Geographic routing protocols for Vehicular Ad hoc NETworks (VANETs): A survey

Souaad Boussoufa-Lahlah a, Fouzi Semchedine a,b,*, Louiza Bouallouche-Medjkoune a

a LaMOS Research Unit, Faculty of Exact Sciences, University of Bejaia, Algeria
b Institute of Optics and Precision Mechanics (IOMP), University of Setif, Algeria

Abstract

During the last decade, many routing protocols have been proposed for Vehicular Ad hoc NETworks (VANETs) by taking into account their specific characteristics. The protocols based on the vehicles’ positions, named geographic routing protocols (GR) or position-based routing protocols (PBR), were shown to be the most adequate to the VANETs due to their robustness in dealing with the dynamic environment changes and the high mobility of the vehicles. Instead of using the IP addresses, as in the case of Mobile Ad hoc NETworks (MANETs) protocols, position-based routing protocols are based on the geographical position of the vehicles when selecting the best path to forward the data. Further, they do not exchange link state information and do not maintain established routes as in MANET routing protocols. This makes the protocols more robust to the frequent topology changes and the high mobility of the vehicles. In this paper, we present a state-of-art of the routing protocols based on the geographic position of the vehicles. We discuss the pros and the cons of these protocols by exploring the motivations of design of such protocols and we define some possible directions for future research related to the use of this class of protocols.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Vehicular Ad hoc NETworks called VANETs are an application of Mobile Ad hoc NETworks (MANETs) [34]. They form the core of an Intelligent Transport System (ITS) designed to rationalize the operation of vehicles in order to improve road safety. Their main advantage is that they bypass the need of an expensive infrastructure since wireless technology becomes pervasive and cheap. Indeed, due to devices installed inside vehicles or placed at roadside, vehicular communications allow drivers to be warned early enough of potential dangers and situations. In addition to improve the road safety, Vehicular Ad hoc NETworks also offer new services to users such as traffic conditions, weather conditions (e.g., ice), and Internet services, making hence the trip more comfortable.

In a VANET (Fig. 1), vehicles are autonomous and move in a self-organized way along roads and, exchange information with other vehicles and road infrastructures within their radio range. They allow, on the one hand, a direct Vehicle-to-Vehicle communication (V2V) and on the other hand, a Vehicle-to-Infrastructure communication (V2I). V2V communication operates in a decentralized architecture, and it is a particular case of mobile ad hoc networks. It is based on the simple inter-vehicle communication without access to any fixed infrastructure. Indeed, a vehicle can communicate directly with another vehicle if it is within its radio range, or through a multi-hop wireless communication using neighboring nodes as relays. It can be used to provide information on traffic conditions and/or vehicle accidents via wireless communication. V2V communications are very efficient for the transfer of information services related to road safety, but they do not ensure a permanent connectivity between vehicles due to the high mobility of the vehicles. A V2I communication, in which vehicles send and receive data to/from fixed road infrastructures, can provide real-time information on traffic conditions, weather, and basic Internet services by communicating with the backbone networks. V2I communication environment makes better the use of shared resources and leverages the services provided through access points RSUs (Road Side Units) deployed on roadsides. However, this communication mode is inadequate for applications related to road safety because the infrastructure networks are not efficient regarding to the delivery times.

A VANET is formed of vehicles (cars, buses, and so on) that are equipped with positioning systems (i.e., GPS devices), wireless communication devices (such as IEEE 802.11p/WAVE network interfaces), and digital maps. The set IEEE 802.11p and WAVE (Wire-
VANETs routing protocols are classified into five categories: ad hoc, cluster-based, broadcast, geocast and position-based. However, position-based or geographic routing tends to be the predominant one. Research [10][19] have shown that the Position-Based Routing (PBR) performs well with the high mobility of the vehicles. PBR uses the geographic position of vehicles to decide in which direction a data packet should be forwarded. Generally, this decision is based on a geometric heuristic which selects the direct neighbor that is the closest to the destination and which is called greedy forwarding [16]. The great advantage of the greedy forwarding mechanism is that it depends only on the direct neighbors. Thus, there are diverse requirements on the availability of position information. First of all, position-based routing requires that each node must be aware of: i) its own geographic position, ii) the position of its direct neighbors and iii) the position of the final destination. A node obtains its position by using the GPS. The position of its direct neighbors is transmitted through beaconing messages. Each node periodically broadcasts a beacon message containing its current position and probably other information like the direction and the speed. So, in order to make the routing decision, a source node also requires information on the current geographic position of the destination to include it in the packet header. The information about the destinations’ position is provided by a location service [4].

The fact that greedy forwarding approach uses only local information could cause the risk that a packet gets stuck in a local optimum (void) i.e. no neighbor, that is closer to the destination than the current node itself, exists. To escape from a local optimum, a recovery strategy must be applied. The overall objective of a recovery strategy is to transmit the packet to a node which is closer to the destination than the node of the local optimum. Once a such node is found, the greedy forwarding can be applied. One of the most used recovery strategies is the right-hand rule [16] for crossing a graph. This rule states that when a node x reaches a local optimum, the next edge to cross is the node that is in the counterclockwise from the virtual arc formed by x and the destination. Once a node closer to the destination is found, the protocol is switched to the greedy mode. Another used approach is called the carry-and-forward [6]. When a local optimum occurs, the node carries the packet until that an eligible neighbor appears or it reaches itself the destination. Instead of using these recovery strategies, some algorithms recalculate a new path from the node of the local optimum. Several others recovery approaches were proposed in the literature giving rise to new routing protocols (see section 3).

In this paper, we explore and we present existing Position-Based Routing protocols in VANETs. Based on a main analysis of these protocols and according to the literature, Position-Based Routing can be divided into three categories: Non-Delay Tolerant Vehicular Ad hoc NETworks (Non-DTVANETs), Delay Tolerant Vehicular Ad hoc NETworks (DTVANETs), and Hybrid. The Non-DTVANETs protocols do not consider intermittent connectivity and are only practical in densely populated VANETs while the DT- VANETs protocols do not consider disconnectivity and are designed from the perspective that networks are disconnected by default. Hybrid types combine the Non-DTVANETs and the DTVANETs to exploit partial network connectivity.

This paper is organized as follows: Section 2 provides the main challenges to consider for the position-based routing in VANETs. In Section 3, we present and we analyze some existing works in the literature. Finally, our paper is concluded in Section 4 with a comparative table.

2. Challenges in the design of position-based routing protocols for Vehicular Ad hoc NETworks

The key difference between VANETs protocols and any other form of Mobile Ad hoc NETworks is the design requirements. VANETs have some unique characteristics which make them different from MANETs. These characteristics must be taken into account when designing protocols for VANETs, it include:

High dynamic topology: Unlike ad hoc networks and sensor networks, the VANETs are characterized by a highly dynamic network situation due to the movement of vehicles at high speed making the topology of the network in constant changing. In fact, if we suppose two vehicles moving away from each other at a speed of...
25 m/s (90 km/h) with radio coverage about 200 m, the link between these two vehicles will be kept for only 8 seconds. The duration of direct communication between them is very short making hence the establishment of possible communication very difficult.

**Frequent network disconnection:** Since the vehicles are moving, the network density varies resulting in a frequent disconnection of the network. This disconnection will most occur in the sparse network and in the presence of radio obstacles. A high vehicles density allows the network to be connected, and therefore there is always a path between two nodes wishing to communicate. Conversely, in the case of low vehicles’ density, frequent network disconnections occur, resulting in high rates of broken communications, longer delivery times or even impossible delivery. To ensure continuous communication quality, a robust routing protocol must recognize the frequent disconnections and provide an alternative link easily and rapidly.

**Mobility modeling and prediction:** The establishment of communications between vehicles requires the knowledge of the node positions and their movements which are very difficult to predict due to the mobility pattern of each vehicle. Nevertheless, a mobility model and a node prediction, based on the study of predefined roadways model, including: traffic environment, vehicle speed, driver’s behavior and so on, are of paramount importance for efficient network design. Since real vehicular traces are not available, a traffic simulator can be used to generate the movement of vehicles. However, the chosen traffic simulator can influence on the relevance and the viability of the obtained results [2][3]. It was observed [26] that when realistic vehicular traces is used in simulation, the packet delivery ratio is considerably reduced compared to results obtained from non-realistic traces. Works in [19][26] have shown the benefits of using realistic vehicular traces. Thus, the more the mobility model and the prediction is close to the reality, the more the performance evaluation of the routing protocol is valid.

**Propagation model:** In VANETs, due to the presence of buildings, trees, and other vehicles that act as obstacles, the propagation model must not be assumed to be free space. So, the VANET propagation model must well consider the potential interference of wireless communication from other vehicles or widely deployed personal access points.

**Communication environment:** The communication environments taken into account in the MANETs are often limited to open or indoor spaces (like the case of a conference or inside a building) unlike the VANETs where the movement of the vehicles are linked to a road infrastructure either in a highway or within a city area. The mobility model highly varies from highways to city environments. Highway environments are characterized by a long straights and a high-speed moving vehicles, whereas city environments are characterized by a moderate speed of vehicles and contain numerous cross roads and junctions as well as obstacles (buildings, trees, etc.) influencing the vehicle-to-vehicle communication.

**Delay constraints:** In some VANET applications, the network does not require high data rates but has hard delay constraints that must be respected. For example, safety warning applications should have a minimum end-to-end delay, because if a warning message is received at a destination with a high delay, this message could not be helpful to prevent an accident or to avoid a car crash. Thus, in this type of applications, the delivery time of the message will be crucial.

**Quality of Service (QoS):** Is defined as a set of service requirements that needs to be met by the network while transporting a packet from a source to its destination. QoS support over VANETs remains a challenge because of the various factors we discussed earlier. As each application has its own QoS requirements, we need to develop adaptive QoS routing approaches that can quickly and efficiently set up new routes when the current routing paths become no longer available due to the changes in the vehicle velocity, the vehicle position, the network topology or the distance between vehicles.

3. **Geographic routing protocols for VANETs**

Several routing protocols have been proposed in the literature for VANETs to address the recalled challenges. These protocols are categorized on the basis of the application where they are most suitable: Non-DTVANETs, DTVANETs, Hybrid.

3.1. **Non-DTVANETs position-based routing protocols**

The Non-Delay Tolerant Vehicular Ad hoc NETworks (Non-DTVANETs) protocols do not consider intermittent connectivity. They suppose a highly populated network and use the greedy strategy to forward the data packets. However, the greedy forwarding strategy can fail if no neighbor is closer to the destination than the current node itself. In this situation, we say that the packet has reached a local optimal. To deal with such situation, different recovery approaches were proposed in the protocols of this category.

**GPSR** – [16] describes a position-based routing protocol called Greedy Perimeter Stateless Routing which became one of the most cited work. It consists of a standard greedy forwarding mode and a recovery method called perimeter forwarding used in the cases where a local optimum occurs. A simple example of greedy forwarding appears in Fig. 2 (a). Here, X wants to send a packet to D or receives a packet destined for D. X forwards the packet to Y, as the distance between Y and D is less than that between D and any of x’s other neighbors. This greedy forwarding process is repeated until the packet reaches D. When a local optimum occurs,
the perimeter forwarding of GPSR uses the long-known right-hand rule for crossing a graph. As shown in Fig. 2(b), x is closer to D than its neighbors w and y. The dotted arc on D has a radius equal to the distance between x and D. If two paths, \( x \rightarrow y \rightarrow z \rightarrow D \) and \( x \rightarrow w \rightarrow v \rightarrow D \), exist at D, then x will not choose to transmit to w or y using the greedy approach. x is a local optimum in its proximity to D. Hence, the right hand rule tries to bypasses this local optimum by browsing a virtual arc (that connects the node of the local optimum to the destination node) in the opposite direction of a clock hand to search a node that is the closest to the destination D than the node of the local optimum. In this case, the node w will be the candidate and the packet will be transmitted on the path \( x \rightarrow w \rightarrow v \rightarrow D \).

However, for obvious reasons, the right-hand rule requires that all the edges are not crossing (the graph must be planar). Taking the example of the Fig. 3, x originates the packet to u. The application of the right hand rule results in the tour: \( x \rightarrow u \rightarrow z \rightarrow w \rightarrow u \rightarrow x \). So, on graphs with edges that cross (non-planar graphs), right-hand rule may not tour enclosed face boundary. However, by deleting the edge \( (w, z) \), the path will be: \( x \rightarrow u \rightarrow z \rightarrow D \). Since GPSR works on an unobstructed plane, the authors propose an approach to obtain a planar graph without crossing the network. However, this leads to an overhead of the network. So, this method of recovery is much more state-full than stateless. Further, planarization of the neighborhood in an urban environment surrounded frequently by obstacles (partitioned another connected graph) can lead to network disconnections and thus, can force GPSR to run in the recovery mode frequently, which deteriorate its performances as demonstrated in many works (such as: [14][19][36][31]).

**GSR** – To overcome the limits of GPSR in the presence of radio obstacles (buildings, trees, etc.) and voids as it is the case for city scenarios, [19] proposes Geographic Source Routing, GSR combines greedy routing and topology knowledge of the road to ensure a promising route in the presence of radio obstacles and uses back to greedy fashion as local recovery method. In GSR, when a source node wants to send a data packet to a destination, it calculates the shortest path to the destination (using the Dijkstra algorithm) and based on the information of the map-street, the source node selects on this path the sequence of intersections through which the data packet must pass. So, the packet travels greedily along the shortest path intersection by intersection to reach the destination. However, the shortest path may not be sufficiently dense to route the data packet and this will certainly cause a significant delivery delay of the data packets and will lead to a high rate of packets’ loss.

**SAR** – Spatially Aware Packet Routing [36] is based on GSR protocol and the use of spatial awareness to predicate permanent local optimum and to avoid routing failures in advance. In SAR, a spatial model is constructed based on the extracted topology information and is represented as a graph \( G(E;V) \) where \( V \) refers to the significant places and \( E \) denotes the interconnections between places. Instead of forwarding packets to the neighbor which is geographically the closest to the next intersection on the shortest path, in SAR each forwarding node maps firstly the positions of its neighbors into the graph model, and chooses the neighbor with the shortest path to the destination along the GSR, as the next hop. In the case of a local optimum, the packet is suspended and buffered for a fixed time by the node of the local optimum which periodically checks for a possible forwarding.

**Fig. 3.** An example of a non-planar graph.

**Fig. 4.** An example of spatial awareness for geographic forwarding. On the left, the source vehicle S wants to send a data packet to the destination vehicle D. Without taking into account the spatial environment and by using geographic forwarding, S will forward the packet to its neighbor A, which is closer to the destination than B. However, the right side of the figure reveals the vehicles’ distribution which is strictly bounded to the underlying road structure. Since vehicle A is actually located on the left road segment, the packet will be greedily forwarded for potentially many hops (as long as a neighbor, closer to the destination, exists), before a greedy failure is recognized and eventually recovered.

By considering the spatial awareness, the vehicle S can avoid the failure by forwarding packets to the more suitable neighboring vehicle B instead of A.

However, despite of the permanent prediction of routing failures in advance, SAR accumulates the same disadvantage of GSR.

**CBF – Contention-Based Forwarding** [7] is a geographic routing protocol that does not require proactive transmission of beacon messages. The data packets are broadcasted to all the direct neighbors. The latter decide if it should forward the packet based on a distributed timer-based contention process which allows to the most-suitable node forwarding the packet and suppressing other potential forwarders.

A CBF data packet contains the position of the node that has just forwarded the packet (called the last-hop), the ID and the position of the final destination, and a packet ID. A node that receives such packet and is not the final destination, sets a timer to determine when it should forward the packet. The timeout value is calculated based on the progress that the node provides towards the packet’s destination. The packet progress for a given node \( i \) is defined as

\[
p_i = \text{dist}(l, d) - \text{dist}(i, d)
\]

where \( \text{dist} \) is the Euclidean distance, \( l \) and \( d \) are the positions of the last hop and the final destination, respectively. The timer value is calculated as follows:

\[
t = \begin{cases} 
\tau \left( 1 - \left( \frac{p_i}{p_{\max}} \right) \right) & 0 \leq p_i < p_{\max} \\
\infty & \text{otherwise}
\end{cases}
\]

where \( p_{\max} \) is the radio range and \( \tau \) is the maximum forwarding delay. The value of \( t \) determines how each forwarder participates in the contention process. If it is infinite, the packet is discarded. Otherwise, the node forwards the packet after \( t \) seconds.

If there is no greedy next-hop, i.e. the initiator does not hear any node forwarding the packet, Perimeter Routing [16] is used as a recovery strategy. Simulation results show that CBF achieves a delivery rate of almost 100%. However, there are some situations where not all the nodes in the greedy area can hear each other (hidden nodes problems) and this will create packet’s duplications which could induce a network congestion.

**A-STAR – Anchor-based Street and Traffic Aware Routing** [31] is designed specifically for urban vehicular environments. The most important feature of A-STAR is the use of information on city bus roads to identify an anchor path with a high connectivity for packet delivery. Anchor points are simply ordinary nodes whose geographical positions are known a priori. In anchor-based routing, the source node includes into each packet a route vector composed of a list of anchors or fixed geographic points, through which the packets must pass. A-STAR attributes to every corner a weight based on the number of bus lines in which it is deserved. Then,
an anchor path of a smallest weight is calculated by using the Dijkstra algorithm. As it has been shown that greedy transfer does not work well in an urban environment, A-STAR proposes a recovery strategy more efficient where a new anchor path is calculated from the local optimum and, the packet is recovered through such a path. Comparing with the GSR greedy approach and the GPSR perimeter mode, A-STAR uses a new local recovery strategy that is more appropriate for a city environment and that shows an excellent improvement in packet delivery while keeping reasonable end-to-end delay. However, to find a higher connectivity path from a source to a destination, it uses static information based on the number of bus lines by which the path is served. This causes connectivity problems on some portions of streets. Indeed, most of the data traffic will be carried out on such paths and consequently, this could increase the chance of data congestions.

**GPCR** – Authors in [22] proposed Greedy Perimeter Coordinator Routing. They benefit from the fact that the structure of urban streets form a natural planar graph to use the right hand rule when a local optimum occurs without using algorithms for planarizing graphs which are very expensive in terms of overhead. To avoid radios obstacles (such as buildings which block radio signals) when selecting the next hop greedily, a coordinator node (if it is a part of the direct neighbors of the current node), which is a node located at an intersection, is preferred that a non-coordinator node even if the coordinator node is not the closest to the destination. Once a packet reaches a coordinator node, this latter transmits the packet with the same restricted greedy routing strategy to reach the destination. As it is illustrated in Fig. 5, where the node \( u \) wants to forward a data packet to the destination node \( D \), by using the usual greedy forwarding, it forwards the packet to the node \( 1a \). Then, the node \( 1a \) forwards the packet to the node \( 1b \) where a local optimum occurs because the node \( 1b \) has no neighbor closest to the destination \( D \) than itself. By forwarding the packet to the node \( 2a \) (which is a coordinator node) an alternative path to the destination node can be found without getting stuck in a local optimum at the node \( 1b \).

GPCR does not require global or external information (graph planarization and static street map), but depends on coordinator nodes where the mechanisms proposed for their selection fails on banking and sparse roads. Further, the combination of a GPS equipped with a digital map and a detailed location of streets and intersections, are nowadays a broadly available feature in automotive basic equipment which can offer information regarding whether a node is located at an intersection or not. As a result, there is no need for extra beacon messages to indicate that a node acts as a coordinator. This is a great improvement which reduces overheads and increases performances of GPCR as shown in [37].

**GPSR+AGF** – Authors of [26] have observed two problems with GPSR in VANETs. First, due to the mobility of the vehicles, a node’s neighbor table often contains outdated information of neighbors’ position. This problem can be corrected by increasing beacons’ frequency but it will certainly increase congestion and potential collisions. The second problem is that the destination’s location included in the packet header is never updated despite the destination is moving. Thus, to address these two problems, the authors suggest a technique called Advanced Greedy Forwarding (AGF) where information on the speed and the direction of the vehicle are

---

**Fig. 4.** The use of spatial awareness for greedy forwarding [36].

**Fig. 5.** Greedy forwarding vs. Restricted greedy forwarding [22].
added in the beacon messages plus the total traveling time and the
time to process the packet. So, the next neighbor is chosen accord-
ing to the information of the velocity vectors stored in the neigh-
bor table. With the total traveling time, each forwarder node can
better estimate the current location of the destination. They pro-
pose also a broadcasting mechanisms called Preferred Group Broad-
casting (PGB) that aims (1) to reduce control messages’ overhead
by eliminating redundant transmissions and (2) to obtain stable
routes with the ability of auto-correcting. In a broadcasting-based
route discovery process, choosing the closest next hop node will
increase the transmission hops, while choosing the farther next
hop node will result in the instability of the link due to the move-
ment of vehicles or the signal interferences. PGB chooses a node
with a moderate distance as the next hop in order to provide a
stable relay node for routing algorithms. In addition, PGB decides
whether to forward the broadcast message based on the receivers
without the extra control information. Simulation results showed
a very large improvement in packet delivery ratio compared to
GPSR. However, the comparison of the simulation results is made
with AODV and GPSR, two protocols known as inappropriate for
VANETS.

**GyTAR** – The key feature of improved Greedy Traffic Aware Routing
protocol [12] is to take into account the density of the road before
forwarding the data packets. It adopts an intersections-based rout-
ing approach to route the data packets. The next intersections are
chosen dynamically and one-by-one in order to take into account
the real time vehicular traffic variation. To select the next inter-
section destination, a vehicle receiving a packet at an intersection
starts by determining the adjacent intersections. Then, it assigns a
score to each intersection. This score takes into account two pa-
rameters: (i) the curvilinear distance \( D_j \), between the candidate
intersections of the destination node, and (ii) the density of the
traffic \( T_j \) at the road segment relying the current intersection to
the candidate intersection. Thus, the intersection with the highest
score \( S(j) \) is the closest to the destination node and for which the traf-
fic density is high enough to reach it.

\[
S(j) = \alpha \times f(T_j) + \beta \times g(D_j),
\]

where \( \alpha \) and \( \beta \) are weighting factors.

Then, the data packets are routed greedily between each pair of
intersections. If a local optimum occurs, the local recovery so-
lution of GyTAR is based on the carry-and-forward approach where
the packet is carried by the vehicle until it hits a node closer to
the destination or reaches itself the concerned intersection. Results
showed that GyTAR outperforms GSR in terms of packet delivery
ratio, routing overhead, and end-to-end delay. However, the next
chosen intersection may not be the real optimal one due to the
limitation of the traffic information on the direct neighboring seg-
ments, which could increase the transmission delay or even, the loss of the data packets.

**CAR** – Connectivity-Aware Routing protocol [27] is based on Pre-
ferred Group Broadcast (PGB) [26] to minimize broadcasts in the
AODV route discovery mechanism when the source node searches
the destination location and an anchor path, and based on Ad-
vanced Greedy Forwarding (AGF) [26] to account the node mobi-
ity. Once the source node founds the destination location and the
anchor path, the data packets are greedily forwarded to the desti-
nation through the set of anchors. Instead of forwarding a packet
to a node closer to the destination, the packet is forwarded to a
node closer to the next anchor. The path is maintained with the
help of guards. These latter are an entry in the periodic Hello bea-
con of the nodes and are maintained by the neighboring nodes of
the destination whenever the destination’s location changes. To
overcome the local optimum, the authors propose two recovery
methods: time-out with active waiting cycle, and walk-around error

![Fig. 6. Routing with GPCR and the critical junction node (coordinator); circle 1 indicates case 1 and circle 2 indicates case 2 [18].](image)

recovery. In the first method, the node of the local optimum buffers
the packet and periodically checks for a possible forwarder. In the
second one, the node of the local optimum buffers the packet, in-
forms the source node about the error and starts a local discovery
for another route. CAR shows a high rate in packet delivery and a
low average data packet delays than GPSR and GPSR+AGF. How-
ever, a high overhead is generated by the path discovery phase.

**PBR-DV** – Distance-Vector-Based Recovery-Strategy [15] combines
the advantages of the greedy routing with the benefits of the Dis-
tance-Vector-based strategy (AODV: Ad hoc On-demand Dis-
tance Vector). This latter is the recovery strategy of PBR-DV in
the case where a local optimum occurs. When the packet reaches
a local optimum, PBR-DV is switched to the AODV mode where
the node of the local optimum broadcasts a route request (RREQ)
message with the intention to find a node that is closer to the de-

tination than itself. Once a such node is found and which is not a
local optimum, it delivers a route reply (RREP) message. This RREP
is sent to the node of the local optimum on the back path built by
the RREQ message. Thus, the route to forward the data packet is
built on such path. The simulation results show that the amount of
packets delivered with PBR-DV is significantly higher compared to
AODV. However, for the non-greedy part, an excessive flooding is
required which causes a network congestion.

**Gpsrf+** – [18] removes the necessity to go through a junction
node in GPCR while keeping the efficient planarity of topological
maps. It uses two-hops neighbor information to predict in which
road segment its neighboring junction node will forward the data
packet. If the prediction indicates that its neighboring junction
node will forward the packet onto a road with a different direction,
it forwards the packet to the junction node; otherwise, it bypasses
the junction node and forwards the packet to its furthest neigh-

boring node reducing also the number of hops. As shown in Fig. 6,
where the destination node is \( D_1 \) and the routing starts from \( A \),
if a packet faces a critical junction and that, the protocol fails to
provide a valid junction path, GPCR will surely encounter a local
optimum (see case 1 in Fig. 6).

Moreover, when a packet does not face a critical junction (the
destination node is \( D_2 \) in Fig. 6), routing to a junction node is
counter-productive since crossing the junction to the relay bring-
ing the maximum distance would have been preferred (see case
2 in Fig. 6). Thus, it would be better if the observation of a criti-
cal junction is made by nodes before the junction. That is precisely what is proposed by Gpsr+.

In the perimeter mode, Gpsr+ uses the right-hand rule to determine the best direction and thereby, the best forwarding node. So, if the furthest node is in the same direction as the best direction, the best forwarding node is the furthest node; otherwise, the best forwarding node is a junction node. By bypassing junction nodes, Gpsr+ aims to increase the packet delivery ratio of GPCR and to reduce the number of hops in the recovery mode by 200% compared to GPSR. However, the concept of junction nodes is also a foreseen problem that must be addressed in a dynamic urban environment. In fact, due to the mobility of the vehicles (dynamic urban environment), the neighborhood table often contains outdated positions of the junction nodes. This complicates the maintenance of paths towards such junction nodes.

**GRANT** – Greedy Routing with Abstract Neighbor Table [29] uses the concept of extended greedy routing where each node knows its x hop neighborhood by adding the neighbor table to every beacon. This gives to each node a farsighted vision of the best route to avoid the local optimum. The metric of selecting a next forwarding neighbor E is based on multiplying the distance between a node N, the x hop away from E to the destination, the shortest path from N to E, and the charge per hop for multi-hop neighbors. The neighbor E that offers the lowest metric will be chosen to be the next hop. To reduce the network overhead, GRANT separates the plane into areas and includes one representative neighbor per area. Simulation results shown that in city scenario with obstacles, the approach of extended greedy routing works well as the usual greedy approach. The number of times the packet is recovered per route is also less in GRANT than in traditional greedy routing. The major inconvenient of GRANT is that its performance evaluation is done on static traces, and there are no absolute performance metrics such as: packet delivery ratio, network overhead, etc, which could validate its real performances.

**LOUVRE** – [20] proposes a geo-proactive overlay routing solution that creates a landmark overlay network on top of an urban topology, ensures an obstacle-free geographic routing on the overlay links and also, reduces the chances of falling in a local optimum. The features of LOUVRE are as follows: (i) landmarks (overlay nodes) are placed at intersections, (ii) a vehicular traffic density between landmarks is distributively estimated among the vehicles, (iii) an overlay network between landmarks is built, using the Dijkstra algorithm, by considering the traffic density-based overlay links, (iv) the best paths from and to any landmark lying on the same grid are maintained for the local routing, and the packets are then forwarded in a greedy stateless fashion towards the destination. A greedy procedure progressively selects the local optimum path, (v) for remote routing, packets are routed to the best neighboring grid. Evaluation results shown that LOUVRE performs better than GPCR and GPSR in terms of packet delivery ratio, and achieves lower hop count due to LOUVRE’s global knowledge of the density distribution on the road segments and the local optimum. The drawback of this approach is obviously its scalability. Broadcasting and maintaining all the density information of all the ever encountered roads is a very complex task.

**JBR** – Junction-Based Routing [37] assumes that every vehicle is equipped with a GPS device specifying its position. It also considers that every vehicle is equipped with a digital map of the city streets where it moves. The combination of these two assumptions can offer information regarding whether a vehicle is at a junction (and thus is a coordinator vehicle) or is placed in the middle of a road (and thus is a simple vehicle). As a result, there is no need for extra beacon messages, like in GPCR, to indicate that a vehicle acts as a coordinator. This is a great improvement over GPCR. The JBR protocol makes use of selective greedy forwarding up to the vehicle that is located at a junction and which is closer to the destination. The source uses selective greedy forwarding and forwards the packet to the neighbor that is selected as a next hop. The process continues until the packet reaches its destination. If a local optimum is reached, a recovery strategy is applied. This latter is based on the minimum angle method for determining the appropriate next hop while being at a recovery mode, which provides an accurate and a safe solution that can be applicable in all cases regardless of the relative position of the source, the destination and the intermediate vehicles. JBR presents interesting results in terms of the end-to-end delay, the delay distribution and the packet delivery ratio which outperforming the GPCR routing algorithm. However, the better performance is achieved only with the higher transmission ranges.

### 3.2. DTVANETs position-based routing protocols

The Delay Tolerant Vehicular Ad hoc NETworks (DTVANETs) aims to support a class of vehicular network applications characterized by the delay tolerance and the asynchronous data traffic. Such applications can tolerate some data losses. It uses opportunistic strategies to overcome frequent disconnections of the network. This category includes:

**VADD** – Vehicle-Assisted Data Delivery [40] is a vehicular routing protocol that aims to minimize end-to-end delivery delays from a moving vehicle to a static destination in sparse vehicular networks by using the idea of carry-and-forward strategy. VADD is based on the use of a predictable vehicle mobility which is limited by the traffic pattern and the road layout. The vehicles are assumed to be equipped with pre-loaded digital maps which provide the street-level map and the traffic statistics such as the traffic density and the vehicle speed on roads at different times of the day. VADD has three packet modes: *Intersection*, *StraightWay*, and *Destination* where each mode is based on the location of the node carrying the packet.

At the *Intersection mode*, the node carrying the packet can sort all the outgoing directions and checks if there is a contact available to help forwarding the packet through that direction. As shown in Fig. 7, vehicle A has a packet to forward to a certain destination. Assuming that the optimal direction for this packet is North, there are two available contacts for the node A: B moving South and C moving North. Both choices aim to forward the packet toward the North: selecting B because B is geographically closer to the North and provides a better possibility to exploit wireless communication, or selecting C because C is moving to the packet-forwarding direction. These two choices lead to two different forwarding protocols: Location First Probe (L-VADD) and Direction First Probe (D-VADD). Given the preferred forwarding direction of a packet, L-VADD tries to find the closest contact toward that direction as the next hop. In D-VADD, the direction selection process is the same as L-VADD. For a selected direction, instead of probing by loca-
tion (like in L-VADD), D-VADD selects the contacts moving toward the selected direction. Among the selected contacts, the one that is closest to the selected direction is chosen as the next hop.

Data forwarding in the StraightWay Mode is much simpler than in the Intersection Mode since the traffic is at most bidirectional. The intersection ahead which is joined by the current road is simply specified as the target and then, GPSR [16] is applied on the target location. If there is no vehicle available to forward the ahead, the current node continues to carry the packet. A packet switches to the Destination Mode when its distance to the destination is below a predefined threshold. The location of the destination becomes the target location, and GPSR is used to deliver the packet to the final destination. By switching between these packet modes, the node takes the best packet forwarding path.

Simulation results shown that VADD outperforms existing solutions in terms of packet delivery ratio, data packet delay, and traffic overhead. However, it cannot permanently ensure multi-hop connectivity; especially, if unexpected changes in the distribution of road traffic flows occur.

**RRP** – The Reliable Routing Protocol proposed in [17] aims to identify the more reliable paths by predicting the existence of the candidate relay nodes when the Link Expiration Time (LET) expires. RRP assumes that the Road Side Unit (RSU) called anchor node is installed at the intersections of the road segments (blocks) and selects the blocks through which the data packets should be passed in order to reach the destination node. Specifically, the shortest path from the source node to the destination node is selected among the paths on which there are no routing holes. Whenever an anchor node receives a data packet, it uses its digital map and selects the closest block to the destination in order to establish the shortest path and then, forwards the packet to a vehicle on the block. If the anchor node finds that the selected block has a routing hole, when receiving a Block Expire (BE) message from a relay vehicle on the selected block, it selects the closest next block to the destination.

Fig. 8 shows an example scenario where a routing hole occurs on the block closest to the destination (i.e., Block 1). In this scenario, the anchor node A reroutes the data from its area to the Block 2 which is the next closest one to the destination once it receives the BE message from the relay node Rc. Rn and Rc denote the newly selected relay node and the current relay node, respectively. Each vehicle maintains a neighbor table in which the position, the velocity and the movement direction of each neighboring vehicle are recorded. This table is updated via Hello beacons exchanged by all of the neighboring vehicles. Using the information in the neighbor table, Rc determines the vehicle closest to the destination as the next relay node Rn. The current relay node also predicts the existence of candidate relay nodes Nc when the LET for the Rc expires (i.e., LET = 0).

RRP does not attempt to create a new route from the source vehicle, but instead reroutes the data packets to a different block. Simulation results shown that the proposed routing protocol reduces the frequency of route failures and data losses while maintaining a low routing overhead.

**RPS – Reactive Pseudo-suboptimal-path Selection routing protocol** [39] is an anchor-based routing protocol which gives the opportunity to the recently passed intersection to renew a path selection instead of using the carry-and-forward solution when a local optimum occurs in order to improve the probability of transmission through wireless channels. RPS works in three modes: Intersection mode, Segment mode and RPS mode. In the Intersection mode, the packet is delivered along the road segment which is dynamically selected one by one with the highest weight as the optimal path. The weight considers the connectivity of the segment and the distance between the segment midpoint and the destination node. Thus, the packet is transmitted, in the Segment mode, greedily along the selected segment. The RPS mode is activated when a local optimum occurs. Thus, the current node sends a special packet to the nearest intersection where the data packet is already traversed to inform the node at the intersection that a local optimum has occurred. So, it selects another more appropriate path. The simulation results shown that in intermittent connectivity scenarios, the delivery ratio of RPS is about 20% higher and the end-to-end delay is 0.1 s lower than that of CyTAR; although, its overhead is about 10% higher.

**GeoSpray** – [35] is inspired from GeoOpps (geographic forwarding routing protocol) [21]. GeoSpray aims to optimize the resources used in the network, including: the storage, the bandwidth, and the energy, while maximizing the delivery probability and, minimizing the delay and the overhead. It uses geographic position information and other mobility parameters with packet destination addresses in order to make sure that the packets are forwarded toward the destination. However, contrary to GeoOpps that maintains at most one copy of a packet in the network, GeoSpray combines the selected replication and forwards it with the explicit delivery acknowledgment. It employs the concept of spray phase and Wait mechanism [33], where a small/fixed number of packet copies are distributed to distinct nodes in the network. However, instead of using blind replication (as proposed in Spray and Wait), GeoSpray guarantees that packet copies are only spread to the network nodes which will be closer to the packet’s destination.

Furthermore, instead of waiting until one of these network nodes meets the destination and delivers its packet copy (as proposed in the Spray and Wait –wait phase-), GeoSpray allows each node forwarding its packet copy further to another node that can take the data closer to the destination. It controls the flooding by setting an upper bound on the number of copies created per packet, while minimizes the transmission overload and the resource consumption. Network nodes send receipts to inform all the nodes they meet about the packets which have already been delivered. These packets, which are buffered at the intermediate nodes, are removed and the storage capacity for upcoming packets is improved. Simulation results shown that the proposed GeoSpray routing protocol improves the delivery probability and reduces the delivery delay compared to other schemes. However, the simulation environment is not presented (the network simulator, the propagation model, the mobility model, etc.) and thus, this causes difficulties to perform direct comparisons with other standard simulator results.

**OSTD – Opportunistic routing based on Symmetrical Traffic Distribution** [24] is an intersection-based multi-hop routing protocol that is capable of finding optimal routes by taking into account the vehicular traffic condition and the vehicles’ driving path. The proposed algorithm considers the type of the vehicular distribution in the calculation of an utility function. This utility function is used
to evaluate the routes in the network. Vehicles’ driving path predictability is also used in the algorithm to forward the packet to a more suitable next hop since the vehicular mobility is a reflection of the human social activity.

In Fig. 9, the source node S wants to send a message to the destination D near the intersection (Id). If the routing algorithm selects a denser route, Ia–Ib–Id will be used for the message forwarding. This selection’s consequence will be a low data delivery ratio in this scenario, because it does not consider the distribution of vehicles on the road. Ia–Id is higher in vehicular density, but Ia–Ic acts better in message forwarding, because vehicles are more uniformly distributed on this route and the wireless communication speed is much higher than the vehicles’ speed of movement. Therefore, the vehicular distribution is important to be considered in the route selection as well as other factors such as the vehicular density and the Euclidean distance. OSTD considers all of these factors simultaneously and gives a more important to the vehicular density and the distribution in comparison with the Euclidean distance. Simulation results shown that OSTD achieves a higher delivery ratio, and a lower end-to-end delay and packet loss compared to other well-known protocols.

### 3.3. Hybrid position-based routing protocols

Hybrid types combine the Non-DTVANETs and the DTVANETs to exploit the partial network connectivity. When the network is dense, the greedy strategy is used for forwarding the data packets and when a disconnection occurs, the mobility of the vehicle is exploited by carrying the packet until an eligible neighbor appears or it reaches itself the destination. Among these protocols, we cite:

**GeoDTN+NAV** – [5] is an hybrid of Non-DTVANETs and DTVANETs approaches that includes the greedy mode, the perimeter mode, and the Delay Tolerant vehicular Network mode. The packets are first forwarded in greedy mode and then, by the perimeter mode when a packet hits a local optimum. If the perimeter mode also fails, it finally switches to the Delay Tolerant vehicular Network mode and bases on the mobility to deliver the packets. The default greedy forwarding strategy is the same as the restrictive greedy forwarding strategy of GPCR, where packets are always forwarded between junction nodes since junctions are the only places where a node can make significant routing decisions. If a local optimum is reached, the recovery mode, called the perimeter forwarding, is used. It switches from Non-DTVANET mode to DTVANET mode by estimating the connectivity of the network based on the number of hops the packet has traveled in the perimeter mode, the neighbor’s delivery quality, and neighbor’s direction with respect to the destination. The delivery quality of neighbors is obtained through the Virtual Navigation Interface (VNI). The authors assume that every vehicle is equipped with a Virtual Navigation Interface (VNI): (i) They classify vehicles based on the traffic pattern into four broad categories:

1. **Deterministic (Fixed) Route**: Vehicles move strictly along pre-configured routes. These vehicles will not deviate away from their routes. Also, the moving direction of vehicles can be derived from their routes.

2. **Deterministic (Fixed) Destination**: Vehicles move strictly toward a pre-configured destination. However, it is possible that the vehicles take different routes to reach the destination.

3. **Probabilistic (Expected) Route/Destination**: Vehicles may move based on suggested routes or destinations. They are allowed to change their route or destination discretionarily.

4. **Unknown**: Vehicles could not provide information about their route, but they do not move randomly either.

(ii) Retrieving route-info and confidence from vehicles:

1. **Route info**: It represents the vehicle’s route information. Note that the route information may either consist of the detailed path, the destination, or the direction of vehicles, depending on the types of underlying data sources. As in Fig. 10, the VNI might be able to retrieve the detailed path information from a navigation system while it may only retrieve vehicle’s direction from an Event Data Recorder (EDR). In addition, the VNI can also retrieve the pre-configured route information.

2. **Confidence**: Confidence indicates the probability that the vehicle’s movement would respect the given route information. More specifically, a confidence with 0% means that the vehicle moves completely in random manner while a confidence with 100% means that the vehicle moves strictly based on its route information. This confidence information can be configured or derived from the vehicles’ movement history.

GeoDTN+Nav outperforms GPCR and GPSR in terms of packet delivery ratio as it improves the graph teachability by using the delay tolerant store-carry-forward solution to mitigate the impact of the intermittent connectivity. However, in a sparse network, GeoDTN+Nav is likely to fall back to the Delay Tolerant mode frequently. This increases the latency and also, decreases the packet delivery ratio.

**CMGR** – Connectivity-aware Minimum-delay Geographic Routing protocol [30] takes into account the connectivity of the roads, when the network is sparse, to maximize the chance of packets’ reception, and minimize the delay by selecting the non-congested roads which have a sufficient level of connectivity. Gate-Ways (GWs), which represent attachment points to the backbone network, are arbitrarily distributed along the roadsides and, any vehicle that wants to set up a route to any GW, generates a Route Discovery (RD) message and broadcasts it in the network. Between all the RDs received at a GW for the same query but coming from
different routes, the GW selects the most appropriate according to the CMGR logic route selection and generates a route reply (RR) message. The RR is sent back greedily to the target vehicle along the selected route or it will be carried using the carry-and-forward approach if a local optimum occurs.

When the RR is returned back to the target vehicle, and if this latter has moved away from its initial position, a mechanism to track it is proposed where the target vehicle has to broadcast its new velocity vector in its beacon packet before moving. The simulation results shown that the packet delivery ratio of CMGR is approximately 25% better than VADD and A-STAR for the high vehicle densities and goes up to 90% better for the low vehicle densities. However, the tracking mechanism proposed for a moving destination will certainly fail if the target vehicle is in a sparse area because no traces of such vehicle will be found.

**ROAMER** - Roadside Units as message routers [23] exploits the presence of Road Side Units (RSUs) to route the packets to the distant locations in VANETs without necessarily knowing their positions. Further, and in order to preserve their privacy, the vehicles employ a pseudonym as in [11] instead of their actual identity when communicating. Hence, a vehicle $S$ requesting to send a packet $P$ to a distant vehicle $D$ can send the packet to the nearest RSU ($R1$) (using a shortest path algorithm if $R1$ is not within its transmission range), which in turn sends $P$ to the nearest RSU ($R2$) through the backbone network. $R2$ can then send the packet to $D$ via a multi-hop technique. This latter combines position-based routing and carry-and-forwarding strategy to opportunistically route messages to and from the RSU in dense and sparse network conditions. The basic motivation of ROAMER behind using RSUs is that the RSUs are a fixed infrastructure. It is much easier to send a packet to a fixed nearest target than to a remote moving object. Also, the delay of sending the packet through the fixed RSU network will be much less than through the VANET, especially, with the high mobility of VANETs that may delay the packet at the intermediate nodes or even lead to dropping it.

### 4. Conclusion

In this paper, we discussed the challenges of designing routing protocols for VANETs and surveyed several geographic routing protocols dedicated for this kind of networks. Table 1 summarizes the characteristics of these geographic routing protocols by considering: their requirements (GPS, GLS, beacon or not), their forwarding strategy, the recovery strategies, the architecture (V2V or V2I), the type of applications where they are most practical (delay tolerant or not, Hybrid) and how they are evaluated (mobility model, network simulators, propagation model, simulation scenario, ...) etc.

As we can see from Table 1, all the geographic routing protocols require the use of a positioning system (e.g. GPS) since it is their principal distinction from other Ad hoc networks where the geographical position is used in order to make the forwarding decision. So each node has to be aware of its position by using a GPS. The GPS allows to determine the actual geographic position of a node, its direction, its speed. The GPSs are also equipped with digital maps with the detailed locations of streets and junctions.
where the vehicles move. The position of its neighbors is gotten from periodical beaconing messages. The beaconing interval is assumed, by the majority of the geographic routing protocols, to be fixed. In [26], it is observed that the inconsistency of neighborhood information tables leads to significant problems and a low flow. The outdated information in the neighborhood tables can be cured with more frequent update, as suggested by the authors of GPSR. However, this will definitely increase the congestion and the collision possibilities. To address this problem, GPSR+AGF protocol suggests a technique called Advanced Greedy Forwarding (AGF) where the information on the speed and the direction of the vehicle are added in the beacon messages. So, the next neighbor is chosen according to the information of the velocity vectors stored in the neighbor table. The position of the final destination can be obtained through a localization system (as GLS) or a geocast application.

Each geographic routing protocol uses the same forwarding strategy which is the greedy approach. The principle of the greedy forwarding is to forward the packet to a neighbor that is the closest to the destination. Some protocols restrict this forwarding to the intersection as is the case for GPCR and GyTAR. Others use it along the path to the destination (A-STAR, GSR), or both as for the case of the RPS routing protocol. If a local optimum occurs, a recovery strategy is applied to recover the data packets. Different recovery strategies are proposed, the well-know are the right-hand rule and the carry and forward strategies.

The presence of a map is a valid assumption since nowadays vehicles can be equipped with onboard navigation systems. For example, in A-STAR and GSR protocols, the road-map knowledge is required for their routing strategy in order to pre-compute the path to the destination.

In this paper, we have focused on the routing problem; especially, on the geographic routing protocols class because it has been demonstrated in many studies that it is the most robust class for Vehicular Ad hoc NETworks that can handle with the frequent topology changes and the high mobility of the vehicles.

One of the main challenges in Vehicular Ad hoc NETworks routing protocols is not only searching an effective path for transporting data packets but also, maintaining this path and recovering it in the case of the local optimum. However, the performance and the efficiency of a routing protocol in VANETs depends highly on many other factors:

- We have noticed that the most existing routing protocols are designed to work in urban environments while driving can also be in highways. The difference between highway and city environments is that the highway scenarios are largely one dimensional which is very advantageous for greedy position-based routing approaches because it can very well deal with the high mobility of the nodes. However, the city environment, that is two dimensional scenario, is characterized by the presence of radio obstacles (buildings, trees, etc.) which often block radio signals and have a significant negative impact on the performance of position-based routing mechanisms. So, the routing protocol must be able to deal with the challenges of the city environments and to find robust routes. So, an efficient position-based routing protocol must consider both type of environments (city and highway) and, must be able to switch in a transparent manner.

- The network density characterizes the type of the application where the routing protocol is most suitable: delay-tolerant, non-delay tolerant or both. There are few protocols which work in the both applications. When the network is sparse, the main challenge is to maximize the chances of packet receptions before they expire; while in dense networks, the aim is to minimize the end-to-end delay. Thus, an efficient routing protocol must be able to adapt and to maintain routes when the network connectivity changes as it was the case for GeoDTN+NAV, CMGR and ROAMER routing protocols.

- The performance of a VANET routing protocol also depends on several other parameters, including the mobility model. The validity of the simulation results depends strongly on the ability of the used models to reproduce as accurately as possible the actual situations. A mobility model reflects the spatio-temporal behavior of the mobile nodes, where the aim is to represent the conditions of traveling in a particular context of the real world. However, the mobility models designed for mobile ad hoc networks, such as the Random Way Point (RWP) model [13], cannot be directly used in the vehicular networks where the vehicle displacements and speeds are delimited and predefined by the roads and by the behavior of the drivers. So, it is very important to use realistic mobility models which reflect the reality. This can be very important in analyzing the performance of the different proposed solutions.

Thus, surveying the existing VANETs geographic routing protocols and comparing their features is absolutely an essential step for designing an efficient new proposition that can deal with the specific characteristics of VANETs. Thereby, future researches must focus on the design of protocols which take into account the cited limits and, not only providing reliable packet forwarding with a minimum delay, a maximum throughput and a low communication overhead but also they should interact with other layers to have a robust cross-layer routing protocols since many applications impose stringent QoS requirements which may not be met by single layer network design solutions. However, any Cross-Layer design should take attention to undesirable effects which can occur due to the cross-Layer exchanges and affect the system performance.

References


