Routing protocols for inter-vehicular networks: A comparative study in high-mobility and large obstacles environments

Yao H. Ho *, Ai H. Ho, Kien A. Hua

School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816-2362, United States

Abstract

An ad hoc network is composed of mobile nodes without the presence of a fixed infrastructure. Communications among nodes are accomplished by forwarding data packets for each other, on a hop-by-hop basis along the current connection to the destination node. In particular, vehicle-to-vehicle communications have been studied, in recent years, to improve driver safety. As more of such applications of high-mobility ad hoc networks emerge, it is critical that the routing protocol employed is capable of efficiently coping with the high-frequency of broken links (i.e., robust with respect to high-mobility). This paper presents a comprehensive comparative study in a city environment of eight representative routing protocols for wireless mobile ad hoc networks and inter-vehicular networks developed in recent years. In a city environment, communication protocols need adapt fast moving nodes (e.g., vehicles on streets) and large obstacles (e.g., office buildings). In this paper, we elaborate upon extensive simulation results based on various network scenarios, and discuss the strengths and weaknesses of these techniques with regard to their support for highly mobile nodes.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Inter-vehicular network; Mobile ad hoc network; Performance comparison; Routing protocol; Mobility

1. Introduction

Wireless networks have become increasingly popular in recent years. There are two variations of mobile wireless networks: infrastructure mobile networks and infrastructureless mobile networks. The latter, also known as an ad hoc network, has no fixed routers. Instead, the mobile nodes themselves function as routers which discover and maintain communication connections. Thus, a mobile ad hoc network (MANET) is a self-organizing multi-hop wireless network where all nodes participate in the routing and data forwarding process. Such a network can be easily deployed in situations where no base station is available, and a network must be built spontaneous. As examples, for applications such as vehicle-to-vehicle communications, battlefield communications, national crises, and disaster recovery, a wired network is not available and ad hoc networks provide the only feasible means of communications and information access.

There are many routing protocols designed to relay data in mobile ad hoc networks (MANETs) such as Ad hoc On-demand Distance Vector (AODV) [30], Asymptotically Optimal Geometric Routing [21], Cluster Gateway Switch Routing (CGSR) [5], Connectionless Approach (CLA) [14], Contention-based Forwarding (CBF) [10], Distance Routing Effect Algorithm for Mobility (DREAM) [1], Dynamic Destination Sequenced Distance-Vector Routing (DSDV) [29], Dynamic Source Routing (DSR) [17], Greedy Perimeter Stateless Routing (GPSR) [19], Location-aware Routing Protocol (GRID) [24], Global State Routing (GSR) [4], Location-Aided Routing (LAR) [20], On-demand Multicast Routing [22], Trajectory-based Forwarding (TBF) [28], Trigger-based Distributed QoS Routing [8], Location-based Routing (TMNR) [2], Wireless Routing Protocol (WRP) [26], and Zone Routing Protocol (ZRP) [11]. However, most of them are not designed for street environments.

Earlier protocols such as Dynamic Source Routing (DSR) [17], Ad hoc On-demand Distance-Vector Routing (AODV) [30] are designed for ad hoc networks. These protocols are well suited for MANETs since they are very efficient in terms of bandwidth usage and delay, and they have good scalability. However, they are not suitable for applications where the networks need to adapt to the movement of nodes. In such cases, routing protocols that take into account the movement of nodes are necessary. These protocols are designed to cope with the high-frequency of broken links in high-mobility environments. They are capable of efficiently adapting to the movement of nodes and maintaining communication connections, even when the network topology changes rapidly.
(AODV) [30], and Location Aided Routing (LAR) [20] require a source to use route request to establish a hop-by-hop route between itself and a destination before sending data. In the street environments, however, obstacles and fast moving nodes result in a very short window of communication between nodes on different streets. The established route expires quickly and the source needs to re-issue another expensive network wide route request after sending only a few data packets via the previous route. These protocols, when applied in the street environments, will incur a high control overhead in terms of route request packets.

To overcome the fragility of multiple-hop routes, one-hop-based approaches, such as Trajectory-based Forwarding (TFB) [28], let each forwarder select the next forwarding node by comparing the positions of all its neighbors with the trajectory defined by a source. This position information is obtained through periodic broadcasts from neighboring nodes. The short window of communication in the street environments, however, means that the nodes need to broadcast more frequently in order to maintain up-to-date location information. This strategy incurs a high control overhead in terms of frequent beaconing packets that also congest the wireless medium. We identify the above protocols as a Connection-oriented approach because each link a packet traverses must first be established through a network wide route request or location information exchange among all the nodes in the network.

Rather than using the expensive control overhead to pre-establish each link, the Connectionless approach allows a node to dynamically participate in a forwarding of data by comparing its current location with headers of the data, which contain location information of a source, a destination, and a previous relayer. Existing Connectionless techniques such as Contention-based Forwarding (CBF) [10], Beacon-Less Routing (BLR) [12], and Connectionless Approach to Mobile Ad Hoc Network (CLA) [14] only allow nodes that have the shortest distance to the destination or are on the shortest geographic path (i.e., a straight line) between the source and destination to relay data. When applying these Connectionless approaches to the street environments, nodes that can relay data around obstacles often do not get to relay the data because they are farther from the destination than the previous relayer or are not on the forwarding path of the data. Thus, these approaches cannot be applied directly to the street environments.

Recently, a new method [25] has modified the CBF technique to address the obstacle problems in a city environment. This scheme allows a source to specify a forwarding path as a list of junctions, and applies CBF between consecutive junctions. This solution requires at least one node at each turning junction, which is often difficult to achieve over an extended communication period. To overcome this drawback, another method (i.e., CLA-S) [13] adapt the CLA to a street environment to utilize multiple communication paths.

A number of papers have been published on performance comparison of these ad hoc routing protocols using different simulation models. However, none of them are studied under street environments. In [3,7], and [18], the ns-2 simulator [16] is used to study communications among 50–100 mobile nodes. Their moving speeds are limited to 20 m/s, with relatively long pause time ranging from 0 to 900 s. It is not clear if the simulation results are still valid under a high-mobility environment such as vehicle-to-vehicle communications and advanced battlefield applications. More recent studies [15,31,34], and [36] focus on scalability by varying the network size and the number of nodes in the simulations. In [31], up to 400 nodes are studied, with node speed and pause time fixed at 14 m/s and 0 s, respectively. Up to 50 mobile nodes are considered in [15]. Their pause time is fixed at 30 s while node speed varied up to 30 m/s. The mobility assumed in these two studies, however, is too slow to reveal the impact of high-mobility on performance. Large networks with up to 50,000 nodes are studied in [36]. In this work, the two highest-mobility scenarios considered are 14 m/s (32 miles/h) with zero pause time, and 30 m/s (67 miles/h) with a 30-s pause time. Since either a relatively low node speed or a long pause time is assumed, the effective node mobility is still too low to capture the characteristics of high-mobility environments such as vehicle-to-vehicle communications (i.e., high node speed with zero pause time). Furthermore, the routing protocols compared in [36] are relatively old (i.e., the most recent protocol was published in 2001).

Many more advanced routing techniques have been developed for MANETs and Vehicular Networks in recent years. In this paper, we present detailed simulation results under a street environment to compare eight representative routing protocols, namely AODV, DSR, LAR, GRID, TMNR, GPSR, CBF-Street Version, and CLA-Street Version. While DSR and LAR were proposed in 1998, other five protocols were proposed within the last few years to address advancements in mobile applications. In particular, our focus is on high-mobility (e.g., vehicles) and street environments, as many important emerging applications of this technology involve high-mobility nodes. As an example, cars in a vehicle-to-vehicle network are typically moving at speeds exceeding 30 miles/h in a city environment with large obstacles (e.g., office buildings). Very little is known about how existing routing methods perform relative to each other in such street environments. Our purpose is to investigate the impact of high-mobility and large obstacles on different routing protocols under various scenarios. The simulation results provide insight into ad hoc routing protocols and offer guidelines for inter-vehicle network applications.

The remainder of this paper is organized as follows. To make the paper self-contained, we briefly review the eight routing protocols, selected for our comparative study, in Section 2. The simulation environment and the performance model are discussed in details in Section 3. In Section 4, we present the simulation results and examine the
shortcomings of each routing technique. Finally, we conclude this paper in Section 5.

2. Routing protocols review

Early generation routing protocols, such as DSDV, WEP, and GSR, establish communication link by maintaining routing information in a routing table at each node. A drawback to this solution is that every node needs to update its routing table and propagate the update, as the network topology changes, in order to maintain a consistent view of the network. This operation incurs excessive network traffic and computation overhead. Later techniques, such as DSR and AODV, attempt to reduce unnecessary network traffic by initiating route request on-demand. This type of routing protocols establishes communication links by flooding the network to find a route to the destination node. This strategy is simple and robust; however, it is not energy efficient and can cause severe media congestion.

Another approach to reduce network flooding is to leverage location information obtained from GPS (Global Positioning System) [9] or other location services [23] and [32]. For instance, LAR uses location information to limit the area of flooding, thereby reducing the number of route request messages. These schemes result in better power conservation and improve network scalability.

Some other techniques reduce not only number of route requests but also route maintenance costs. This type of approach, such as GRID and CGSR, organizes mobile nodes into clusters. Each cluster has a cluster-head and a number of gateways. Two clusters communicate via a gateway node within their communication range. An obvious advantage of this environment is that only cluster-heads and gateway nodes need to rebroadcast messages. However, the network throughput can be limited by the number of gateway nodes. Furthermore, cluster management incurs overhead.

To address mobility issues, one-hop approaches, such as TBF, TMNR, and GPSR, have been proposed. In these schemes, instead of the need to establish a complete connection from the source to the destination, the node only needs to establish the connection to the next hop (i.e., one hop) and forward the data. To determine the next hop, a node compares the distances of its neighbor nodes to the destination node (i.e., GPSR), the next waypoint (i.e., TMNR), or a trajectory (i.e., TBF).

CBF and CLA are more recent techniques developed for routing in MANETs. While GPSR, TMNR, and TBF need to maintain (proactively or reactively) neighbor nodes location information and establish a connection to the next hop before forwarding a data packet, CBF and CLA simply forward data packets without first establishing the link to the next node. Any node that happens to be in the general direction towards the destination node can compete for the “right” to forward data packets. Since there is no need to pre-establish multi-hops or one-hop connections, we classify these protocols, i.e., CBF and CLA, as a connectionless-based approach. The advantages of this approach are twofold. First, there is no need to maintain communication connections; and second, nodes do not need to maintain information about their neighboring nodes. Recently, CBF and CLA have been extended to address the obstacle problems in a city environment. In the modified CBF technique, a source specifies a forwarding path as a list of junctions and applies CBF between consecutive junctions. In the modified CLA (i.e., CLA-S), a source specifies a forwarding zone as an area of streets and intersections; and applies CLA approach.

To make the paper self-contained, we explain the eight techniques selected for our study in greater detail in the remainder of this section. In summary, they are grouped into four categories, as illustrated in Fig. 1. The shaded boxes represent the eight routing methods selected for our comparative study. These techniques are selected as representatives of their category. We decided not to include proactive techniques in this study because it would not be feasible to maintain up-to-date routing tables in a high-mobility environment where the network topology changes rapidly.

2.1. Dynamic source routing

DSR [17] is a simple and efficient routing protocol. It allows nodes to dynamically discover a multi-hop route across the ad hoc network to any destination node. To do so, a source node S first sends out a ROUTE REQUEST message with the request ID to all of its neighbors. If the nodes receiving this message have not seen it before, they add themselves to the route and forward the message to their neighbors. If a receiving node D is the destination or has information on a route to the destination, D sends a REPLY message containing the full route. D may send the REPLY along the route. After receiving one or several routes, the source node picks the best route, stores it, and starts to send messages along this route. If the network topology changes due to a broken communication connection along the route, source node S can attempt to use any other route S happens to know about, or can issue
2.2. Ah hoc On-demand Distance Vector

AODV [30] has the same on-demand routing characteristics of DSR. Similar to DSR, when a node \( S \) needs to send data to a node \( D \), node \( S \) first needs to broadcast a ROUTE REQUEST message including the last known sequence number for that destination node \( D \). This ROUTE REQUEST message will flood the network until it reaches the node \( D \) or a node with information on a route to node \( D \). These nodes then replies with a ROUTE REPLY message containing the number of hops needed to reach \( D \) and a sequence number for \( D \). Similar to traditional routing tables, each node participating in the route creates an entry in the routing table, where each route entry is maintained using timer-based states. To maintain the state, each node periodically broadcasts a HELLO message. If a node does not receive the HELLO message from a neighbor node for three consecutive periods, then the link is considered broken. When a link breaks, any upstream node recently using this broken link will be notified using an UNSOLICITED ROUTE REPLY message with an infinite metric for the destination. When node \( S \) receives this UNSOLICITED ROUTE REPLY, it will start another route discovery as described previously.

We note that it is unfair to compare AODV and DSR in this study because they do not require GPS. Nevertheless, including these two well-known routing techniques in our study provides a reference to assess the benefit of using GPS.

2.3. Location-aided routing

The LAR [20] protocol leverages location information to limit the area of route request when searching for a new route. Using a previously known location and average speed of the destination node, this scheme estimates the “expected zone” of the destination node as a circular region that may contain the destination node. The request zone is then defined as the smallest bounding rectangle that includes the current location of the source node and the expected zone. Once the request zone has been defined, only nodes in the request zone need to forward the route request message ROUTE REQUEST.

A variant of the LAR scheme is LAR2 [20]. In LAR2, the distance \((\text{DIST}_s)\) between the source node and the destination node, and the location information of the destination node are included in the route request message. When any node receives a routing request, it compares its distance \((\text{DIST}_n)\) to the destination node and the distance \((\text{DIST}_s)\) in the route request message. If \(T_s + \delta \geq \text{DIST}_s\), this node will forward the route request and change the \(\text{DIST}_s\) to \(\text{DIST}_n\) in the request message; otherwise, it will discard the route request. The parameter \(\delta\) determines the tolerant distance for a node to forward or not. Since this scheme still needs to maintain hop-by-hop routes, it does not ease the route maintaining problem.

2.4. GRID-based routing

GRID [24] is a cluster-based approach. This technique reduces the cost of route maintenance by dividing the network area into fixed-size grid cells with nodes within each cell forming a cluster. For each cluster, a node is selected as the gateway, and only gateway nodes may rebroadcast messages. When gateway nodes leave their current cell, they call a handover procedure to select a new gateway. This strategy requires each gateway node to periodically broadcast its existence to other nodes within the cell.

In this environment, nodes establish connections by selecting a list of consecutive grid cells as a route. A gateway node in each of the selected grid cell participates in forwarding the data packets. Thus, the number of hops in a route is determined by the number of selected grid cells. Although the number of hops can be reduced by using a larger grid cell, this can result in a higher frequency of broken links because two gateway nodes of two adjacent cells might move out of the radio range of each other. In the worst scenario, the re-establishment of the connection might not be possible if the same situation is observed with all of the eight neighboring cells. To avoid this problem, we set the grid size to “\(x \times x\)”, where \(x = (\text{nominal radio range} R)/(2\sqrt{2})\), in our simulation study. This ensures that any node in one cell can communicate with any node in the neighboring cells within the same street or within in two intersected streets.

2.5. Greedy perimeter stateless routing

To address the scalability issue, Greedy Perimeter Stateless Routing (GPSR) [19] makes a greedy decision at each forwarding node to minimize the total number of necessary hops. This is achieved by having each node broadcast its location information to its neighboring nodes. To send a data packet, the source node includes the location information of the destination node with the data packet. This information enables a forwarding node to greedily select a neighbor that is geographically closest to the packet’s destination as the next hop. This process is repeated for each hop until the packet reaches its destination.

2.6. Terminode routing

Terminode Routing (TMNR) [2] is a hybrid of location-based routing and the table-driven technique. Each node maintains a routing table with the location information of its one- and two-hop neighbors. To establish a connection, a source node first attempts to determine the route...
to the destination using only information in its local routing table. If this fails, a Direct Path technique is used. This scheme determines the direct path as an approximation of the straight line from the source to the destination. Each data packet is forwarded by nodes alone this direct path until the packet reaches its destination. A perimeter mode is used to circumvent any topology hole by planar graph traversal, similar to GPSR [19]. In the initial stage of this mode, the next forwarding node is generally farther from the destination than the last forwarding node. This mode is switched back to the Direct Path technique when a forwarding node is found to become closer to the destination. Since we do not consider obstacles in this study, we did not include the perimeter mode in our simulation.

We note that predetermined geographical anchor locations are also used in [2] as waypoints towards the destination node. In this environment, a direct path is established between two adjacent waypoints such that a data packet is forwarded by intermediate nodes in the direction of the next waypoint in the list until the packet reaches a node close to that waypoint, at which point the next waypoint is attempted in the same manner until the data packet eventually reaches the destination node.

2.7. Contention-base forwarding for MANET and vehicular network

In Contention-based Forwarding (CBF) [10], a forwarding node transmits a data packet as a single-hop broadcast to all its neighbors. These neighbors compete with each other for the “right” to forward the packet. During this contention period, a node determines how well it is suited to be the next hop for the packet. The node that wins the contention suppresses the other nodes, thus establishing itself as the next forwarding node. This contention is based on the distances of the nodes to the destination. A drawback of this strategy in a high-density environment is that several neighboring nodes might have similar distances to the destination. Consequently, they all establish themselves as the next hop and forward the data packet. This incurs unnecessary network traffic and wastes power of the mobile nodes. A solution, suggested in [10], is for each contestant node to report its qualification for forwarding the data packet and wait for the current forwarding node to select the winner for the next hop. We did not consider this strategy in our study since it is similar to GPSR. In the extended CBF [25] for a city environment, a source needs to specify a forwarding path as a list of junctions. Between two junctions, CBF approach is applied.

2.8. Connectionless approach for MANET and vehicular network

In the Connectionless Approach (CLA) [14], the network area is divided into small “virtual cells”. Instead of maintaining a hop-by-hop route between the source and destination nodes, the source selects a list of grid cells that form a “connecting” path between the source and destination. Nodes within each of these cells alternate in forwarding data toward the next cell using a delay function. This function computes a shorter delay for a node farther from the sender and closer to the destination. In this environment, a connecting path is considered broken if one of its cells becomes empty. This is addressed by replacing the empty cell with a neighboring cell. The fundamental advantages of CLA are twofold. First, a connection path (i.e., a list of grid cells) is much less likely to become broken than a standard route used in conventional techniques; and second, unlike standard routes, the robustness of connection paths is not sensitive to the mobility inherent in the network. In the modified CLA for a street environment [13] (i.e., CLA-S), streets are divided into small “virtual cells.” These cells are divided according to intersections and blocks. A source selects a list of cells as forwarding zones that form a “connecting” forwarding area between the source and destination. Nodes within each of these cells alternate in forwarding data toward the destination node.

3. Simulation environment and methodology

The overall goal of our experiments was to measure the ability of the routing protocols in handling network topology changes in a street environment with high-mobility of nodes. In a street environment, the radio range and mobility are constrained by buildings and streets. If a mobile node is within a block, we can see that a mobile node’s “effective” radio range is only within the street it is on (see Fig. 2). Similarly, if a mobile node is at an intersection, the mobile node’s “effective” radio range will be along the two intersecting streets (see Fig. 3).

The simulators used for our study were implemented using the GloMoSim [35] library, a packet-level simulator specifically designed for ad hoc networks. It follows the OSI 5-layer network communication model. The field configuration is a 1000 m x 1000 m field with a street width of 10 m and a building block size of 100 m x 100 m, unless it is
specified otherwise by the network scenarios. The radio
propagation range for each node is 375 meters and channel
capacity is 2 Mbits/s. Initially, nodes are placed uniformly
with two nodes per intersection and eight nodes per block.
Then the nodes move in the directions permitted in the
streets. Upon arriving at an intersection, a node pauses
for a period (i.e., specify in the different simulation setups)
of time, and then the node probabilistically changes its
direction of movement (e.g., turn left, turn right, or con-
tinue in the same direction). Traffic applications are con-
stant bit rate sessions involving 1/10 of all the nodes.
Each data packet is 512 bytes and the senders are chosen
randomly among the nodes. Multiple simulation runs
(100 runs per setup on average) with different seed numbers
were conducted for each scenario and collected data were
averaged over those runs.

The routing protocols are compared according to the
following four metrics.

(i) Fraction of packets delivered. Measures the ratio of
the data packets delivered to the destinations and
the data packets generated by the CBR source. This
number indicates the effectiveness of a protocol.
(ii) End-to-end delay. Measured in milliseconds, includes
processing, route discover latency, queuing delays,
retransmission delay at the MAC, and propagation
and transmission times. This number measures the
total delay time from a sender to a destination.
(iii) Normalized routing load. Measures the number of rout-
ing packets transmitted per distinct data packet deliv-
ered to a destination. The routing overhead is an
important metric for comparing these protocols as it
measures the scalability of a protocol, and its efficiency
in terms of throughput and power consumption.
(iv) Packet duplication. Measures the average number of
duplicate packets per distinct data packet received
by the destinations. A protocol with a high number
of duplicate packets can congest the network and
waste power of mobile nodes.

The first three metrics were suggested by the IETF
MANET working group for routing protocol evaluation
[6], and were also used in [3] and [7].

4. Simulation results

In our simulation study, we performed sensitivity analy-
sis to investigate the effect of various network parameters
on the eight routing protocols – AODV, DSR, LAR,
GRID, TMNR, GPSR, CBF (Street Version), and CLA
(Street Version). We present our simulation results in this
section.

4.1. Effect of mobile speed

This study is based on 200 nodes with 20 communica-
tion sessions. We set up our simulation with zero pause
time to stress the mobility in the network. To understand
the effect of mobile speed on performance, we varied the
speed of the mobile nodes between 10 m/s (or 22 miles/h)
and 25 m/s (or 56 miles/h).

The simulation results are presented in Figs. 4–7. They
show performance trade-off in some techniques. Although
DSR performs comparably to CLA (street version) and
CBF (street version) in terms of end-to-end delay (Fig. 5)
and number of control packets transmitted per data packet
(Fig. 6), DSR does poorly in delivering data to their desti-
nation (Fig. 4). This can be attributed to the fact that DSR
needs to rediscover routes more frequently as node mobil-
ity increases. Similarly, GRID, TMNR, GPSR, LAR, and
AODV have high end-to-end delay and control packet
overhead because links break frequently due to high node
mobility (Figs. 5 and 6). Under this condition, they need
to send more ROUTE DISCOVERY messages. In addi-
tion, LAR suffers from inaccurate prediction of the request
zone (used to limit the flooding area), which makes flood-
ing the entire network more common.

In the cases of TMNR and GPSR, the high control
overhead is caused by maintaining neighbor information

![Fig. 3. Effective radio range of a mobile node at an intersection.](image-url)
UNCORRECTED PROOF

534 (locations), and high end-to-end delay is caused by the
535 inaccurate (outdated) neighbor information. The inac-
536 curate neighbor information causes TMNR and GPSR to
537 forward to non-existing neighboring nodes. In Fig. 4,
as
538 the mobility increases, the performance of TMNR and
539 GPSR degrades rapidly due to outdated information. Simi-
540 larly, GRID also has high control overhead caused by
541 maintaining information on the gateway node for each grid
542 (see Fig. 6). In Fig. 5, the high end-to-end delay in GRID is
543 not only caused by the inaccurate gateway information but
544 also the fact that only a few selected gateway nodes can for-
545 ward data. The limited number of forwarding nodes (gate-
546 way nodes) causes the network throughput to decrease in
547 Fig. 4. This demonstrates that the Connection-oriented
548 approaches such as Multi-hop based (e.g., AODV, DSR,
549 and LAR), Cluster-based (e.g., GRID), and One-hop based (e.g., TMNR and GPSR) are not suitable to be used
550 in street environments. For example, it only takes 1/2 of a
551 second to traverse an intersection of 10 meters wide at a
552 speed of 20 m/s. This means that the Connection-oriented
553 approaches either drop a large amount of data packets
554 (in the case of DSR in Fig. 4) or require a large amount
555 of control overhead to keep routes from the sources to
556 the destinations up to date (in the case of AODV and
557 LAR in Fig. 6), neighboring nodes information (in the case
558 of TMNR and GPSR in Fig. 6), and cluster membership
559 (in the case of GRID in Fig. 6).

560 In contrast, since Connection less-oriented approaches
561 such as CLA and CBF have no connection to break and
562 maintain or neighbor information to update, they have
563 low control overhead, short end-to-end delay, and high
564 successful delivery ratio. Between these two techniques,
565 Fig. 7 shows the number of packet duplication for CBF
566 is three times higher compared to that of CLA. This is
567 due to the fact that CLA only allows nodes in selected grid
568 path to forward data packets. In addition, CBF has the
569 “fan-out” effect that is similar to the broadcast storm prob-
570 lem [27] and [33] when forwarding data packets. We note
571 that we did not study the effect of mobility beyond 25 m/
572 s (or 56 miles/h) because the performance comparisons
573 can be extrapolated from the trends in the performance
574 behavior.

4.2. Effect of pause time

575 In this study, we fixed the number of nodes at 200, their
576 speed at 20 m/s, the number of communication sessions at
577 20, and varied the pause time between 0 and 600 s to inves-
578 tigate its effect on performance.

579 The simulation results are plotted in Figs. 8–11. We note
580 that as the pause time becomes very long, communication
581 connections are less likely to break and most protocols dis-
582 play about the same performance. Nevertheless, we observe
583 trade-off in the performance metrics among different proto-
584 cols as the pause time is shorter, i.e., higher mobility. We
585 discuss these conditions as follows.

586 In Fig. 8, AODV has very high fractions of packet deliv-
587 ered under short pause time. This is due to the fact that
588 AODV periodically maintains a local routing table in each
589 node, and a data package can be dynamically rerouted to a
590 new next hop if the current “next hop” has moved away.
591 This helps to reduce the number of lost packets. This strat-
EGY, however, incurs a high number of control packets per data packet due to the maintenance of the local routing tables, as seen in Fig. 10.

Fig. 9 indicates that DSR performs well in terms of end-to-end delay. This is, however, due to the fact that DSR takes relatively longer time to establish a route. Longer routes take too long to connect and many of them become broken soon after they are established under high-mobility (i.e., short pause time). Consequently, we observed mostly short routes, two to three hops, in our simulation study with small end-to-end delay. This also explains the low fractions of packets delivered in DSR because many packets delivered over long routes are lost (see Fig. 8). DSR also has low control overhead according to Fig. 10. Nevertheless, this is due to the high percentage of packet loss (see Fig. 8) and we do not take into account the control packets for these lost packets. For LAR, frequently flooding the entire network is caused by inaccurate predication of the request zone due to short pause time. Flooding the entire network will cause high end-to-end delay (see Fig. 9) and high number of packet duplication (see Fig. 11). As the pause time increase, LAR performs better due to the more accurate predication of the request zone.

For GRID, the periodic update of a gateway node is required to notify its existence to the other nodes in its grid even if it stays at the same location. Therefore, the number of control packets stays the same after pause time reaches 100 s in Fig. 10. Since only gateway nodes can forward data packet, the performance of end-to-end delay also stays the same after pause time reaches 100 s in Fig. 9. Similar to GRID, TMNR and GPSR have basic maintenance cost (i.e., control overhead) associated with periodic update of neighbor information in Fig. 10. Compared to GPSR, TMNR has higher control overhead caused by maintaining additional routing table information. This routing table is used when the destination is near.

From Figs. 8–10, we see that the performance curves of CLA are essentially flat. CBF has similar performance in terms of Fraction of Packet Delivered, End-to-End Delay, and Normalized Routing Load. This indicates that both CBF and CLA are unaffected by node mobility (i.e., speed or pause time). This means that CBF and CLA are very robust and suitable for a wide range of mobile applications. In terms of the number of duplicate packets received by destinations, most of the routing protocols have on average two packets, except for CBF and CLA (Fig. 11).
these two, CBF has three times more packet duplication (10–12 duplicate packets) compared to CLA.

4.3. Effect on number of communication sessions

In this study, we performed sensitivity analysis with respect to the number of communication sessions. We ran our simulation with speed fixed at 20 m/s, pause time at zero, and number of nodes at 200. We varied the number of communication sessions between 5 and 40.

The simulation results are shown in Figs. 12–15. Again, they show the trade-off among fractions of packets delivered, end-to-end delay, control overhead, and packet duplication. In Figs. 12 and 14, AODV and LAR achieve high packet delivered ratio with the cost of high control overhead. For AODV, each node periodically maintains the state of the routing table. This is the reason that AODV has high overhead in Fig. 14. For LAR, the overhead cost comes from maintaining and updating request zone and expected zone in Fig. 14. Inaccurate request zone causes LAR re-issue route request by flood the network. As the result, LAR has higher end-to-end delay and higher control overhead.

From the simulation, we notice that DSR does not perform well when mobility reaches above 20 m/s (see Figs. 12–14). In DSR, only source nodes maintain the route. When a route breaks, the source node will attempt to use any other route that it happens to know about or issues another route request to find a new route. However, with mobility that reaches above 20 m/s, a source node cannot robustly adapt to the changes of topology due to high mobility.

For GRID, TMNR, and GPSR, the low fractions of packet delivered ratio and high control overhead are the result of outdated information. For GRID, this out of date information is the gateway nodes in each of the grid cell. For TMNR and GPSR, this out of date information is the neighboring nodes of each node. And since only gateway nodes will forward the data packets, the gateway nodes become bottleneck due to unbalance workload in GRID.

Only Contention-based Forwarding (CBF) and connectionless approach (CLA) can robustly adapt to the changes in the number of communication sessions to maintain good performance regardless of the network conditions. By robust, we mean that both CBF and CLA achieved high successful delivered rate (see Fig. 12), low end-to-end delay (see Fig. 13), and low control overhead (see Fig. 14). Between these two, CLA has a significantly lower number of packet duplications compared to CBF. This can be attributed to the fact that CLA limits the forwarding area to a grid path.

4.4. Effect of network density

In this study we assumed the nodes move constantly at 20 m/s without pausing and that each maintains 20 concurrent communication sessions. To examine the effects of network density, we ran the simulation with different numbers of nodes: 50, 100, 150, 200, and 400 nodes.

The results of this study are plotted in Figs. 16–19. In terms of fraction of packets delivered in Fig. 16, AODV and LAR tend to perform the best out of eight routing protocols under density of 100 and 200 nodes, respectively. However, as number of node increases, the number of control packets per data packet also increases for AODV and LAR (see Fig. 18).

Again, DSR does not perform well, even if the number of node increases. For GRID, the number of gateway nodes is fixed due the number of grid is fixed (i.e., one gateway node per grid). Therefore, since only gateway nodes allow to forward data packets, increase the number of nodes in the network did not improve the performance of GRID.

For TMNR and GPSR, the increasing number of nodes in the network causes the two protocols to have more neighbor nodes to maintain. As the results, the number of control packets is also increased in Fig. 18. As density increases, the time to determine which neighboring nodes that is closer to destination/next anchor to be the next hop also increase. This selection process become more time consuming as number of nodes increase. Therefore, end-to-end delay also increases for TMNR and GPSR in Fig. 17.

Since this greedy forwarding approach will choose the next forwarding hop closest to destination and furthest from
721 current node, the connection between the current node and
722 the selected next hop will be very weak (i.e., faster out of
723 radio rage of each other). Therefore, this causes low frac-
724 tion of packets delivered rate show in Fig. 16.
725 When the node density is sufficiently high (i.e., 150
726 nodes or more), the CLA is the only scheme that consist-
727ently displays good performance under all four metrics.
728 We note that “150 nodes in a 1000 m × 1000 m terrain or
729 100 grid cells” is still a reasonably practical scenario. We
730 observe that the “end-to-end delay” curve of the connec-
731tionless technique (CLA) behaves irregularly when the
732 node density is very low, i.e., 50 and 100. This is due to
733 the fact that data packets that fail to reach their final des-
734 tination are not taken into account in the computation of
735 the end-to-end delay. As a result, the average end-to-end
736 delay is small because only packets relayed over a few hops
737 make it to the destination. This measure increases when
738 there are 100 nodes in the network because the node den-
739 sity now becomes sufficiently high to support longer hop-
740 by-hop connections. In fact, when the density drops below
741 50 nodes in 1000 m × 1000 m field (i.e., 140 m × 140 m per
742 node or three grids per nodes), TMNR, CBF, and CLA
743 can no longer forward the data packets. If we continue
744 to increase the number of nodes in the network, the Con-
745 tention-based Forwarding (CBF) and connectionless
746 approach (CLA) eventually have the option to select the
747 shorter hop-by-hop connections for each packet transmis-
748 sion, resulting in very good end-to-end delay. From this
749 point forward, the performance of the Contention-base
750 Forwarding (CBF) and the connectionless technique
751 (CLA) becomes “flat” given the fixed terrain dimensions.
752 As the density increases, the number of packet duplications
753 increases rapidly for CBF. Thus, CBF is not scalable for
754 high-density environments.

4.5. Effect of terrain area (scalability)

To study if the techniques under consideration can scale
756 up to facilitate large-area deployment, we increased both
757 the network area and the number of nodes in order to
758 maintain a constant node density (i.e., averaging
759 70 m × 70 m per node or two nodes per grid cell):
760 • 500 m × 500 m area and 50 nodes;
761 • 1000 m × 1000 m area and 200 nodes;
762 • 1500 m × 1500 m area and 450 nodes;
763 • 2000 m × 2000 m area and 800 nodes.
764
765 In these four simulation runs, we fixed the node speed at
766 20 m/s, the pause time at zero, and the number of commu-
767 nication sessions at 20.

The results are presented in Figs. 20–23. As we increase
770 the network area, the performance of DSR, GRID,
771 TMNR, and GPSR degrade very quickly in the case of
772 the fraction of data packets delivered successfully. Simi-
773larly, LAR, AODV, GRID, TMNR, and GPSR degrade
774 rapidly in the cases of end-to-end delay and control over-
775 head. When scale up the network in terms of area and
776 number of nodes, the number of hop or the distance of a
777 connection between source and destination becomes
778 longer. Therefore, route maintenance becomes more costly
779 in term of control packets per data packet for most routing
protocols. Also, a longer route has more chance to break due to any one of the node in a connection fail or out of reach. Thus, fraction of packets delivered is also lower as the scale of the network increases for DSR, GRID, TMNR, and GPSR. Although, AODV and LAR can achieve high fraction of packet delivered, both protocols do not perform well in the cases of control overhead and end-to-end delay.

Only the CBF and CLA perform well under three metrics – fraction of packets delivered, end-to-end delay, and normalized routing load. However, we observe that only CLA does not increase the number of packet duplications with the increases in the network area and the number of nodes. Thus, only CLA can scale up to support larger networks. In contrast, CBF degrades rapidly in the case of number of packet duplication when the network area and number of nodes increase.

5. Conclusion

We present a comparative study of eight routing protocols for vehicular ad hoc networks in streets environments with high mobility. The detailed simulators, implemented using GloMoSim, allow us to perform fair and accurate comparisons of these techniques with a broad range of network parameters including mobility, pause time, the number of communication sessions, density, and size of terrain.

We summarize the performance characteristics of these techniques in Table 1. In our study we observed that AODV, DSR, LAR, GRID, TMNR, and GPSR have to make a trade-off between the fraction of packets delivered, the end-to-end delay, and the normalized routing load. Although both CBF and CLA perform well in terms of all three metrics; CBF has much higher number of packet duplications compared to all of the other protocols. CLA, on the other hand, is not suitable to low-density environments (i.e., below 50 nodes in 1000 m × 1000 m field).

In Table 2, we summarize the characteristics of environments suitable for each protocol. Multi-hop based approaches, such as DSR, AODV, and LAR, are more suitable for conference or meeting applications where mobility is low, pause time is long, communication load is light, density is low, and terrain size is small. Notice that only AODV and DSR do not require any location information provided by GPS. For the Cluster-based approach in general, long pause time is needed to maintain up-to-date cluster membership. GRID can support moderate node speed because it only maintains the gateway nodes. How-

![Fig. 23. Effect of terrain area on packet duplication.](image-url)
ever, short pause time can cause GRID to constantly reelect the gateway nodes. Thus, GRID adapts well in an environment with a moderate speed and long pause time. Our simulation shows, the pause time can have great effect on GRID’s performance. However, network throughput is limited by only gateway nodes allow to forward for the cluster approach. For one-hop based approach such as TMNR and GPSR, the need to maintain neighboring information causes the performance to decrease under high mobility. Compare the two protocols, GPSR can adapt to low-density due to geographical anchor locations. For connectionless approach, both CBF and CLA perform well and suitable for most environments. However, CBF tend to have higher number of packet duplication. This leads to media congestion, waste of power, and lower network throughput. Cluster-based (i.e., GRID) and one-hop based (i.e., TMNR and GPSR) are suitable for mid-range scale of network such as disaster recovery or sensor network with location service available. Connectionless based approach (i.e., CBF and CLA) is suitable for vehicle network or battle field that has a large terrain, high node density, and nodes moving in a high speed with short pause time. Connectionless based approach can also perform well in a more static environment where nodes move in low speed and have long pause time. Between CBF and CLA, CLA tend to have higher communication load and low end-to-end delay which is ideal for voice and video application. While the results of this paper can provide guidelines, the final selection of a routing protocol should also take into account considerations specific to a given application.

References


