ and broadcast it to discover the route and collect vehicular density information on its way to destination. Upon receiving the first RREQ, the destination node waits for the other RREQs forwarded through other paths for a certain duration. Afterwards, the different routes are compared in terms of vehicle density, then the route with the highest density is selected and an RREP is sent back to the source. However, flooding the network with RREQ packets to discover the best route might result in high network load, specially when many nodes are issuing route discovery at the same time. In addition, due to the dynamic environment of VANET, route discovery is issued so frequently in order to maintain broken routes. Thus, flooding the network with
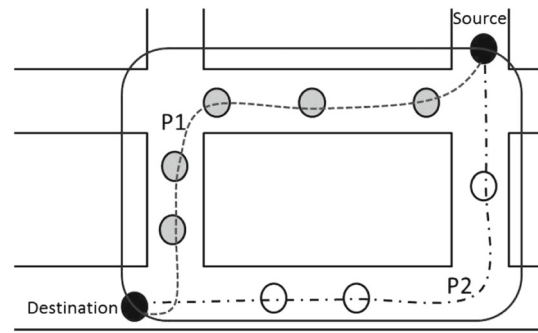

Fig. 3 Sub-graph construction in GeoSVR

route discovery packets is not efficient in VANET and might degrade network performance severely.

The Geographic Stateless VANET Routing (GeoSVR) protocol was proposed in [29], which introduces two main algorithms, Optimal forwarding path (OFP) algorithm and restricted forwarding algorithm (RFA). OFP aims at solving the problems of local optimum and sparse connectivity. On the other hand, RFA is to mitigate the impact of unreliable wireless channel problems.

OFP calculates a path by assigning a weight for each road on the map based on distance to destination and vehicle density. Next, the Dijkstra algorithm is applied to find the optimum forwarding path (e.g. P1 in Fig. 3), which is the shortest path with minimum weight. The optimum forwarding path is inserted into the packet header then the packet is sent towards destination. However, in large city maps the Dijkstra algorithm needs to traverse the whole weighted graph to find a connected path, which result in high computational complexity. Therefore, to reduce complexity and overhead, GeoSVR subtract a connected sub-graph from the whole weighted graph, which is defined according to source and destination positions. In particular, the junctions which are closer to source and destination are determined, then a rectangle surrounding the sub-graph is constructed based on these junctions, as shown in Fig. 3. Afterwards, the OFP can be applied on the sub-graph instead of the whole city map.

In [30] Traffic Aware Routing Algorithm (TARA), which finds the minimum delay and hop routing path to route packets from source to destination, was introduced. TARA was simulated using a data-set of two-year GPS tracking data from over 4000 taxis in Shanghai. Obviously, using real tracking data in simulation contributes towards more realistic results in evaluation process. TARA applies Dijkstra algorithm on weighted city map to find the minimum cost path. Road weighting was based on four parameters: road length and average travel time of a packet to pass that road, while considering the impact of vehicular density and velocity on packet transmission time. Afterwards, data packets

are forwarded based on the calculated routing path, which is a series of road intersections inserted in the packet header. TARA sets a time threshold to update roads weights then recalculates the minimum weight shortest path.

5.2 Routing metrics measurement

ACAR utilizes two routing metrics; connectivity and transmission quality. The connectivity of a road segment is calculated using road length and the number of vehicles on that road (i.e. vehicular density). Afterwards, the transmission quality is estimated based on road segment connectivity and the calculated packet error ratio. The number of neighbours, transmission range and number of lanes are used to calculate packet error ratio. First ACAR depend on statistical vehicular density obtained from maps, then start collecting real-time density while forwarding data packets along selected path. Thus, no special packets were used, such as tester or collector packet, for collecting vehicular density information.

On the other hand, DAR-RH forwards test packets to collect density information from every intermediate node along the two shortest path calculated by the source. Each intermediate node checks the number of neighbours in its neighbour list and specifies the number of neighbours which are ahead of it, then attaches the number to the test packet. Upon receiving a test packet at the destination from one route only, and if the density count is less than a threshold at any intermediate node, the destination waits for a test packet from another route. Afterwards, destination node replies through selected route and data packets forwarding can be resumed.

In MHR-GS route stability is measured while forwarding GA in the specified geographical area. The stability parameter is the estimated route expiration time, whereas the lifetime of a route is the minimum lifetime link between two consecutive vehicles along that route. The the link expiration time (LET) of two consecutive vehicles can be predicted using vehicles motion parameters (i.e. current position, speed and direction). Afterwards, the route expiration time (RET) can be estimated to be the minimum LET along that route. Upon receiving a GA message, a vehicle within the specified geographical area sets a timer before rebroadcasting the message. In addition, it will calculate the LET for its link with previous forwarder, then update the LET field in the GA if it has higher value than calculated LET. While waiting for timer, if another GA message was received from the same gateway with the same sequence number (i.e. re-broadcast by others), the timer is cancelled and both messages are discarded. Otherwise, the message is re-broadcasted once the timer expires and the vehicle is selected as a relay, also neighbour vehicles cancel their timers. The timer is computed based on the stability parameters received via GA.

CMGR selects routes based on their connectivity. Node local density (i.e. number of neighbors), average expected density changing rate and trip time are the metrics used to evaluate each route connectivity. In addition, these metrics are used to distinguish between dens and sparse situations. The measurement of these metrics is carried out while forwarding RD packets from source vehicle to GWs. The route which has the highest density and least trip time is considered to have high connectivity. The average expected value of vehicles density changing rate is calculated and attached to the RD at each forwarder before rebroadcasting. The calculation is carried out in two steps:

1. Each vehicle computes the expected value of density changing rate in its vicinity over a number of beaconing periods based on received beacons from its neighbours.
2. By exchanging expected values via beacons, the average of all expected values from the vehicle's neighbours can be used to calculate the average expected value of density changing rate.

In order to compute trip time for each of the received RDs, the GW calculates the difference between generation and reception time of individual RDs.

VLBR utilizes two main routing metrics for path selection: distance and pre-defined vehicular density from a digital map. However, the used vehicular density was based on statistics, which is not accurate in VANET dynamic environment [12]. In addition, network load was considered an indication to switch between paths. In other words, network load was not an input for the k-shortest path algorithm. For each individual node, network load is measured based on contention window size, which increases with each retransmission trial. In particular, a collision probability was calculated based on current contention window size and the previously calculated probability. Afterwards, each node exchanges with neighbours the calculated collision probabilities. Accordingly, a road is considered a heavy loaded area if all its nodes have collision probability higher than a specified threshold. However, the author did not specify what is the threshold value or how to calculate it. Moreover, no mechanism was defined for measuring collision probability for all nodes on heavy loaded road.

In a similar approach to CMGR, VR-RVD broadcasts RREQ packets toward destination and records the minimum node density on their ways. Vehicle density calculation is done at each traversed node based on the received periodic beacons from neighbor vehicles and a Road Information (RI) table. Beacon messages include vehicle's movement direction, total number of reverse cars (TRC), vehicle identifier and position. Accordingly, a vehicle can get the number of vehicles ahead of it, from the TRC number received via reverse direction cars beacons. Each vehicle calculates its TRC value based on the value of reverse cars field in RI table, driving distance from road beginning and road length.

A RI table is created by each vehicle upon road entry, to store road vehicle density, number of reverse cars, road ID and entrance time. The number of reverse cars recorded in the RI table is increased by one upon receiving a beacon from a vehicle moving in opposite direction. Thus, by using received TRC (i.e. the received TRC from an opposite direction vehicle is the number of ahead vehicles) and the number of reverse cars recorded value, vehicles on a certain road can estimate the current road vehicular density.

In GeoSVR two routing metrics are used: distance to destination and vehicular density. However, the author reckons that a wider road implies a higher probability of high vehicle density. Accordingly, the OFP algorithm uses road types as measurement of vehicle density. Therefore, based on the digital city map, GeoSVR assigns roads (i.e edges) weights according to their length and width. However, there is no guarantee for having high vehicular density in wide roads.

In order to find the minimum delay and hop routing path, TARA makes use of four routing metrics. The first routing metric is distance which can be obtained from the city digital map. The second routing metric is the average packet transmission delay which shows the delay of carrying the packet by vehicles along that road. As the value of packet transmission delay is affected by vehicular density and velocity, these two routing metrics were used in the calculation of such delay value. However, the author used statistical data to predict the future traffic patterns in order to calculate the impact of vehicular density and velocity on packet transmission delay. Consequently, statistical data may result in inaccurate results as vehicular density and speed are highly dynamic metrics and difficult to predict [12].

5.3 Forwarding mechanisms

ACAR selects the neighbour, which is closest to destination and has the lowest transmission error rate, as a next forwarder. This is achieved by keeping a record at each node for packet error rates associated to each one of its neighbours. On the other hand, DAR-RH adopts greedy unicast as a forwarding mechanism to forward test, reply and data packets.

In MHR-GS data packets are forwarded based on the selected path, which was defined in the received GA messages. On the other hand, GA packets are forwarded based on the link lifetime and distance metrics. Thus, in case of forwarding GA packets, the next forwarder is the neighbor which makes the highest progress in packet delivery (i.e. the furthest neighbor), and has the most stable link with the current forwarder.

Similar to DAR-RH, CMGR applies greedy forwarding along discovered routes to forward packets to and from GWs. Moreover, TARA adopts geographic greedy forwarding for data transmissions between two adjacent road intersections.
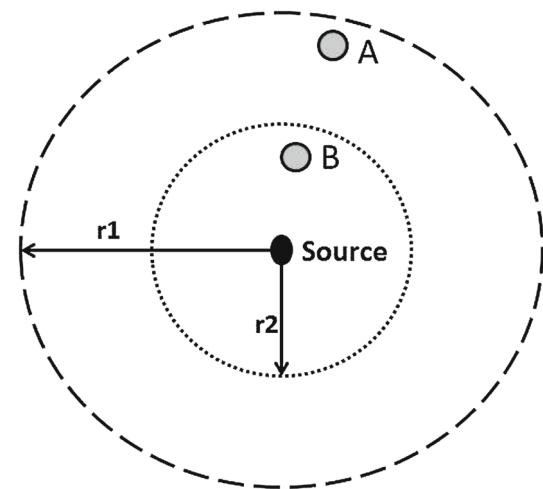

Fig. 4 Restricted forwarding algorithm

Thus, DAR-RH, CMGR and TARA inherits the problems of greedy forwarding.

VLBR has two forwarding mechanisms based on network conditions. By comparing road's density to a threshold, each forwarding vehicle will either consider the road as sparse or connected. If the road status is sparse, the vehicle will forward packets to the neighbour that is closest to next intersection and moving towards it. And if the road is considered connected (i.e. dense), then packets are forwarded to a neighbour that is closest to next intersection and has less network load compared to other neighbours. However, the author did not provide the threshold value or how to calculate it. In addition, the balance between distance and network load on node, while selecting next forwarder, was not clear.

In VR-RVD, once the source receives the RREP message, it will send data packets to destination through the path provided in the RREP. On the other hand, GeoSVR applies a restricted forwarding algorithm, which chooses a neighbour in a specific range for forwarding packets along the calculated OFP. Thus, reducing packet loss due to distant neighbours selection. Figure 4 describes the restricted forwarding algorithm. r_1 is the wireless communication range and r_2 is the restricted forwarding range which was calculated using the free-space path loss and channel frequency of the wireless communication standard. Accordingly, node B can be selected as a next forwarder but node A can not be considered as a potential forwarder.

5.4 Recovery techniques

ACAR adopts the carry-and-forward mechanism to overcome network partitioning problems, where as packets are carried until a suitable next forwarder is found.

DAR-RH utilizes the alternative route in case of primary route failure and a new route calculation process is issued.

However, DAR-RH did not define a mechanism to early inform source node about route failures, which may result in high packet loss ratio. Similarly, MHR-GS can switch to an alternative route when a route fails or its life time expires. However, upon selecting an alternative route an update message must be sent to the gateway, which will update the intermediate vehicles routing tables as well. Alternative routes are always updated via receiving new GA messages.

In case of local maximum in sparse area CMGR apply the concept of carry and forward to deliver packets. However, when a vehicle carrying a packet reaches a junction, the aim is to forward to a vehicle on the first priority out-going road if one is detected. There are two options for forwarding based on the carrying vehicle position:

1. The vehicle did not reach the junction center: if no first priority vehicle is found, the vehicle will forward to any vehicle present in the second priority road as shown in Fig. 5a. However, it will not remove the packet from its buffer after forwarding just in case any opportunity to forward the packet to a first priority vehicle comes up.
2. The vehicle has passed the junction center: the vehicle either continues to carry the packet if it has turned into the first priority road, or looks for vehicles on a higher priority road until it exits the junction, as shown in Fig. 5b.

VLBR adopts the carry and forward technique for forwarding packets in sparse areas. In addition, when a road is considered a heavy loaded area with a lot of re-transmissions, VLBR switches to another less congested path. However, VBLR does not apply any route re-calculation or maintenance mechanism. In particular, VLBR does not define when the k-shortest path algorithm need to be applied or a path re-calculation is required.

In VR-RVD the author did not provide a recovery mechanism or a route maintenance technique. On the other hand, GeoSVR has no recovery technique as the OFP algorithm

avoids choosing sparse routes and the restricted forwarding was intended to eliminate local maximum problem.

In a different approach, TARA utilizes the carry and forward technique in sparse areas or disconnected routes. However, as a vehicle might get faraway from the calculated route, an allowed area was defined around the planned path. Accordingly, recovery is executed as follows:

– If the forwarding vehicle deviates from the planned path but is still inside the allowed area, no correction.
– If the forwarding vehicle is outside the allowed area, TARA calculates the shortest path towards the allowed area.

5.5 Challenges and issues to consider

This section highlights and discusses the challenges and critical issues that need to be considered regarding the mechanism and performance of aforementioned protocols. Moreover, in order to improve the full path traffic aware routing protocols various open research issues were introduced for further investigation in the following points:

● Because of the VANET dynamic environment the main issue is how to select the most stable routes (i.e. the longest lifetime), and how to initiate route recalculation only when it is necessary, in order to adapt to network and traffic conditions and reduce routing overhead.
● Although exploiting route request or route discovery broadcasting mechanisms will discover the best routes, in terms of traffic and network status awareness, it has the following shortcomings:

1. Generates high network overhead.
2. Consumes long time.
3. These mechanism are always tested in simulation environment, where a limited number of vehicles are initiating route discovery. However, in reality broadcasting might be performed by a large number of

Fig. 5 Forwarding a carried packet at junction in CMGR protocol
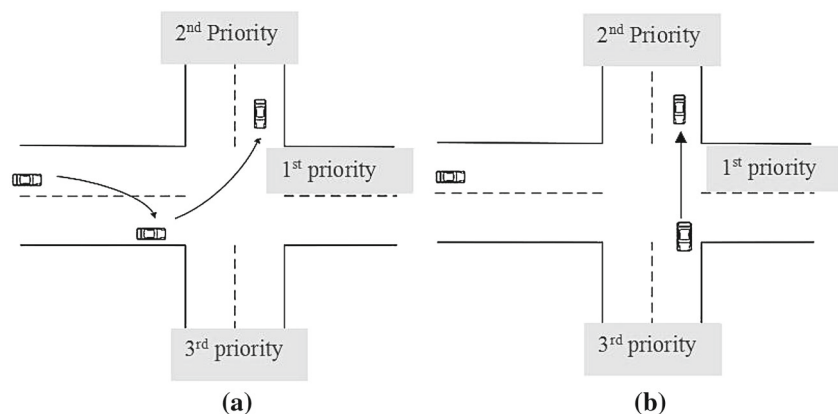
vehicles specially in high density scenarios. Therefore, further investigation is required to measure the performance of these route discovery mechanisms in a more realistic scenario. In addition, mechanisms for mitigating broadcast storms and congestion control are necessary to be integrated with such protocols.

- As many vehicles tend to select the highest quality routes (i.e. routes with high density and good connectivity), such routes will get congested and form network bottleneck. Therefore, network load must be considered while selecting or constructing a route towards destination. Although VLBR apply a load balancing mechanism, it still has the following drawbacks:

  1. Despite of being a load balancing protocol, VLBR does not prevent network congestion on the used routes. In particular, VLBR switches to the second route when the status of all nodes on the first route reach a congestion threshold. Thus, based on such load balancing mechanism, some parts of the route might be already congested and formed network bottleneck.
  2. Network load was not considered as a metric while calculating paths. Thus, the source node has no information about network load on paths while calculating and classifying the k-shortest paths.
  3. The source node will not get any feedback about the status of the first path while using the second path.
  4. The source node can not know the network load status for a path without trying to use that path, which may result in packet delay or loss.

- Finding a group of shortest paths then measuring routing metrics on such routes might cause the following problems:

  1. It may not result in using the optimum routes in terms of traffic and network awareness. In other words, there might be some other routes which are not among the shortest but have better network and traffic conditions.
  2. When route recalculation is applied, the same group of shortest paths might be selected again unless source or destination node has moved to a different road.

  Thus, For route calculation or recalculation, applying shortest path Dijkstra algorithm may not be efficient without considering traffic and network status as inputs for the algorithm.

- For large city maps, the Dijkstra algorithm needs to traverse the whole weighted graph (i.e. city map), which result in high computational complexity. Therefore, to reduce complexity and overhead it's recommended to apply the algorithm on a sub-graph of the city map, such mechanism was utilized in GeoSVR.

- For full path traffic aware routing protocols, which utilize more than one routing metric, giving equal weighting factors for all metrics might not be the optimum solution. Therefore, further investigation need to be carried out to specify the right weighting ratio for each routing metric based on network and traffic conditions.

- Many traffic aware routing protocols depend on measuring vehicular density by collecting the number of neighbours along a road. However, such a mechanism may lead to wrong indication for high density on a road. For instance, sometimes vehicles are more condensed in some areas, such as areas near intersections or traffic lights, than the rest of the road, which result in low and high density areas. Therefore, vehicular spacial distribution must be considered along with vehicular density.

- While calculating vehicular density direction of vehicles should be taken into account. In particular, routes with vehicles moving towards a destination in high density will have less breakage probability than routes with the same density but with vehicles moving away from destination, as stated in [31].

- Forwarding collector packets, which measure routing metrics, along a path using greedy forwarding may result in loosing such packets. Accordingly, the path might be considered disconnected because of collector packet lose. Thus, it is recommended to avoid using greedy forwarding to forward collector packets, in order to get more accurate measurements for route conditions.

- Some full path TAR protocols utilizes statistical density, such as ACAR, VLBR and TARA. However, statistical density is not accurate because VANET is a highly dynamic environment and predicting vehicular density is difficult as stated in [12].

- Periodic route discovery or proactive routing for full paths (e.g. MHR-GS and CMGR) might result in high network overhead. In particular, in all the reviewed protocols a limited number of vehicles was considered applying the route discovery mechanisms in VANET simulation scenario. However, in reality most of vehicles need to apply this mechanism to communicate with other vehicles. Therefore, further research is required to study the effect of having a large number of vehicles, employing route discovery mechanisms, on the routing protocol and network performance.

- Frequent route recalculation might be unnecessary and result in extra network overhead, which affects packet delivery ratio. On the other hand, routes have to be updated in order to adapt to network and traffic conditions. Therefore, route recalculation can be based on timer (i.e. route lifetime DAR-RH) or variations in traffic and network status, to avoid unnecessary recalculation.

- Many of the aforementioned protocols apply carry and forward technique in order to overcome network disconnection problems. However, it must be considered that the carrying vehicle might drive faraway from the right route to destination. Thus, restrictions on packets carrying (e.g. allowed area definition in TARA recovery technique), must be applied specially if the carrying vehicle is approaching a junction or turning (e.g. CMGR).
- The accuracy of routing metrics measurement mechanisms were not evaluated in the aforementioned protocols. The measured metrics were used in the routing process and the evaluation was for the routing performance only. Therefore, in order to provide a more clear view about the performance of the measuring mechanism, such mechanisms must be tested independently and the measured metrics can be compared to the given real values. For instance, measured density can be compared to the real density, which can be given as a simulation parameter.
- Table 1 summarizes and compares the previously mentioned full path TAR protocols main characteristics. In addition, by studying Table 1, the following points can be concluded:

  1. Non of the aforementioned full path TAR protocols considers network load balancing except VLBR.
  2. In ACAR, DAR-RH, VLBR, CMGR and MHR-GS, the scope of traffic condition awareness was limited on the calculated paths or specific geographical areas.
  3. Considering protocol performance evaluation, non of the previously mentioned protocols was compared to a TAR protocol. In other words, the benchmarking was against non TAR protocols. Thus, the performance measurements can not be used to evaluate the level of awareness in these protocols. Moreover, TAR protocols will always have better performance compared to other protocols as they consider traffic conditions. Therefore, comparing improved TAR protocols to exiting TAR protocols should be considered in future research work in order to analyse and evaluate the performance more accurately.

6 Junction-based traffic aware routing protocols

The most recent junction-based TAR protocol are discussed in this section, in terms of routing process, routing metrics measurement, forwarding mechanisms and recovery techniques. Thereafter, protocols' critical issues, challenges and open research problems were highlighted in the challenges and issues to consider section. Afterwards, Table 2 summarizes and compares the characteristics and performance of the following reviewed protocols. In addition, for each junction-

based TAR protocol, Table 2 shows to which protocol it was compared and what was its evaluation in terms of network overhead, end-to-end delay and packet delivery ratio.

6.1 Routing process of junction-based TAR protocols

In [32] junction-based geographical greedy traffic-aware routing protocol (GyTAR) was introduced. GyTAR dynamically chooses intermediate intersections that a packet has to traverse to reach destination, considering both vehicular density of roads and curvemetric distance towards destination. Thus, the highest score road is the highest density and shortest route towards a destination vehicle.

The enhanced GyTAR (EGyTAR,) was introduced in [33] which considered vehicle's direction in calculating road vehicle density. In particular, the calculated density is for vehicles moving towards the destination only, which result in a more accurate estimation for the quality of packet transmission along roads. Accordingly, choosing a road with high vehicle density moving in opposite direction and low density towards destination, was avoided.

A connectivity-aware routing protocol (iCAR) for vehicular ad hoc networks (VANETs) was proposed in [34]. iCAR takes into account real-time vehicular traffic density and packet delivery delay for each road, in order to enable infotainment and interactive applications, as well as multi-hop Internet access in urban environments. Upon reaching a junction, adjacent roads are evaluated based on their routing metrics scores and progress towards destination. Afterwards, the road which leads to more progress towards destination and have the highest density with least delay is selected. Although, the routing overhead of iCAR is reduced in high vehicular density, it is still higher than the overhead of GyTAR and GPSR.

In [35] a vehicle density and network load aware routing (VDLA) protocol was proposed. VDLA utilizes real-time vehicular density and network traffic load collection mechanism. Although, VDLA is a junction-based routing protocol, the major routing decision is made before the packet reaches the junction. Intermediate junctions are dynamically selected based on three routing metrics: distance to destination, real-time vehicle density and network traffic load measurements. After evaluating each adjacent road, weights (i.e.scores) are assigned to them by a junction node. Afterwards, the junction node broadcast hello messages containing the adjacent road segments weights as an extra field.

An enhanced VDLA protocol was proposed in [36], where the accuracy of neighbour tables was increased to improve the performance of geographic greedy forwarding. The author defined two problems caused by traditional hello scheme which has a fixed broadcast interval:

Table 2 Junction-based traffic aware routing

Protocol	GyTAR [32]	EGyTAR [33]	iCAR [34]	VDLA [35] and enhanced VDLA [36]	LSHR [37]
Routing metrics	Distance, vehicular density	Distance, directional vehicular density	Distance, vehicular density and road transmission delay	Distance, vehicular density, vehicular distribution and network load	Distance in 3D scenario and communication delay
Routing metrics measurement	Initiated by a cell leader leaving road	Initiated by a cell leader leaving road	Initiated when the validity period of previous measurement is about to finish	Initiated when required for routing by a vehicle in junction	Initiated when required for routing by a vehicle in junction
Measurement announcement scope	At the opposite Junction	At the opposite Junction	Junction areas at both road end	Junction areas at both road end	Junction areas at both road end
Measurements availability	Announced once	Announced once	Announced once	Announced once	Announced once
Network load balance	No	No	No	Yes	No
Suitable for bidirectional/ unidirectional roads	Bidirectional	Bidirectional	Bidirectional/ unidirectional	Bidirectional/ unidirectional	Bidirectional/ unidirectional
Performance					
Compared to other TAR protocols	No (GSR, LAR [49])	Yes (GSR, GyTAR)	Yes (GyTAR, GPSR)	No (GPCR [45])	No (GPSR, AODV)
Network overhead	Lower	same as GyTAR	Higher specially in low density	Not measured	Not measured
End-to-end delay	Lower	Lower specially in higher	Lower	Lower	Lower
Packet delivery ratio	Higher specially in high density	Higher	Higher	Higher	Higher

– First, the entries of past neighbours are not deleted in time when they move out of the transmission range.
– Second, some vehicles have become neighbours but they are not recorded in the neighbour table as they are waiting for their next hello packet to be sent.

These two kinds of nodes often located close to the transmission range boundary and are referred to as marginal nodes. As greedy forwarding strategy chooses the farthest neighboring node to be the next hop, the inaccurate information of these nodes will directly affect the successful transmission of packets. Thus, the enhanced VDLA calculates a life time for each neighbour record in the neighbour table. In addition, the intermediate vehicle, between two vehicles approaching each others transmission range, will calculate the time when both vehicles can communicate directly then sends this information in a small packet without waiting for the next hello. However, depending on lifetime estimation is critical, as it can be affected by vehicles variable speed and driver behaviour. Moreover, the added enhancement produces extra routing protocol overhead which increases with vehicles higher speed.

2D routing is usually considered in cities environment. However, 2D routing might not be efficient for cities with bridges and complex hierarchy of roads. Therefore, a 3D routing protocol named Link State aware Hierarchical Road routing (LSHR) was introduced in [37]. LSHR utilizes 3D coordinates for vehicles positions aiming at reducing the number of hop count and transmission delay, while increasing packet delivery ratio. As LSHR is a junction based protocol, junction's adjacent roads are weighted based on their routing metrics measurement and the road with the least weight is selected to forward packets. Distance towards destination and road connectivity are the two routing metrics used in LSHR.

6.2 Routing metrics measurement

In GyTAR, each road is dissected into fixed location cells with radius equal to transmission range, whereas the closest car to the cell's center is the group leader. Vehicle density is measured locally within each cell by the group leader, then transferred through group leaders using a collector density packet (CDP). The CDP is generated by a group leader leaving the road as shown in Fig. 6. This CDP generation criteria was used to restrict CDP generation and adapt to the variable vehicular traffic (i.e. high density reduces vehicles speed as well as CDP message generation and vice versa).

After CDP generation, it is sent backward to the first cell leader in the targeted road on the basis of improved greedy strategy, which is explained in Sect. 6.3. Afterwards, the CDP will pass through cell leaders until it reaches the road's end. Once the CDP is received by a group leader, it is updated by including the density of the corresponding cell (i.e. the number of Hello message responses received from group leader's neighbours), then forwards it to the next cell. The road vehicle density measurement is achieved by delivering the CDP to the opposite intersection, where the CDP packet is analyzed to calculate the average vehicle density of cells. Finally, the evaluated road can be given a score based on its vehicular density and distance to destination.

EGyTAR adopts the same mechanism, as GyTAR, for measuring vehicular density. However, EGyTAR only considers vehicles moving towards destination (i.e. in-direction) while measuring vehicular density. Although, high vehicular density moving towards destination leads to higher delivery ratio, vehicles moving in the opposite direction can be a good option when in-direction density is low.

iCAR adopts the same mechanism of sending a collector packet (CP) as GyTAR. However, iCAR does not form fixed cells, instead it forwards the CP along the street using iCAR forwarding mechanism. The delivery of CP at the next intersection indicates the instantaneous connectivity of the road. While forwarding the CP, each forwarder counts the number of vehicles located between itself and the vehicle chosen as the next forwarder, then add it to the number of vehicles recorded in CP. The CP generation time and the number of traversed nodes are also recorded in CP. Before sending the packet to the next hop, the forwarder calculates the validity period and updates the lifetime field if calculated value is lower. The validity period is to define the validity of measured data of a given road segment. In other words, it aims at predicting the time at which a disconnection may occur. Once a CP is received at the center of the next road

Fig. 6 Generating and forwarding a CDP packet in GyTAR
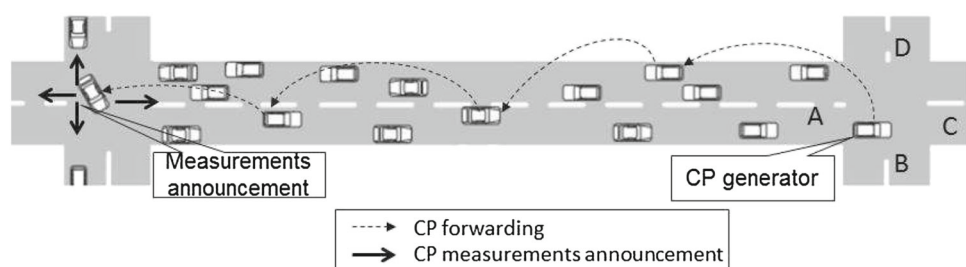
intersection, the closest vehicle to the center of the intersection, is responsible of generating the updated score. Then the updated score with its validity time is announced across the intersection using beacon messages, and sent back to the location where the CP packet was created. In case of not receiving back the CP, the road score is set to zero as it is considered disconnected.

Similarly, VDLA routing metrics (i.e. Vehicular density and network load) are measured using the collector packet mechanism. Once a vehicle reaches a junction it initiates the measurement process by sending a CP to each adjacent road. To measure network traffic load, the author considered that the number of packets stored in the buffer queue of a node can be used to evaluate the node load status. The buffer queue length and the number of neighbours of the initiator vehicle is first included in the CP. In addition, the minimum neighbour number (MNN) field, which is used to evaluate vehicle distribution along road, is initialized by the value of the initiator's number of neighbours. Upon receiving the forwarded CP, each node will add the number of its right side neighbours to the number of neighbours in the CP, also the length of buffer queue is added similarly. In addition, each forwarding node will compare its number of neighbours with the MNN; if the MNN field has greater value, it will be updated with the number of neighbours of the current forwarder. Afterwards, the updated packet will be forwarded till it reaches next junction. After processing the CP packet at next junction, another CP will be initiated and sent backward to the first junction in the same described mechanism. Afterwards, the initiator vehicle will get vehicle density and traffic load measurement for all adjacent road segments. Thereby, road evaluation can be carried out using the collected data, then scores can be given to adjacent roads. However, if a vehicle has already received a CP before it enters the junction, then it will start a timer for one second then generates a new CP. Thereby, network overhead caused by CP is reduced in dense situations.

The measurement mechanism of iCAR is also used in LSHR. However, LSHR depends on greedy forwarding to forward CPs. In particular, when a node at intersection receives a data packet, if it does not have a sorting table it will generate one. The sorting table includes the progressive distance towards destination through adjacent roads and their communication delays. This table is generated by sending collector packets towards adjacent intersections and recording the sending time in the CP. The last node which receives the CP will calculate the communication delay of the road, then the result is broadcasted with the corresponding road ID to neighbours. The communication delay for adjacent roads is recorded and sorted in the sort table, which is also broadcasted to other nodes in the junction area. Finally, each road is given a score (i.e. weight) based on distance and connectivity (i.e. communication delay), which are the two routing metric used in LSHR.

6.3 Forwarding mechanisms

GyTAR and EGyTAR utilize improved greedy mechanism to forward packets towards junctions. First, the forwarding vehicle predicts the position of each neighbour using its corresponding motion information (velocity, direction, and position), which was obtained by hello messages and saved in the neighbour table. Then, the neighbour with the closest predicted position to destination junction is selected as the next-hop.

For next-hop selection in iCAR, the improved greedy forwarding mechanism of GyTAR and EGYTAR is adopted to forward packets between junctions. However, the selected next-hop is not only the closest to destination, but also which has the highest received signal strength indication (RSSI).

In VDLA and enhanced VDLA, the classical greedy forwarding is used to forward packets between junctions. However, in these two protocols routing decisions can be done before reaching a junction. In such a case, if the next road segment lies on the same direction as the current one, the packet will be forwarded directly in greedy mode; otherwise, the packet will be forwarded to the junction node in order to be transmitted towards a different direction.

LSHR calculates the virtual distance for all neighbour nodes which are closer to the destination than current node. Here, the virtual distance refers to the largest two-hop distance, which equals to the distance between the current node and its neighbour node plus the distance between neighbour node and its closest neighbour to destination. Finally, the neighbour with the largest virtual distance is chosen as the next hop.

6.4 Recovery techniques

The carry and forward technique was adopted in GyTAR, EGyTAR, LSHR and iCAR. In such a technique, a vehicle will carry the packet to next junction or until another vehicle closer to destination junction enters or reaches its transmission range. However, for VDLA and enhanced VDLA no recovery technique was proposed.

6.5 Challenges and issues to consider

Junction-based TAR protocols critical issues and challenges are presented and discussed in this section. Moreover, a number of open research problems are highlighted for further investigation and future work in the following points:

- Junction-based TAR protocols depend on weighting adjacent roads (i.e. giving scores to roads), to select the best road for forwarding. Weighting is usually done using weighting factors, such as α_1, α_2 and α_3 in Eq. 1. In order to give a weighting score for roads in the range

[0, 1] the summation of weighting factors has to equal 1. Consequently, choosing equal values for weighting factors (e.g. $\alpha_1 = \alpha_2 = \alpha_3 = 0.33$)results in giving the same effect and priority, for each of the utilized routing metrics (e.g. distance, density and connectivity), in the routing decision. Nevertheless, there is no guarantee that equal weighting factors will result in the best routing performance. In particular, giving higher weighting factors to some routing metrics (e.g. distance) may result in better routing performance. However, calculating the optimum weighting factor for each routing metric in the aforementioned protocols was not carried out. Although, in GyTAR the author defined optimum weighting factors values, these values were derived based on excessive simulations not mathematical analysis. Therefore, this issue needs further research and investigation in order to improve the performance of junction-based TAR protocols.

$$RoadWeight = \alpha_1(Distance)$$
$$+ \alpha_2(Density) + \alpha_3(Connectivity)$$
$$(1)$$

- Routing metrics measurement process introduces extra network overhead, which results from generating and forwarding collector packets. Therefore, this process need to be done when it is necessary only. However, initiating the measurement process when the measurement is requested for routing decisions, such as in LSHR and VDLA, might result in routing and delivery delays. Thus, measurement results must be always updated and available at junctions with the lowest network overhead. Therefore, adapting the initiation of measurement process to network status, such as iCAR, GyTAR and EGyTAR, results in avoiding unnecessary network overhead produced by the measurement process.
- Assigning a validity time for the routing metrics measurement result, as in iCAR, is important to prevent using outdated measurements. In addition, calculating a validity time based on network conditions is more efficient than fixed validity intervals, as it can adapt to variable network conditions.
- Road transmission delay was considered as an indicator for network connectivity in some protocols (e.g. iCAR, LSHR). However, transmission delay is not only caused by network disconnections, it can result from network load as well. Therefore, to measure real road connectivity a different indication, such as inter-vehicle link life time, can be used.
- Forwarding a collector packet using greedy forwarding between junctions may result in loosing the packets because of neighbours mobility. Subsequently, loosing a collector packet gives a false indication of network dis-

connection along that road. Therefore, the improved forwarding mechanisms, such as the one used in iCAR and GyTAR are preferable.
- Accuracy in measuring routing metrics provides higher awareness of road and network conditions, which makes routing protocols more adaptive to VANET dynamic environment. However, in all the aforementioned protocols no evaluation was provided for the accuracy of routing metrics measurement process, except for the density measurement evaluation in GyTAR. Therefore, evaluating measurement process accuracy is essential in future research work as it gives an indication on the level of awareness in TAR protocols.
- Keeping neighbour tables updated is essential for accurate routing metrics measurement. However, depending on predicting neighbour positions is not always accurate. On the other hand, increasing the frequency of beacons or hello messages is not recommended specially in high density situation. Therefore, utilizing beacon frequency adaptation mechanisms can enhance TAR protocols performance [38,39].
- The previously mentioned junction-based TAR protocols depend on carry and forward technique, in order to deliver packets in case of network disconnection. However, it might not be efficient to continue in the same route. Therefore, other solutions for recovery must be considered for future research work, such as stepping backward to a previous junction and re-evaluating roads to update routing decision.
- A performance and characteristics comparison for the previously reviewed junction-based TAR protocols is provided in Table 2. Accordingly, the following points can be concluded based on Table 2 inspection:

1. GyTAR and EGyTAR do not consider unidirectional roads as the scope of result announcement will be limited to one junction only. Obviously, routing measurement is initiated at one road end and result is announced at the other road end. Thus, in unidirectional roads only one road end will have the measurements. In addition, no mechanism was introduced for sending back results to the initiation junction, such as the one used in iCAR, VDLA and LSHR.
2. Three of the reviewed protocols performance was not compared to TAR protocols, which does not give a clear evaluation for the level of their traffic awareness. Thus, comparing traffic aware routing protocols to each other is necessary in order to get a more significant evaluation.
3. It is obvious that measurement results are announced once in all reviewed protocols. Thus, if there was no vehicles in the junction area at announcement time or after announcement all vehicles left junction, then

measurement results are lost. Therefore, this issue need to be considered in future improvements of junction-based TAR protocols.

4. The junction-based TAR protocols which consider network load while making routing decisions are VDLA and enhanced VDLA only. However, taking network load into consideration is important to prevent congestion and network performance degradations.

5. Network overhead was not measured in the evaluation of some junction-based TAR protocols. However, evaluating junction-based protocols performance in terms of network overhead is necessary.

7 Hybrid traffic aware routing protocols

The most recent hybrid TAR protocols are explored in this section, based on their routing process, routing metrics measurement, forwarding mechanisms and recovery techniques. Afterwards, challenges, critical issues to consider and future research problems were discussed in the challenges and issues to consider section. Finally, Table 3 presents a comparison between the reviewed hybrid protocols in terms of characteristics and performance.

7.1 Routing process of hybrid TAR protocols

In [8] a Hybrid Traffic-Aware Routing (HTAR) protocol was proposed, which selects a vehicle, called Junc-Tracker, at each junction to collect traffic information for adjacent roads. First, HTAR apply Dijkstra algorithm to find the shortest path, whereas road weights are assigned based on vehicular traffic density and network traffic load metrics. Afterwards, data packets are forwarded based on the calculated best path. However, the forwarding path is dynamically adjusted at each junction based on updated traffic information.

HTAR employs Dijkstra algorithm and needs traffic information about all possible full paths between source and destination. Therefore, the information obtained at each Junc-Tracker will be periodically broadcasted within the junction area and disseminated to other adjacent Junc-Tracker. As a result, high extra network load will be generated, which may degrade network performance.

A similar protocol called Junction-Based Traffic-Aware Routing (JTAR) protocol was introduced in [10]. In order to reduce the time required to collect traffic information and to increase it's accuracy, JTAR introduced the concept of collector packet merge. In particular, when two collector packets meet at a certain node in the traversed road, they will exchange traffic information about their unvisited part of the

Table 3 Hybrid traffic aware routing protocols

| Protocol	HTAR [8]	JTAR [10]	ASAR [40]
Routing metrics	Distance, vehicular density and network load	Distance, vehicular density and network load	Distance, vehicular density, velocity and transmission delay
Routing metrics measurement	Periodically initiated by Junc-tracker	Periodically initiated by Junc-tracker and adopt probe-merge	Fixed equipment record vehicular density and velocity based on vehicles mobility
Measurement announcement scope	Junc-tracker junction area and to adjacent Junc-trackers	Junc-tracker junction area and to adjacent Junc-trackers	To all fixed equipment along the path from destination to source
Measurements availability	Periodically announced	Periodically announced	When requested by destination node
Network load balance	Yes	Yes	No
Awareness scope	Current and adjacent junctions	Current and adjacent junctions	Along the path from source to destination
Performance			
Compared to other TAR protocols	Yes (GyTAR, GSR)	Yes (GyTAR, HTAR)	No (GPSR, GSR)
Network overhead	Not measured	Not measured	Higher
End-to-end delay	Not measured	Lower	Lower
Packet delivery ratio	Higher	Higher	Higher

At time t1

At time t2
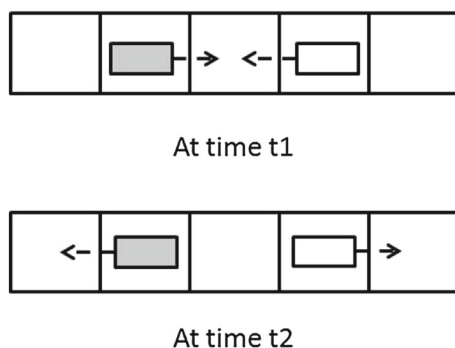
Fig. 7 Two IC packets meeting (probe-merge method)

road, then return back to their original junction, as shown in Fig. 7. In this case the collector packet do not have to traverse the whole road. JTAR was compared to GyTAR and HTAR and it gives higher delivery ratio than both, as well as lower end-to-end delay. However, no measurement was provided for routing protocol overhead.

On the other hand, the adaptive state aware routing (ASAR) protocol, which was introduced in [40], utilizes fixed equipments laid at each junction to collect real-time traffic information. ASAR adaptively selects an optimal path based on an estimated transmission delay model that takes into account real-time vehicle density and average speed. When a source node wants to send data, it will inform the destination node through cellular network. Afterwards, the destination node informs the fixed equipment in the closest junction to update the estimated transmission delay. The fixed equipment will broadcast update messages to its adjacent junctions periodically by cellular network. The broadcast range depend on the size of the city and the number of junctions.

For the purpose of obtaining global optimum path, when a data packet reaches a junction it will be forwarded for the fixed equipment. Subsequently, the fixed equipment will calculate the optimum path towards the destination based on two types of transmission delays: (1) the estimated transmission delays for adjacent roads, which was done by this fixed equipment; (2) the received estimated transmission delays from other adjacent fixed equipment. Afterwards, the optimum path will be adjusted at each junction, in order to adapted to changes in network conditions. Although, ASAR achieved higher packet delivery ratio and lower end-to-end delay than GPSR and GSR, it's generated network overhead was slightly higher specially in high vehicular density or velocity.

7.2 Routing metrics measurement

HTAR and JTAR adopt the mechanism of collecting network information by sending collector packets between adjacent junctions. The Junc-Tracker vehicle is responsible for sending, receiving and processing collector packets as well as spreading data to neighbours and Junc-Trackers at adjacent

junctions. The mechanism of dividing the road to cells then collecting traffic information from cell headers , which was used in GyTAR, is applied in both HTAR and JTAR.

Since Junc-tracker are responsible for generating and processing collector packets, the author in HTAR proposed an algorithm for selecting a Junc-tracker, which was adopted in JTAR as well. The algorithm guarantees the uniqueness of Junc-tracker at each junction, as well as the selection of the longest life Junc-Tracker. In the Junc-Tracker selection algorithm a node can be in one of four states: monitor, compete, spread, and sleep. The node state is based on its current location and status. The algorithm can be summarized in the following points:

1. If a node is not in a junction, it is in sleep state.
2. When a node enters the junction it will change its state from sleep to monitor and sets a timer T_m.
3. During T_m, if the node did not receive any of the Junc-Tracker periodic update messages, it will change to compete state, otherwise it will remain in the monitor state and resets the timer.
4. Upon entering the compete state, the node sets a timer T_c. When T_c expires, the node sends a declaration message to become the new Junc-Tracker and enter the spread state.
5. Receiving a message from another Junc-Tracker while being in the compete state will change the node status to monitor.
6. If the Junc-Tracker is leaving the junction, it chooses the successor with the longest time to live in junction to inherit its up-to-date weighting information.

Vehicular density and network load are the two routing metrics measured by the collector packets in HTAR and JTAR. Network load is measured by observing channels, whereas each node calculates the ratio of channel busy time during each measurement period. Subsequently, the highest value for channel load in a cell is attached to the collector packet. Vehicular density is measured by collecting the number of neighbours for each traversed cell on road, then average vehicular density is calculated in the same mechanism adopted in GyTAR [32]. However, JTAR apply the probe-merge method to decrease delay time and overhead generated from information collection. The probe-merge method can be applied when two collector packets, generated from opposite road sides, meet at a certain node, as shown in Fig. 7. Consequently, each collector packet can obtain information about the other non-traversed part of the road from the opposite direction collector packet. In this case, the collector packet will be forwarded back after information exchange without reaching the destined next junction.

On the other hand, ASAR depends on the fixed equipment to collect information about its two routing metrics: vehicular density and velocity. When vehicles leave or enter a section,

they send their ID to the fixed equipment in the junction. By cooperating with other fixed equipment at adjacent junctions, the number of vehicles at each road and their average velocity can be calculated. In addition, the fixed equipment will calculate the transmission delay for adjacent roads based on vehicle density and average vehicular velocity. Afterwards, the fixed equipment will exchange the calculated transmission delay with other fixed equipment at adjacent junctions.

7.3 Forwarding mechanisms

Between two consecutive junctions on the path, geographical greedy forwarding is used to select forwarding nodes on the road in HTAR, JTAR and ASAR. However, in ASAR , when reaching a junction a packet has to be forwarded to the fixed equipment. On the contrary, HTAR and JTAR do not restrict packets to be forwarded to certain nodes at junctions.

7.4 Recovery techniques

In HTAR and JTAR, if a road node can not find a next forwarding node it will carry the packet until a suitable next hop becomes available. On the other hand, if a junction node can not find any next forwarding node, it will re-compute the routing path then look for a suitable next forwarder again. Similarly, ASAR apply the concept of carry and forward when a node can not find a suitable next forwarder. However, when a fixed node can not find a suitable next node to forward the packet, it will keep storing the packet until a suitable node becomes available.

7.5 Challenges and issues to consider

A discussion for hybrid TAR protocols challenges and critical issues is provided in this section. In addition, some open research problems have been highlighted in the following points:

- In the aforementioned hybrid protocols, the full path is recalculated at each junction to adapt to network changes. Afterwards, the full path is attached to data packet header before forwarding, which increases consumption of network resources and might result in extra delay.
- Disseminating network information (i.e. collected at Junc-Tracker or fixed equipment), results in extra network overhead specially when large area is considered for broadcast.
- Although hybrid traffic aware routing protocols select the globally optimum next junction, this result in higher network overhead compared to junction based protocols (e.g. GyTAR), which selects the locally optimum next junction. Therefore, solutions for reducing network overhead must be considered in future research.

- In order to reduce the routing protocol network overhead the recalculation of paths must be done only when required (i.e. when network conditions change). However, this issue was not investigated in the aforementioned protocols.
- Generating periodic collector packets is not efficient as it increases network overhead. However, adapting collector packet generation based on network and traffic conditions might result in lower network overhead.
- A comparison for the performance and characteristics of the previously reviewed hybrid TAR protocols is provided in Table 3. Accordingly, the following points can be concluded:

  1. Hybrid TAR protocols have more traffic awareness scope compared to junction-based protocols.
  2. In some of the reviewed hybrid protocols, the network overhead was not measured. However, network overhead is one of the critical issues in hybrid TAR protocols.

8 Node-based traffic aware routing protocols

This section explores the most recent node-based traffic aware routing protocols. In addition, critical issues, challenges and open research problems were discussed in the challenges and issues to consider section. Finally, Table 4 compares the characteristics and performance of the reviewed protocols.

8.1 Routing process of node-based TAR protocols

A Multi-Metric, Map-aware Routing (MMMR) protocol was introduced in [3] for urban area VANETs. MMMR enhances the packet forwarding mechanism in geographical routing by utilizing four metrics in the process of selecting next forwarding nodes: vehicle density, trajectory, distance to destination and available bandwidth. Map awareness is achieved by taking into account the presence of buildings while selecting the next forwarding node. The trajectory of a node describes its moving direction and speed. The available bandwidth metric is used to eliminate data congestion problems.

Although the MMMR enhanced forwarding is more reliable and result in higher delivery ratio, it has some drawbacks. Based on simulation results of MMMR, it is clear that the number of hops and end-to-end delay increased, specially in low vehicular densities. However, if reliability and high throughput is the main concern, then enhanced forwarding mechanism of MMMR is sufficient.

In [41] a Link State aware Geographic Opportunistic routing protocol (LSGO) was introduced. The protocol aims at

increasing packet delivery rate and improving data transmission reliability. First, LSGO selects a candidate nodes set based on the vehicles geographic location and the links quality. If the network state is good, the link between any two nodes is relatively stable, so the set of candidate nodes is small. Second, a priority scheduling algorithm, which prioritizes forwarder nodes based on timer scheduling method, was utilized.

Opportunistic routing can enhance the reliability of a routing protocol by making full use of the additional backup links with other neighbours. However, this might be at the cost of creating extra routing overhead as it depend on multicast to a group of neighbours. This was clear in the simulation results of LSGO as it has higher routing protocol overhead compared to GyTAR. Nevertheless, LSGO provides higher network throughput which drops in sparse situations only.

8.2 Routing metrics measurement

In MMMR, to calculate the available bandwidth each node estimates the percentage of idle time on the common wireless medium. Then, the estimated value is included in the hello messages. Thereafter, the available bandwidth estimator (ABE) algorithm [42], was used to estimate the available bandwidth of a link based on the transmitter and receiver estimated idle time. For density measurement, each node measures its local density based on counting the number of available neighbours. Next, density measurement results are exchanged with neighbours, along with direction and speed, via hello messages. In order to enhance the accuracy of neighbouring tables, only neighbours currently in the transmission range and their received power is above a certain threshold are considered stable neighbours. The neighbour power measurement was obtained from the periodic hello messages.

LSGO measures link's quality using the enhanced expected transmission count (ETX) metric, which considers mobility to adapt to network dynamicity. Depending on periodic hello messages, LSGO calculates the ETX of a link. In particular, each node records the number of received hello packets and divide it by the number of packets that should have been received in a certain time interval. Thereby, the link ETX for one-way transmission rate can be calculated.

8.3 Forwarding mechanisms

To send a packet using MMMR, the current node has to find the optimal next forwarding node from its neighbors list. Therefore, each neighbor is assigned a score based on the four evaluation metrics. Then the node with the highest score is selected as the next forwarder. However, each neighbour considered in the evaluation process has

to satisfy the following conditions: (1) Be within current node transmission range; (2) In the line of site of current forwarder; (3) Closer to destination than current forwarder.

On the other hand, LSGO selects a candidate set of neighbours after calculating the link ETX for each neighbour. The set of neighbours contains the neighbours which are more close to destination and have the potential of sending data to the next hop in the required transmission rate. In other words, it calculates the probability that source sends data to the next hop successfully using that candidate set. Afterwards, the current node will record the candidate nodes IDs and their priority numbers in the packet header.

LSGO uses timer-based priority scheduling algorithm, whereas priority is given based on distance to destination, and the ETX value for the link between current node and candidate node. The highest priority node sends the packet directly when it is received, while the lower priority nodes need to set a timer. However, if the timer expires and a higher-priority node is not transmitting, the lower priority node can send the packet.

8.4 Recovery techniques

In order to avoid routing loops caused by perimeter mode, MMMR stores packets in a local buffer of the current relaying node when there is no neighbour that satisfies all next forwarder conditions. Then the carrying node periodically looks for a neighbour that can satisfy the conditions to be the next forwarding node. On the other hand, as LSGO employ opportunistic routing, alternative links are always available to forward packets unless network disconnections occur. In this case, no recovery technique was proposed for LSGO.
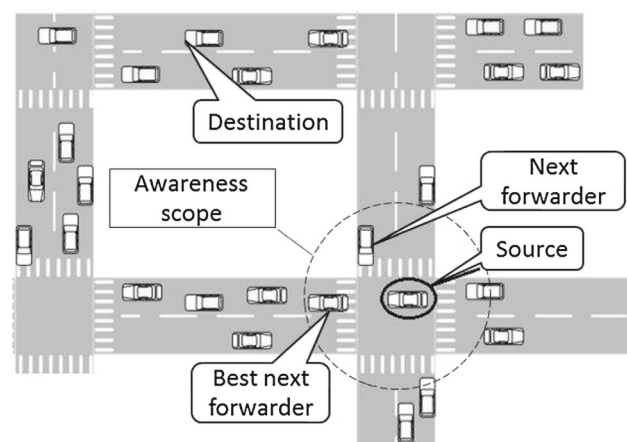

Fig. 8 Node-based limited awareness scope

Table 4 Node-based traffic aware routing protocols

Protocol	MMMR [3]	LSGO [41]
Routing metrics	Vehicular density, trajectory, distance and available bandwidth	Distance and link state
Routing metrics measurement	Periodic based on hello messages	Periodic based on hello messages
Measurement announcement scope	One hop neighbours	No announcement
Network load balance	Yes	Yes
Awareness scope	Local (one hop)	Local (one hop)
Performance		
Compared to other TAR protocols	No (AODV, GBSR [50], GPSR, I-GPSR [51])	Yes (GyTAR, GpsrJ+ [52])
Network overhead	Low	Higher
End-to-end delay	Higher specially in low density	Higher than GpsrJ+
Packet delivery ratio	Higher	Higher

8.5 Challenges and issues to consider

In node-based traffic-aware routing there are a number of critical issues that need to be considered in order to improve the performance, which are highlighted in the following points:

- In opportunistic routing, higher number of neighbours gives more opportunity for packet forwarding, this is true until reaching a certain density, above which the frequency of collisions increases and result in packet drops [3].
- Although, node-based traffic aware routing enhances the forwarding mechanism, it has no awareness of network status on street or network level. Thus, it may not select the shortest or the best path. For instance, in Fig. 8 the selected next forwarder is in a road which has network disconnection. However, choosing the best forwarder was not possible because the awareness scope of the current node is limited.
- Based on Table 4, opportunistic routing usually results in higher network overhead as it uses multi-casting.

9 Conclusion

This paper explains the concept of traffic aware routing and reviews the most recent TAR protocols in VANET. The main objective is to highlight critical issues related to traffic awareness and their effect on packet routing. As traffic aware routing was introduced to overcome the shortcomings of geographical routing, this paper explains the concept of geographical routing and its main limitations. Afterwards, traffic aware routing concepts and classification was explained in detail. In particular, traffic aware routing was classified into four categories: full path TAR protocols, junction-based TAR protocols, hybrid TAR protocol and node-based TAR protocols. For each category, the routing process, routing metrics measurement, forwarding mechanisms and recovery techniques were inspected. In addition, special effort has been devoted to discuss critical issues, challenges and open research problems for each category in the "Challenges and issues to consider" sections.

Obviously, traffic aware routing protocols are more suitable for VANET environment as they adapt to network and traffic variations. However, the main issue in such protocols is how to get sufficient level of awareness with the least network overhead, in order to get high packet delivery rate and low end-to-end delay. Thus, the routing metrics measurement process has to be improved. In addition, the way of utilizing routing metrics measurement in routing packets needs more enhancement. Moreover, packet forwarding mechanisms in TAR protocols must be more aware of network and traffic conditions. Finally, to improve routing protocols performance more efficient recovery techniques, specially designed for VANETs and have consideration for traffic conditions, are essential.

References

1. Hossain, E., Chow, G., Leung, V. C., McLeod, R. D., Mii, J., Wong, V. W., & Yang, O. (2010). Vehicular telematics over heterogeneous wireless networks: A survey. *Computer Communications,*

33(7), 775–793. doi:10.1016/j.comcom.2009.12.010. http://www.sciencedirect.com/science/article/pii/S0140366410000022.

2. Seredynski, M., & Bouvry, P. (2011). A survey of vehicular-based cooperative traffic information systems. In *2011 14th international IEEE Conference on, intelligent transportation systems (ITSC)* (pp. 163–168). doi:10.1109/ITSC.2011.6083055.

3. Tripp-Barba, C., Urquiza-Aguiar, L., Igartua, M. A., Rebollo-Monedero, D., de la Cruz Llopis, L. J., Mezher, A. M., et al. (2014). A multimetric, map-aware routing protocol for VANETs in urban areas. *Sensors*, *14*(2), 2199–2224.

4. Asefi, M., Céspedes, S., Shen, X., & Mark, J. W. (2011). A seamless quality-driven multi-hop data delivery scheme for video streaming in urban VANET scenarios. In *2011 IEEE international conference on (IEEE, 2011), communications (ICC)* (pp. 1–5).

5. Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H., & Zedan, H. (2014). A comprehensive survey on vehicular ad hoc network. *Journal of Network and Computer Applications, 37*(0), 380–392. doi:10.1016/j.jnca.2013.02.036. http://www.sciencedirect.com/science/article/pii/S108480451300074X.

6. Zeadally, S., Hunt, R., Chen, Y. S., Irwin, A., & Hassan, A. (2012). Vehicular ad hoc networks (VANETS): Status, results, and challenges. *Telecommunication Systems*, *50*(4), 217–241. doi:10.1007/s11235-010-9400-5.

7. Hanshi, S. M., & Kadhum, M. M. (2013). Geographic routing protocol issues in vehicular ad hoc networks. In *2013 IEEE international conference on (IEEE, 2013), RFID-technologies and applications (RFID-TA)* (pp. 1–7).

8. Lee, J. W., Lo, C. C., Tang, S. P., Horng, M. F., & Kuo, Y. H. (2011). A hybrid traffic geographic routing with cooperative traffic information collection scheme in VANET. In *2011 13th international conference on, advanced communication technology (ICACT)* (pp. 1496–1501).

9. Fonseca, A., & Vazão, T. (2013). Applicability of position-based routing for VANET in highways and urban environment. *Journal of Network and Computer Applications*, *36*(3), 961–973.

10. Lo, C. C., & Kuo, Y. H. (2013). Junction-Based Traffic-Aware Routing scheme for vehicular ad hoc networks. In *2013 IEEE 24th international symposium on, personal indoor and mobile radio communications (PIMRC)* (pp. 3001–3005). doi:10.1109/PIMRC.2013.6666661.

11. Saleet, H., Basir, O., Langar, R., & Boutaba, R. (2010). Region-based location-service-management protocol for VANETs. *IEEE Transactions on Vehicular Technology*, *59*(2), 917–931. doi:10.1109/TVT.2009.2033079.

12. Darwish, T., & Bakar, K. A. (2015). Traffic density estimation in vehicular ad hoc networks: A review. *Ad Hoc Networks, 24*(Part A(0)), 337–351. doi:10.1016/j.adhoc.2014.09.007. http://www.sciencedirect.com/science/article/pii/S1570870514001966.

13. Cadger, F., Curran, K., Santos, J., & Moffett, S. (2013). Communications surveys & tutorials. *IEEE*, *15*(2), 621–653.

14. Karp, B., & Kung, H. T. (2000). GPSR: Greedy perimeter stateless routing for wireless networks. In *Proceedings of the 6th annual international conference on Mobile computing and networking (ACM, 2000)* (pp. 243–254).

15. Stanica, R., Chaput, E., & Beylot, A. L. (2011). Local density estimation for contention window adaptation in vehicular networks. In *2011 IEEE 22nd international symposium on, personal indoor and mobile radio communications (PIMRC)* (pp. 730–734). doi:10.1109/PIMRC.2011.6140062.

16. Giudici, F., & Pagani, E. (2005). Spatial and traffic-aware routing (STAR) for vehicular systems. In L. Yang, O. Rana, B. Di Martino, & J. Dongarra (Eds.), *High performance computing and communications. Lecture notes in computer science* (Vol. 3726, pp. 77–86). Berlin, Heidelberg: Springer. doi:10.1007/11557654_11.

17. Al-Shugran, M., Ghazali, O., Hassan, S., Nisar, K., & Arif, A. S. M. (2013). A qualitative comparison evaluation of the greedy for-

warding strategies in Mobile Ad Hoc Network. *Journal of Network and Computer Applications*, *36*(2), 887–897.

18. Bilal, S. M., Bernardos, C. J., & Guerrero, C. (2013). Position-based routing in vehicular networks: A survey. *Journal of Network and Computer Applications, 36*(2), 685–697. doi:10.1016/j.jnca.2012.12.023. http://www.sciencedirect.com/science/article/pii/S1084804512002706.

19. Nasir, M. K., Sohel, M. K., Rahman, M. T., & Islam, A. K. (2013). A review on position based routing protocol in vehicular adhoc network. *American Journal of Engineering Research*.

20. Li, G., Wang, W., Yao, X., & Chen, W. (2013). SOBP: A sender-designated opportunistic broadcast protocol for VANET. *Telecommunication Systems*, *53*(4), 453–467. doi:10.1007/s11235-013-9709-y.

21. Yang, Q., Lim, A., & Agrawal, P. (2008). Connectivity aware routing in vehicular networks. In *Wireless communications and networking conference, 2008. WCNC 2008. IEEE* (pp. 2218–2223). doi:10.1109/WCNC.2008.392.

22. Yang, Q., Lim, A., Li, S., Fang, J., & Agrawal, P. (2010). ACAR: Adaptive connectivity aware routing for vehicular ad hoc networks in city scenarios. *Mobile Networks and Applications*, *15*(1), 36–60. doi:10.1007/s11036-009-0169-2.

23. Mouzna, J., Uppoor, S., Boussedjra, M., & Pai, M. (2009). Density aware routing using road hierarchy for vehicular networks. in *IEEE/INFORMS international conference on, service operations, logistics and informatics, 2009. SOLI '09* (pp. 443–448). doi:10.1109/SOLI.2009.5203974.

24. Benslimane, A., Barghi, S., & Assi, C. (2011). An efficient routing protocol for connecting vehicular networks to the Internet. *Pervasive and Mobile Computing*, *7*(1), 98–113.

25. Shafiee, K., & Leung, V. C. M. (2011). Connectivity-aware minimum-delay geographic routing with vehicle tracking in VANETs. *Ad Hoc Network*, *9*(2), 131–141. doi:10.1016/j.adhoc.2010.06.003.

26. Wu, D., Luo, J., Li, R., & Regan, A. (2011). Geographic load balancing routing in hybrid vehicular ad hoc betworks. In *2011 14th international IEEE conference on, intelligent transportation systems (ITSC)* (pp. 2057–2062). doi:10.1109/ITSC.2011.6083019.

27. Hashemi, H., & Khorsandi, S. (2012). Load balanced VANET routing in city environments. In *2012 IEEE 75th, vehicular technology conference (VTC Spring)* (pp. 1–6). doi:10.1109/VETECS.2012.6240114.

28. Yu, H., Yoo, J., & Ahn, S. (2013). A VANET routing based on the real-time road vehicle density in the city environment. In *2013 fifth international conference on, ubiquitous and future networks (ICUFN)* (pp. 333–337). doi:10.1109/ICUFN.2013.6614836.

29. Xiang, Y., Liu, Z., Liu, R., Sun, W., & Wang, W. (2013). GeoSVR: A map-based stateless VANET routing. *Ad Hoc Networks, 11*(7), 2125-2135. doi:10.1016/j.adhoc.2012.02.015. http://www.sciencedirect.com/science/article/pii/S1570870512000455.

30. Cao, T., Zhang, X., Kong, L., Liu, X. Y., Shu, W., & Wu, M. Y. (2013). Road traffic density estimation in vehicular networks. Traffic aware routing in urban vehicular networks. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE (IEEE, 2013)* (pp. 2004–2009).

31. Oliveira, R., Furtadoa, A., Bernardoa, L., Pintoa, P., et al. (2013). Improving path duration in high mobility vehicular ad hoc networks. *Ad Hoc Networks*, *11*(1), 89–103.

32. Jerbi, M., Senouci, S. M., Rasheed, T., & Ghamri-Doudane, Y. (2009). Towards efficient geographic routing in urban vehicular networks. *IEEE Transactions on Vehicular Technology*, *58*(9), 5048–5059. doi:10.1109/TVT.2009.2024341.

33. Bilal, S., Madani, S. A., & Khan, I. (2011). Enhanced junction selection mechanism for routing protocol in VANETs. *The International Arab Journal of Information Technology*, *8*(4), 422–429.

34. Alsharif, N., Cespedes, S., & Shen, X. (2013). iCAR: Intersection-based connectivity aware routing in vehicular ad hoc networks. In *2013 IEEE international conference on, communications (ICC)* (pp. 1736–1741). doi:10.1109/ICC.2013.6654769.

35. Zhao, C., Li, C., Zhu, L., Lin, H., & Li, J. (2012). A vehicle density and load aware routing protocol for VANETs in city scenarios. in *2012 international conference on, wireless communications signal processing (WCSP)* (pp. 1–6). doi:10.1109/WCSP.2012.6542825.

36. Li, C., Zhao, C., Zhu, L., Lin, H., & Li, J. (2013). Geographic routing protocol for vehicular ad hoc networks in city scenarios: a proposal and analysis. *International Journal of Communication Systems*. doi:10.1002/dac.2602.

37. He, Y., Li, C., Han, X., & Lin, Q. (2014). A link state aware hierarchical road routing protocol for 3D scenario in VANETs. In R. H. Hsu & S. Wang (Eds.), *Internet of vehicles technologies and services. Lecture notes in computer science* (Vol. 8662). Berlin: Springer. doi:10.1007/978-3-319-11167-4_2.

38. Ghafoor, K. Z., Lloret, J., Bakar, K. A., Sadiq, A. S., & Mussa, S. A. B. (2013). Beaconing approaches in vehicular ad hoc networks: A survey. *Wireless personal communications*, *73*(3), 885–912.

39. Ghafoor, K. Z., Bakar, K. A., van Eenennaam, M., Khokhar, R. H., & Gonzalez, A. J. (2013). A fuzzy logic approach to beaconing for vehicular ad hoc networks. *Telecommunication Systems*, *52*(1), 139–149.

40. Qin, H., Yu, C., Guo, L., & Zhou, Y. (2013). Adaptive state aware routing protocol for vehicle ad hoc network in urban scenarios. In *2013 10th IEEE international conference on (IEEE, 2013), control and automation (ICCA)* (pp. 1376–1381).

41. Cai, X., He, Y., Zhao, C., Zhu, L., & Li, C. (2014). LSGO: Link state aware geographic opportunistic routing protocol for VANETs. *EURASIP Journal on Wireless Communications and Networking*, *2014*(1), 96. doi:10.1186/1687-1499-2014-96.

42. Sarr, C., Chaudet, C., Chelius, G., & Lassous, I. (2008). Bandwidth estimation for IEEE 802.11-based ad hoc networks. *IEEE Transactions on Mobile Computing*, *7*(10), 1228–1241. doi:10.1109/TMC.2008.41.

43. Zhao, J., & Cao, G. (2008). VADD: Vehicle-assisted data delivery in vehicular ad hoc networks. *IEEE Transactions on Vehicular Technology*, *57*(3), 1910–1922. doi:10.1109/TVT.2007.901869.

44. LeBrun, J., Chuah, C. N., Ghosal, D., & Zhang, M. (2005). Knowledge-based opportunistic forwarding in vehicular wireless ad hoc networks. in *Vehicular technology conference, 2005. VTC 2005-Spring. 2005 IEEE 61st, vol. 4 (IEEE, 2005)* (vol. 4, pp. 2289–2293).

45. Lochert, C., Mauve, M., Füßler, H., & Hartenstein, H. (2005). Geographic routing in city scenarios. *ACM SIGMOBILE Mobile Computing and Communications Review*, *9*(1), 69–72.

46. Hamidian, A., Krner, U., & Nilsson, A. (2005). Performance of internet access solutions in mobile ad hoc networks. In G. Kotsis & O. Spaniol (Eds.), *Wireless systems and mobility in next generation internet. Lecture notes in computer science* (Vol. 3427, pp. 189–201). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-31963-4_14.

47. Seet, B. C., Liu, G., Lee, B. S., Foh, C. H., Wong, K. J., & Lee, K. K. (2004). A-STAR: A mobile ad hoc routing strategy for metropolis vehicular communications. In N. Mitrou, K. Kontovasilis, G. Rouskas, I. Iliadis, & L. Merakos (Eds.), *Networking 2004. Lecture notes in computer science* (Vol. 3042, pp. 989–999). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-24693-0_81.

48. Das, S. R., Belding-Royer, E. M., & Perkins, C. E. (2003). Ad hoc on-demand distance vector (AODV) routing. *RFC 3561*.

49. Ko, Y., & Vaidya, N. (2000). Location aided routing (LAR) in mobile ad hoc networks. *Wireless Networks*, *6*(4), 307–321. doi:10.1023/A:1019106118419.

50. Tripp Barba, C., Urquiza Aguiar, L., & Aguilar Igartua, M. (2013). Design and evaluation of GBSR-B, an improvement of GPSR for VANETs. *Latin America Transactions, IEEE (Revista IEEE America Latina)*, *11*(4), 1083–1089.

51. Xiao, D., Peng, L., Asogwa, C. O., & Huang, L. (2011). An improved GPSR routing protocol. *International Journal of Advanced Computer Technology*, *3*, 132–139.

52. Lee, K. C., Härri, J., Lee, U., Gerla, M. (2007). Enhanced perimeter routing for geographic forwarding protocols in urban vehicular scenarios. In *Globecom Workshops, 2007 IEEE (IEEE, 2007)* (pp. 1–10).
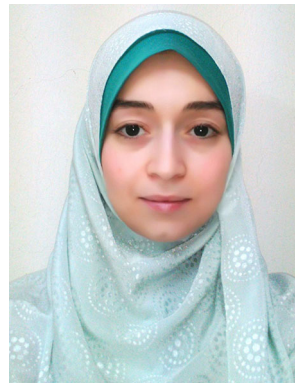
Tasneem Darwish is a Ph.D. student at the Faculty of Computing, University Technology Malaysia, Malaysia. She is a member of the Pervasive Computing research group. She received her B.Sc. degree in computer engineering from Islamic University of Gazah in 2005 and M.Sc. degree in electronics and electrical engineering from the University of Glasgow, United Kingdom, in 2007. From 2008 to 2013 she was working as a research and lecturer assistant at the Faculty of applied engineering, University of Palestine. Her current research focuses on vehicular communications, wireless ad hoc networks, and mobile computing.

Kamalrulnizam Abu Bakar is an Associate Professor in Computer Science at University Technology Malaysia, Malaysia, and member of the Pervasive Computing research group. He received his B.Sc. degree in computer science from the University of Technology Malaysia, Malaysia, in 1996, M.Sc. degree in computer communications and Networks, from Leeds Metropolitan University, United Kingdom, in 1998, and his Ph.D. degree in computer science from Aston University, United Kingdom, in 2004. His research interest includes mobile and wireless computing, ad hoc and sensor networks, information security and grid computing. He is a member of ACM and International Association of Engineering. He involves in many research projects and is a referee for several scientific journals and conferences.