Technical Report, IDE0748, June 2007

# State of the art on network layer aspects for inter-vehicle communication

**Annette Böhm**

## CONTENTS:

# Introduction

In future vehicular networks, the amount of help a driver can get in avoiding dangerous traffic situations or minimizing the potential damage will increase dramatically. Information about weather and road conditions, digital maps and navigation systems are combined with sensor data from the own vehicle as well as surrounding vehicles. Fast and reliable communication between cars (vehicle-to-vehicle) and/or between a car and a road side unit (vehicle-to-infrastructure) are essential for future vehicle alert systems.

From a network perspective, this means that messages have to be routed from the information source to one or several destinations without too much administrative overhead and delay. Fast topology changes and high mobility in a vehicular network need to be considered when developing new routing protocols. In vehicle-to-infrastructure communication, vehicles enter and leave the transmission range of roadside units at a fast pace and handover techniques need to be adapted in order to maintain connectivity. The motivation for many of the current approaches is to provide the driver and/or passengers in a vehicle with the quality of service needed for multimedia applications but the ideas behind fast and seamless communication can likewise be the basis for future vehicle safety application.

This report gives a short survey over the work that has been done on adapting multi-hop routing and handover techniques to a vehicular network environment and explains some of the ideas behind the proposed solutions.

# A) Routing in vehicular networks

In Intelligent Transportation Systems (ITS) warning messages between vehicles may concern other vehicles in the immediate neighborhood. In this case, it is sufficient that the sending node can reach all nodes within its transmission range. In other cases warnings have to be forwarded to vehicles that are hundreds of meters or even kilometers away. To reach these destinations, relaying techniques have to be employed and intermediate nodes (other vehicles or communication infrastructure on the roadside) have to be used as routers to forward the data.

## *1 Requirements of routing in highly mobile ad hoc networks*

In a vehicular network, routing challenges arise mostly from the high mobility of the nodes and the need for dependable and high speed communication. Apart from that, scalability is an important issue as the network can go from scarce to dense in a very short time. We will keep these points in mind when looking at different routing strategies and solutions. Other desirable characteristics are loop-free routing paths, dynamic topology maintenance (i.e. dynamic adaptation of routing information in highly mobile topologies), minimal control overhead, low complexity and multicast capabilities.

## *2 Ad hoc routing techniques*

Ad hoc routing protocols can be divided into different groups. In a network with a flat routing algorithm, each node has the same role and responsibility in the routing process, whereas hierarchical routing assigns this responsibility only to a certain number of nodes that then provide the rest of the nodes with the required routing information. Location-based routing makes use of the node’s location information. Routing decisions may be influenced by other parameters like power efficiency or application specific requirements.

Surveys of MANET (Mobile Ad hoc NETwork) routing protocols are given by Hong et al. [Hong] and Royer et al. [Royer]. A comparison between different MANET algorithms can be found in Boukerche et al. [Boukerche].

### 2.1 Flat routing

In flat routing, each node is equally involved in the routing decision. This technique suffers from scalability issues but comprises still some of the most popular ad hoc routing techniques. Two general approaches are used, proactive or table-driven routing and reactive or on-demand routing.

## 2.1.1 Proactive, table-driven approach

Proactive routing protocols offer up-to-date routes to any destination in the network at any given time by frequently updating the routing tables. The main advantage of this approach is the minimal delay an application experiences when it wants to send information across the network. Keeping track of all the routes through the network, on the other hand, introduces a large protocol overhead. The routing table updates are usually periodic but can be extended by event-driven updates to quickly react to network changes.

There are table-driven solutions for mobile ad hoc networks, but these are not really suitable for inter-vehicle communication. Considering the quickly changing topology of a vehicular ad hoc network, route updates must be issued with very short time intervals, which influences the total throughput over the network and adds overhead. Even with very frequent route updates, the risk of using outdated routes would still be high.

## 2.1.1 Reactive, on-demand approach

Reactive routing means that routing information is not gathered in advance, but first when requested by an application. This saves control overhead but introduces the need for a discovery phase each time an application needs to send data over the network and increases the communication delay. In the route discovery phase, the network is partially or entirely flooded with route request messages to find the shortest path to the destination. In reactive routing, scalability is better because routing information is only exchanged when needed and a node can be certain to use the most recent routing information which leads to a higher transmission success than with proactive routing where routes might be outdated.

Flooding the network with request messages is only meaningful if the actual data to be sent is considerably larger than the request message. Otherwise, it is more bandwidth efficient to flood the data itself. Short warning messages with real-time requirements between vehicles might be a borderline case.

AODV (Ad Hoc On Demand Distance Vector Routing) [Perkins] and DSR (Dynamic Source Routing) [Johnson] are the two reactive ad hoc routing protocols that gained most importance and are considered as leading candidates for standardization [Hong].

- **Dynamic Source Routing (DSR)**

  In Dynamic Source Routing, the sender node determines a suitable path to the destination before sending the data (route discovery). This is done by broadcasting route requests to its neighbors, which in their turn broadcast it further throughout the network. The route request registers all nodes it passed on the way and, when received by the destination, it is returned to the sender. The actual data is then sent along this path with routing information in the packet header. The intermediate nodes only follow the instructions in the header and do not participate actively in the routing decision. Each node monitors the paths to its neighbors (route maintenance) and in case of a link failure, source nodes of ongoing transmissions are informed and a new route discovery phase is started.

- **Ad Hoc On Demand Distance Vector Routing (AODV)**

In Ad Hoc On Demand Distance Vector Routing, each node maintains a traditional routing table listing next hops to possible destinations. A sender sends out a route request which is broadcasted throughout the network until it reaches the destination. A route reply is sent back to the sender via the same paths used by the request. Each intermediate node compares the hop count with the information in its routing table and discards the proposed path if it already has a shorter path in its table. Sequence numbers are used to separate the most recent information from older routing entries. Route maintenance is accomplished by sending Hello messages between neighbors and link failures are reported throughout the network by broadcast from neighbor to neighbor similar to DSR. Boukerche et al. [Boukerche] compare (among others) DSR and AODV in terms of throughput, delay and control overhead and state that DSR has higher throughput, considerably smaller control overhead and handles the quick changes of high mobility better than AODV. On the other hand, this introduces a higher delay due to the route discovery phase.

Other flat, on-demand routing algorithms for mobile ad hoc networks, e.g. ***Temporally Ordered Routing Algorithm (TORA***) [Park], have been proposed. See the survey by Royer et al. [Royer] for information about other routing protocols.

## 2.2 Hierarchical routing

In hierarchical routing, the network is divided into clusters of nodes [Hong]. Routing decisions are made by certain nodes (cluster heads), while the rest of the nodes in a cluster receive or relay data via the cluster head. Both routing table size and the size of control packets for route discovery or maintenance are reduced considerably. This is even an advantage from a scalability point of view [Royer], as only the cluster heads need to be concerned about nodes entering or leaving the network. Hierarchical routing techniques lead in general to shorter response times as the flooding of routing information is restricted to the cluster heads and routes are discovered more quickly. On the other hand, in highly mobile networks like vehicular ones, frequently updating and renegotiating the hierarchical partition of the network complicate the routing process and add delay.

Most hierarchical routing algorithms are based on the concept of proactive routing. Examples are ***Clusterhead-Gateway Switch Routing (CGSR)*** [Chiang] that uses clusters with gateways belonging to two clusters to forward information throughout the network and cluster heads to make the routing decisions based on a routing table and a table of cluster members. ***Hierarchical State Routing (HSR)*** [Pei] is another example. Updating both the hierarchies and the routing tables frequently enough to cope with high mobility gives these traditional types of hierarchical routing protocols a disadvantage in vehicular networks.

The ***Zone Routing Protocol (ZRP)*** [Haas] can be seen as a variation of a hierarchical protocol and uses a combination of reactive and proactive routing. Each node defines a zone around itself and maintains routing information proactively inside the zone. If the destination cannot be found inside the zone a path outside the zone must be found on-demand. By changing the size of the zone, routing delay and scalability can be balanced.

Royer et al. [Royer] describe the general concepts behind hierarchical routing, while examples of hierarchical routing protocols for MANETs are given in the survey of Hong et al. [Hong].

## 2.3 Location-based, geographical routing

Flooding the whole network with route requests or routing updates is a waste of resources. If the position of the destination node is known the performance can be enhanced by integrating this position in the routing process and restrict flooding to the area where the destination is expected to be found. Vehicles equipped with a GPS receiver (Global Positioning System) have access to their coordinates in a globally known reference system. As it is not always possible to reach a sufficient number of GPS satellites to produce a correct position (e.g. in tunnels or in urban areas), other ways of positioning like triangulation or local reference systems can be used.

Many location-based routing protocols have been suggested. A common feature is that they all imply that each node knows its position at any given time and use this information to reduce routing overhead. Many of the location-based algorithms are applicable to the requirements of inter-vehicle communication. An overview of location-based routing strategies is provided by Mauve et al. [Mauve] or Stojmenovic [Stojmenovic].

Location-based routing (other frequently used names are georouting, position-based or geographical routing) consists of two stages [Mauve]. First a location service is responsible for providing a source with the position of the desired destination. In the second stage, the forwarding phase, the route from source to destination is determined.

### 2.3.1 Location Services

Popular location services are ***Distance Routing Effect Algorithm for Mobility (DREAM)*** [Basagni] and ***Grid Location Service (GLS)*** [Li].

- ***DREAM***
  In DREAM, each node maintains a location database with the positions of every other node in the network. The database entries are periodically broadcasted to all nodes within a predefined radius and the location databases are updated. A source node can access a destination's position without requesting information from other nodes, which reduces communication delay and makes it a candidate for local communication between cars in emergency situations [Mauve].

- ***GLS***
  GLS uses a different strategy. It divides the network into hierarchical squares. Each node keeps a table with location information of all other nodes in the same first-order square (the smallest square in the hierarchy). In each of the first-order squares surrounding the source node, the node with an ID closest (but higher) than the source node's ID is chosen as location server and stores the location information. Each node is assigned its unique and randomly generated ID. In each of the surrounding second-order squares a location server is chosen

in the same fashion and so on. Closer to the node, the location server density is high but the further away you get, the scarcer the location servers get. This strategy scales better but suffers from the delay introduced by finding the right location server when a node's location has to be determined.

In a vehicular network, information about a vehicle's speed and direction (maybe even in combination with knowledge of speed limits, road type and other map-based information) could be included in location services. In that way, probable positions could be calculated and frequent location table updates could be avoided.

### 2.3.2 Forwarding strategies

**Greedy packet forwarding**

The sending node includes the position of the destination in the data packet and forwards it to a node closer to the destination. Each intermediate node does the same, i.e. it chooses a next hop that brings the packet closer to the destination. Different greedy routing strategies are described in the surveys of Mauve et al. [Mauve] and Stojmenovic [Stojmenovic]. The following list describes the most commonly used ones:

- *Most forward within r* relays the packet to the node within the transmission range that is closest to the destination.
- *Nearest with forward progress* chooses the node closest to the sender but still closer to the destination than the sender itself. This strategy pays off in terms of energy savings when the transmission range is adaptable. In inter-vehicle communication, where energy aspects are not of big importance, this type of greedy routing has no advantages.
- *Compass routing* forwards the data to the node closest to a straight line between sender and receiver and thereby tries to minimize the spatial distance the packet travels.
- Even a randomly chosen next hop has been proposed.

The disadvantage of greedy packet forwarding is that it cannot be guaranteed that a path can be found even if it actually exists. The ***Greedy Parameter Stateless Routing Protocol (GPSR)*** [Karp] defines a recovery mode when this case occurs and constructs a planar subgraph of the network between sender and receiver. If a path between sending and destination node does exist, it can be found along this subgraph. As soon as a node closer to the destination than the sender is reached, the recovery mode is stopped and normal greedy forwarding is reentered. As such failures are not memorized, the same recovery process has to be repeated for each packet, adding to a large overall delay [Tian].

**Restricted Directional Flooding**

In contrast to greedy forwarding, where a packet was forwarded to a single node only, in restricted flooding the message is flooded to all nodes within a given region. There are different ways to determine such a region:

- The ***DREAM*** protocol mentioned as a location service, even defines one possible restricted flooding strategy. An *expected region* is defined by a circle around the position where the destination is expected to be found. The radius can be adapted based on the known speed of the destination and the freshness of the information. The more likely, the location information is outdated, the larger the radius. Two straight lines between the sender and each side of the circle define the area to be flooded. See figure 1.

- Location Aided Routing (***LAR***) [Ko] uses a similar strategy. It is not a routing protocol itself but it aims at restricting the flooding of route requests in the route discovery phase of a reactive protocol. In LAR, a circular *expected zone* around the destination is created based on available position information and a *request zone* is defined as the smallest rectangle containing both the source and the expected zone. The route request is then flooded within the request zone. See figure 2.

None of the strategies scales very well to large networks but are very robust against node failure and position inaccuracies [Mauve].

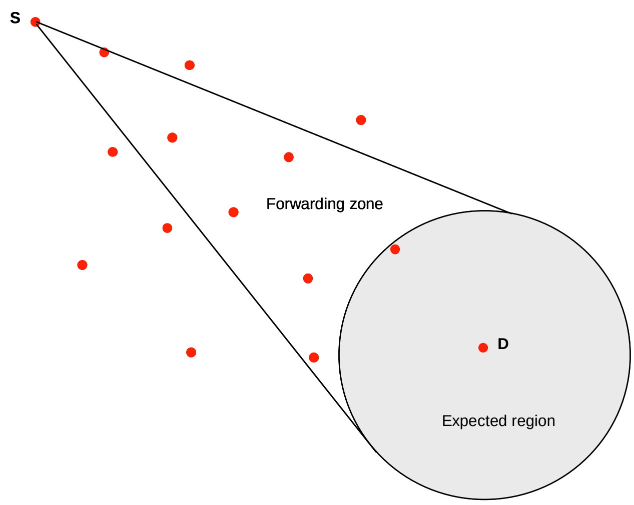


Figure 1. Restricted Directional Flooding - DREAM

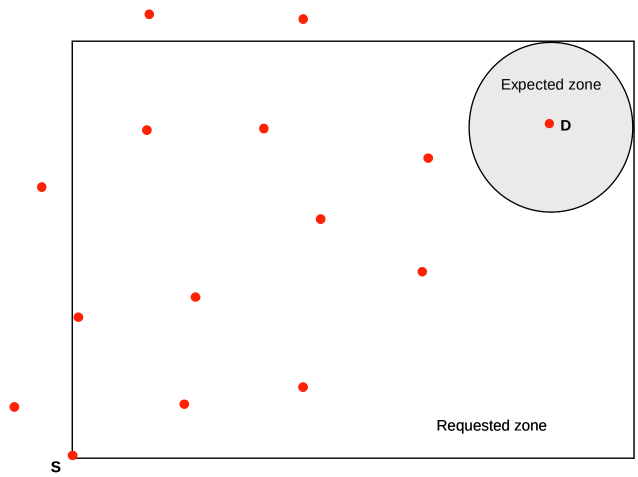


Figure 2: Restricted Directional Flooding – LAR

### 2.3.3 Hybrids or improvements to location-based routing

- ***Context Assisted Routing (CAR) [Dumitrescu]***

  [Dumitrescu] describes an extension to both location services and the greedy forwarding strategy by introducing context-assisted graphs. These graphs are based upon information about the road infrastructure and vehicle-specific data as speed and direction.

  Normal greedy forwarding as described in the GPSR protocol is often forced into recovery mode because the greedy choice of the next hop not always confirms with the actual road infrastructure. Figure 3a shows a scenario where source node S forwards its packet to the node that would bring the most geographical progress towards destination D, ultimately leading into a dead end. Recovery mode finds the way back to the right path but this means unnecessary delay. The Context Assisted Routing (CAR) protocol sets up a graph according to the road infrastructure (see figure 3b) and forwards the packets along this path (see figure 3c).

  The vertices of the graph can also be used to minimize communication overhead in the location service. A node sends out a location update when it enters a new vertex stating its direction (which edge it comes from and which edge it travels towards) and its speed. Based on this information, the position within the vertex zone can be estimated and no further location update is needed until the vehicle enters a new vertex.

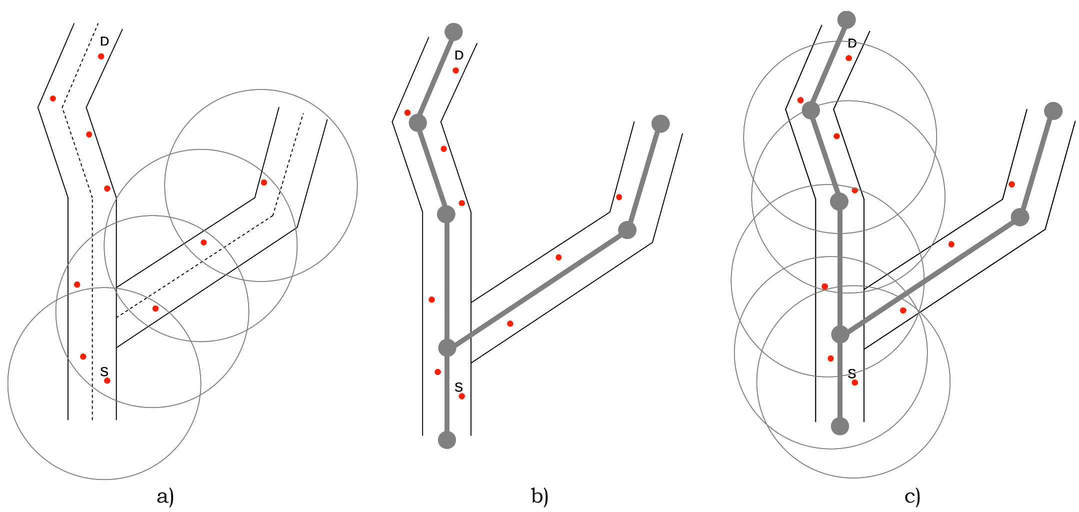


Figure 3: Context Assisted Routing

- ***Location Routing Algorithm with Cluster-Based Flooding (LORA_CBF) [Santos]***

  The LORA_CBF protocol is developed with an ITS scenario in mind and combines the advantages of georouting, hierarchical routing and reactive routing. Clusters of nodes are established with a simple Hello message approach finding cluster heads and gateways between neighboring clusters. When a node wants to transmit data to a certain destination and there is no (or outdated) location information available at the cluster head, a location request is sent out to other cluster heads in the network via gateway nodes. As soon as a cluster head can provide the information, a location reply is sent back to the sender. Both sender and destination send location updates in data and acknowledgement packets to keep the location information as correct as possible during the ongoing transmission.

The actual data transfer follows the *most forward within r* approach but does not define any recovery mode to deal with hidden paths. Nevertheless, the LORA_CBF outperforms traditional ad hoc routing algorithms like AODV and DSR in terms of route discovery time, end to end delay and delivery ratio.

(The same protocol and simulation evaluation was published by the authors as ***Cluster-Based Location Routing (CBLR)*** [Santos2]).

- ***Position-based Quality of Service (QoS) Routing [Erbas]***

The solution by Erbas et al. uses a grid-based hierarchy with cluster heads maintaining not only location information, but even information on available bandwidth and the link state in the network. Updates are sent to the cluster head periodically or on demand within each cell of the grid and then forwarded to other cluster heads in the network. Thereby, a global knowledge is achieved and a cluster head can provide the sender with information about both the destination's position and possibility of setting up a QoS link. There must of course be support from other layers (e.g. Medium Access Control protocols) to give any QoS guarantees.

- ***Spatially Aware Packet Routing (SAR) [Tian]***

In order to overcome the problem of topology holes (i.e. areas that lack active nodes), Tian et al. suggest the use of a spatially aware routing algorithm. As the occurrence of topology holes often is closely related to the road infrastructure, relevant spatial information is extracted from a Geographic Information System (GIS) to build a graph of intersections (vertices) and road segments (edges). In this aspect the approach is very similar to the ***CAR*** algorithm suggested by Dimitrescu et al.

Depending on the memory capability, only major roads and intersections can be used or the information can be extended to the entire road net including additional data on the length of the road segment, speed limits, number of lanes etc. This information can be used to assign weights to the edges of the graph. The sender uses the graph to find the shortest path to the destination and the packet forwarding is done via vehicles along the path. In case no suitable intermediate node can be found, normal greedy forwarding is tested, an alternative graph is computed or the packet is suspended in a buffer until a forwarding vehicles comes along.

The authors compared the SAR protocol and a version with suspension buffer (called SARB) to standard GPSR and showed that SAR gives the shortest delay and highest packet delivery ratio.

### 2.3.4 Advantages and disadvantages

For georouting purposes it is usually sufficient for a node to know its own position, the position of its immediate neighbors and the position of the destination node acquired through the location service. Therefore, no routes have to be established or maintained and most of the routing overhead is avoided. This makes most georouting approaches scale well and adapt quickly to a highly mobile network [Mauve].

The main disadvantage is that the position of the destination node has to be known for georouting to work and that location information has to be spread through the network. Maintaining a location service itself and/or supplying neighbors (or other parts of the network) with location information adds overhead and takes away much of the gain from low routing overhead. The design of location update schemes is often more difficult than the routing decisions themselves.

Apart from that, the position offered by e.g. a GPS receiver can be inaccurate or unavailable due to obstacles. To work properly, all location-based algorithms rely on correct positions.

Location-based routing protocols have problems handling topology holes [Dumitrescu]. In scarcely trafficked areas or in a situation where only few of the vehicles are equipped with communication devices, this might become a problem for inter-vehicle communication scenarios.

In a vehicular network with very high node mobility, location-based routing can avoid extensive flooding of routing information, especially when traffic patterns, speed etc. are known and can be used in determining restricted flooding areas.

## 2.4 Multicast/Geographical multicast

Considering vehicle alert messages, the information sent out from one source often concerns not only one destination node but all nodes in a certain alert area. Multicast is therefore an important feature. Maihöfer gives a survey of current geocast protocols [Maihöfer].

- ***Location-Based Multicast [Ko2]***

  Location-based multicast (LBM) is based on the concepts of location-based routing (LAR) and directional flooding. It defines a destination region and a flooding zone and discards packets outside this zone. As in LAR, the flooding zone is a rectangle including at least the sender and the destination area. Another approach is to only allow intermediate nodes that are closer than a certain distance threshold from the destination area. A node calculates its geographical distance from the center of the destination region and determines if it is to forward the packet or to drop it. The distance is decreased with each hop.

- ***GeoGRID [Liao]***

  GeoGRID enhances the LBM approach by dividing the forwarding region into grids where only preselected gateway nodes are relaying the packets until the destination region is reached where the packets are flooded.

- ***Unicast Routing with Area Delivery [Maihöfer]***

  In his survey, Maihöfer proposes to use any unicast georouting protocol of choice to reach the destination area and simply flood the region. He calls this suggestion Unicast Routing with Area Delivery.

- ***Inter-Vehicle Geocast [Bachir]***

  The Inter-Vehicle Geocast Protocol (IVG) is developed for relaying emergency messages in a highway accident situation. A broken down vehicle sends out an alarm to approaching vehicles. The packets are only relayed by nodes inside a risk area, i.e. amongst vehicles approaching the accident site, not by vehicles driving away from the site. A node inside the risk area waits a given amount of time until it broadcasts the message in its transmission range. If it during this time receives the same message from a node behind it, it knows that another node has taken over the role as relay node and drops the packet. As long as this in not the case, the node rebroadcasts the message with a certain time interval. There is no time-to-live limit to prevent infinite broadcasting in order to keep messages alive in light-crowded network as a highway at night time. By defining the risk area depending on the vehicles speed, the packets are not spread to vehicles that do not need the information.

# B) Mobility Management

In a vehicle alert system, different kinds of information have to be collected to help the system (and in the long run the driver) to make appropriate decisions. Direct communication of emergency messages between vehicles has to be supported, as well as access to digital map updates and weather information via WLAN at the roadside or the cellular network etc. The main challenge of mobility management is to seamlessly integrate several categories of wireless network technologies that differ amongst others in terms of bandwidth, cellular coverage and quality of service.

The term mobility management comprises two stages: location management and handover (handoff) management [Sun]. Location management deals with locating the mobile node, tracking its movement and updating the location information. Handover management is mostly concerned with the efficient and smooth change of a mobile node's access point during an ongoing data transmission and is the main focus of this chapter.

## *1 Location Management*

Techniques to determine a node's current position include Time Of Arrival (TOA), Time Difference Of Arrival (TDOA) and Angle Of Arrival (AOA), amongst others [Kyriazakos]. In the case of vehicular networks, we can assume that each communicating vehicle has a GPS receiver that gives the current position with a precision of less than 10 meters. A comparison of satellite-based positioning (GPS, Gallileo) with different positioning techniques in terms of cost, latency and accuracy is not in the scope of this report.

## *2 Handover management*

Handover mechanisms deal with nodes moving from one access point or base station to another. This procedure is usually associated with packet loss and delay as it takes both time and resources to terminate a connection, to establish a new one and to handle problems like addressing. Optimally, the handover is handled so smoothly that the user or the application does not experience any effects at all. Research in this field has been done for cellular networks where mobile phone users are handed from base station to base station and for Wireless LAN (WLAN) where e.g. a continuous connection from a laptop to the Internet is maintained via changing access points.

For a vehicle alert system, the communication delay in a cellular system is large for emergency messages and the relevant handover procedures are only presented very shortly here. The financial cost in terms of service provider fees associated with this type of wireless network is another limiting factor.

WLAN on the other hand, especially with the upcoming, vehicular network adapted standard IEEE 802.11p, is better suited for high speed, communication, both for a road side unit and for inter-vehicle communication. Internet Access via WLAN can provide both entertainment and important information about e.g. weather and road situations and can be used to update digital maps that might play a central role for routing protocols and application layer decisions. As WLANs normally are IP-based networks, a short introduction to the handover procedure in Mobile IP (especially Mobile IPv6) is given and important enhancements are described.

## 2.1 General handover procedure

In general, handover can be divided into three stages [Sun]:

- ***Handover triggering***
  The handover process is initiated by a change of circumstances. Examples are the deterioration of the signal strength, an announcement from a new and better access point, a network topology change or the decrease of available bandwidth. Handovers can be triggered by the network (e.g. Access Points) or by the mobile node itself.
- ***Connection re-establishment***
  A new access point is discovered and a new connection between the mobile node and the access point is established.
- ***Packet routing***
  The route between the mobile node and its correspondents is adapted.

Intra-technology handover, i.e. the change to another access point within the same wireless technology, is called horizontal handover, while inter-technology handover process between different technologies is called vertical handover [Sun]. (Vertical) handovers are not only triggered by loss of connectivity but also by the need for a different loss rate, bandwidth, power consumption, transmission range etc.

## 2.2 Handover in IP networks

A IP address of a device is used as a unique identifier of the device's attachment to a net via a specific access point. When the mobile node connects to another access point, a new, temporarily assigned IP address is created and datagrams addressed to the old address are lost [Stallings]. To overcome this, the IETF (Internet Engineering Task Force) presented Mobile IP and later Mobile IPv4 and Mobile IPv6 [D.Johnson]. Recently, additional enhancements to manage high node mobility (Hierarchical Mobile IPv6 [Castelluccia] and the Fast Handover Protocol [Koodli]) were presented.

Each mobile node, i.e. each vehicle in the vehicular network, needs an individual, global IP address to be uniquely identifiable [Bechler]. The address space offered by the 32-bit IPv4 addresses is not enough and therefore the focus here lies on the deployment on 128-bit IPv6 addresses which offers enough capacity for future, large-scale vehicular networks.

### 2.2.1 Mobile IP

Each device is assigned a home address within its home network [Stallings]. When the device moves into a foreign network, it registers with the local agent (foreign agent). The foreign agent contacts its counterpart in the device's home network (home agent) and announces a care-of-address that can be used to relay packets to the network where the device currently resides. Packets to the mobile node are sent to its home address, where the home agent encapsulates it into a new IP datagram with the care-of-address as destination address in the header and tunnels it to the foreign network. The movement of a mobile node is thereby made transparent to higher layer applications. See figure 4.

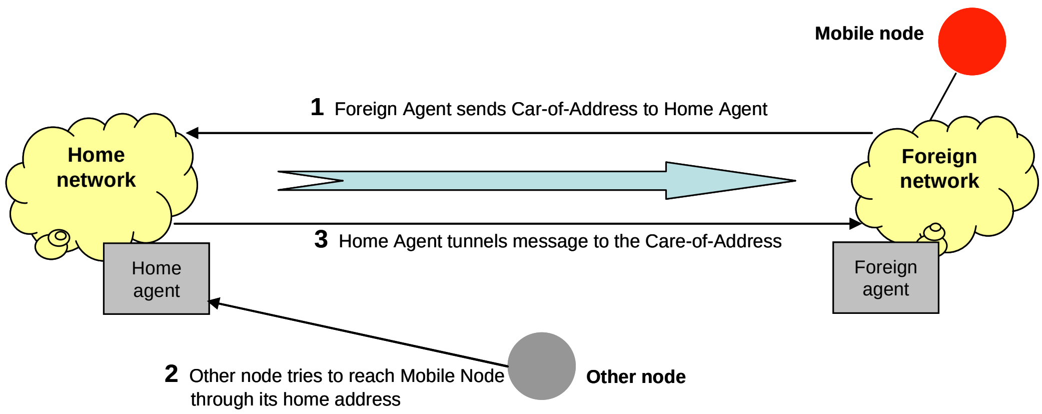


Figure 4: Mobile IP

### 2.2.2 Enhancements to Mobile IP

Establishing the care-of-address and notifying the home agent introduces a handover latency that is considered to be too long for certain real-time applications [Montavant]. Protocols to improve the handover procedure were proposed by the IETF and are published as RFCs (Request For Comment).

- **Mobile IPv6**

  Mobile IPv6 introduces enhancements to the Mobile IP mechanisms that reduce both signaling overhead and handover latency. Two enhancements are interesting for a vehicular network with its high node mobility.

  ***Route optimization mode***
  The protocol allows information about the binding between a mobile node and its care-of-address to be stored at other communication nodes [D.Johnson]. A node that tries to reach a mobile node learns about the new care-of-address and sends its further traffic directly to the mobile node's current network. This implies that the corresponding nodes implement IPv6 but reduce on the other hand the route each packet takes through the network by avoiding unnecessary tunneling via the home agent. Further, possible congestions at the home agent are avoided, as well as the risk of the home agent as a single point of failure. Knowledge about new care-of-addresses can also be obtained by binding updates sent out by the mobile node.

  ***Multiple care-of-addresses***
  A mobile node can be associated with several care-of-addresses [Montavont]. Packets are then duplicated and simultaneously sent to all the different networks specified in the care-of-addresses. This is especially interesting when the mobility is high and the direction of movement is known in advance. The mobile node can then obtain and manage care-of-addresses of upcoming networks before even entering it.

- **Hierarchical Mobile IPv6**

  A protocol for hierarchical Mobile IPv6 was proposed by Castelluccia et al. [Castelluccia] and published as RFC in 2005 (see figure 5). The performance of

Mobile IP during handover is improved by introducing a Mobility Anchor Point (MAP) that is common for several access routers (AR). When a mobile node moves into a MAP domain, it obtains a regional care-of-address (RCoA) from the MAP and a local care-of-address (LCoA) from the access router. The MAP maps the regional and the local care-of-addresses. The mobile node's home agent only gets to know the RCoA and forwards packets to the MAP which in its turn sends them further to the appropriate access router in its domain.

Local changes (i.e. when the mobile node moves between access routers in the MAP domain) are hidden from the home agent and binding updates sent to the home network are reduced. The communication delay for handover between access routers within one domain is thereby decreased.

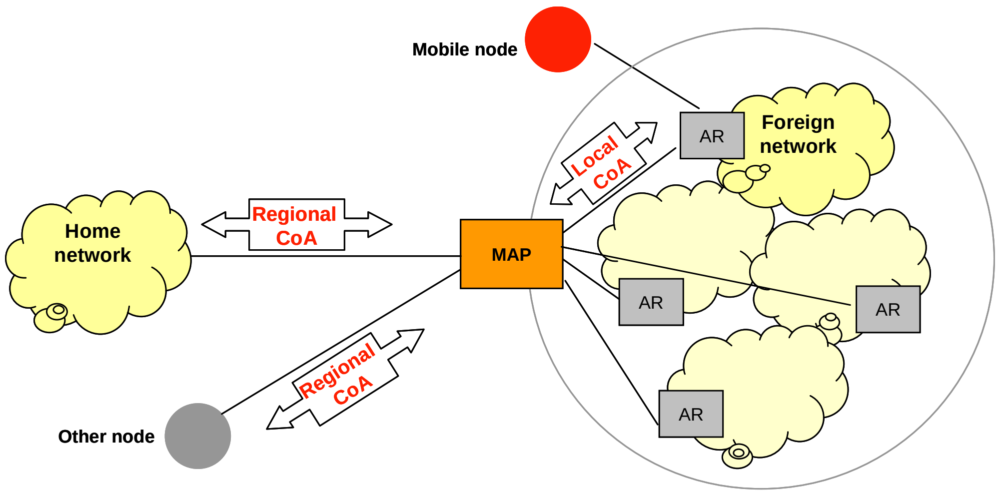


Figure 5: Hierarchical Mobile IPv6

- **Fast Handover Protocol for Mobile IPv6**

The specification for the Fast Handover Protocol [Koodli] deals with the issue of allowing the mobile node to send packets immediately after the detection of a new link and thereby shortening the handover latency. This latency is a combination of the movement detection latency, the time it takes to configure a new care-of-address and the latency from sending and processing a binding update at the corresponding nodes. If these processes can be started while the mobile node still is connected to the old network, unnecessary latencies can be avoided.

To perform layer 2 handover (i.e. handover between two access points at the link level), the node scans its environment for other access points. This information can be used to start a layer 3 handover process in time (i.e. handover between different subnets on network level). The mobile node immediately sends a request for a new care-of-address to the access router of the current network. The new care-of-address is forwarded to the mobile mode and the new network for validation and when the mobile node moves into the new network, the address can directly be used for communication. This mechanism is called anticipated handover [Montavont].

A contrasting solution called tunnel-based handover lets the mobile node enter a new network without its knowledge. Only layer 2 handover is performed and the packets to the node are sent via the old access router and a tunnel established between the old and the new network. First later, in parallel with the communication, a new care-of-address is determined and announced to the mobile node.

- **Hybrid solutions**

  Pérez-Costa et al. [Pérez-Costa] compare the performance of Mobile IPv6, Hierarchical Mobile IPv6 and Fast Handovers for Mobile IPv6 to their own protocol which is a combination of the latter two Mobile IPv6 enhancements.

## 2.3 Handover in cellular networks

Due to the reasons stated before, handover in cellular networks is not treated in depth, neither is any kind of vertical handover. Only short examples of different handover approaches are given.

In [Aust], Aust et al. describe a solution for vertical handover in a highly mobile network. The paper deals with handover procedures between WLAN and a cellular network, in this case GPRS, and includes AODV routing in the process.

Markopoulos et al. [Markopoulos] evaluate existing handover protocols for handover in cellular networks on a GSM and UMTS network simulator and compare the results to their proposed, location-based solution ***Location Aided Handover (LAH)***. A mobile node's position and velocity are used to determine if it is moving towards another cell, a handover procedure is then initiated and resources are bound for this user before it crosses over to the new cell. Another approach taken in the paper is to divide the cell into pixels and to keep a record with parameters such as the packet drop rate in each pixel. These parameters are combined in a formula to determine a cost threshold for the pixel. As soon as a mobile node experiences costs above the threshold value, it examines the possibilities for handover. Simulations show that the LAH algorithm reduces the call drop rate considerably. In [Kyriazakos], Kyriazakos et al. take a similar approach.

## 2.4 Handover in vehicular networks

The fact that cars move along a predefined road and tend to keep their direction of movement can be used in mobility management for vehicular networks.

***Forward Loss Recovery***

In [Paik], Paik et al. minimize packet loss during handover by predicting the next access point along the path of a bus in order to give passengers seamless access to the Internet via heterogeneous wireless technology (cellular network, WLAN or satellite). The regularity of routes in public transport makes predictions even easier, but it is stated that the scheme is adaptable to the irregular travel patterns of normal vehicles by employing a digital map of the access points and a navigation system. The scheme is based upon Mobile IPv6.

As soon as a mobile node looses connectivity, a ***Forward Loss Recovery*** scheme anticipates the next access point likely to associate with the mobile node and announces this to the node's home agent. All traffic to the mobile node is then routed to the expected future access router and stored until the vehicle reaches the transmission range. The paper does not give any details about how the scheme is initiated or how the routing and storing are done.

### *Mobility management with multi-hop capabilities*

Vehicle-to-vehicle communication can be used to connect a roadside unit to mobile nodes that are not in its immediate transmission range. In [Bechler], Bechler et al. propose a mobility management protocol called ***MMIP6*** that relies on IPv6 global addressing and a modification of Mobile IP to ensure that messages to a certain vehicle is always routed via an appropriate access point even if the vehicle itself is not in direct reach of any access point.

Each mobile node has a home agent associated with the node's home network. Unlike Mobile IPv6, even a foreign agent is introduced located at the access point at the road side unit. The foreign agent represents the vehicle at its current location, hiding the multi-hop capabilities of the VANET. Each mobile node keeps its original home address no matter where in the network it resides and no care-of-addresses are used. Packets from a corresponding node are always sent to the vehicle's home network where the home agent tunnels them to the foreign network where a multi-hop VANET routing protocol of choice takes over.

Foreign agents announce their services periodically to mobile nodes in their vicinity by hop- or lifetime-restricted broadcast or a geocast technique. The mobile nodes decide when to initiate a handover process which is then handled by the foreign and the home agent.

Although the authors of the paper do not try to find the fastest handover solution, the integration of vehicle-to-vehicle and vehicle-to-infrastructure is very interesting from a mobility management point of view.

### *Moving base stations*

A very different solution was proposed by Gavrilovich et al. [Gavrilovich]. In order to reduce the number of handovers, a number of access points move along with the traffic stream on rail-based loop situated in between the two directions of a major highway (see figure 6). Thereby, handover is only necessary when entering a new loop. Including an additional loop to adapt to a different traffic speed in the opposite direction, the authors estimate the total installation costs per mile to be around 4 million US dollars which reduces the feasibility of the approach.

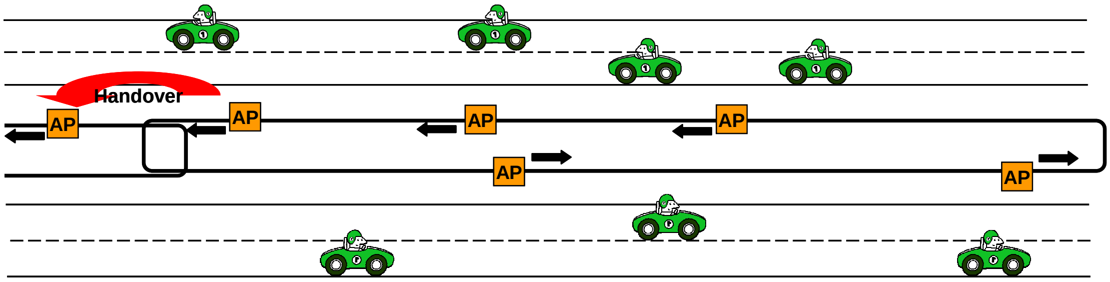


Figure 6: Moving access points

# Conclusion

Although much research has been done in both ad hoc routing and handover enhancement, there are not many tailored solutions for the special demands of a vehicular network, like e.g. maintaining a high QoS despite the highly mobile and dynamic nature of the network. The fact that the mobility amongst cars is restricted to a predefined road infrastructure and underlies certain patterns of regularity could still be explored more. The access to fine grain location information, sensor data from the car and advanced digital maps can be exploited in both routing protocol design and handover enhancements.

# References

[Aust]
S. Aust, C. Görg and C. Pampu, *Mobile IPv6 Ad Hoc Gateway with Handover Optimization,* Proceedings of 61st IEEE Vehicular Technology Conference VTC 2005-Spring, vol.4, pp. 2439-2443, Stockholm, Sweden, June 2005.

[Bachir]
A. Bachir and A. Benslimane, *A Multicast Protocol in Ad Hoc Networks Inter-Vehicle Geocast,* Proceedings of the 57th IEEE Vehicular Technology Conference VTC 2002-Spring, pp. 2456-2460, Jeju, Korea, April 2002.

[Basagni]
S. Basagni, I. Chlamtac, V.R. Syrotiuk, B.A. Woodward, *A Distance Routing Effect Algorithm for Mobility (DREAM),* Proceedings of the 4th ACM/IEEE International Conference on Mobile Computing and Networking MOBICOM '98, pp. 76-84, Dallas, TX, USA, 1998.

[Bechler]
M. Bechler and L.Wolf, *Mobility Management for Vehicular Ad Hoc Networks,* Proceedings of the 61st IEEE Vehicular Technology Conference VTC 2005-Spring, vol. 4, pp. 2294-2298, Stockholm, Sweden, June 2005.

[Boukerche]
A. Boukerche, *Performance Comparison and Analysis of Ad Hoc Routing Algorithms,* Proceedings of the IEEE International Conference on Performance, Computing and Communications 2001, pp. 171-178, Phoenix, AZ, USA, April 2001.

[Castelluccia]
C. Castelluccia, K. Malki, H. Soliman, L. Bellier, *Hierarchical Mobile IPv6 Mobility Management (HMIPv6),* RFC 4140, Aug. 2005.

[Chiang]
C.C.Chiang and M. Gerla, *Routing and Multicast in Multihop, Mobile Wireless Networks,* Proceedings of the 6th International Conference on Universal Personal Communications, vol. 2, pp. 546-551, San Diego, CA, USA, Okt. 1997.

[Dumitrescu]
V. Dumitrescu and G. Jinhua, *Context Assisted Routing Protocol for Inter-Vehicle Wireless Communication,* Proceedings of the IEEE Intelligent Vehicles Symposium 2005, pp. 594 – 600, Las Vegas, NV, USA, 6-8 June 2005.

[Erbas]
F. Erbas, J.E. Garcia and K. Jobmann, *Position-Based QoS Routing in Mobile Ad Hoc Networks: Problem Statement and a Novel Approach,* Proceedings of the 3rd IEEE International Conference on Performance, Computing and Communications, pp. 619-623, Phoenix, AZ, USA, April 2004.

[Füssler]
H. Füssler, J. Widmer, M. Mauve, H. Hartenstein, *A Novel Forwarding Paradigm for Position-Based Routing with Implicit Addressing*, Proceedings of the IEEE Computer and Communications Workshop CCW2003, Dana Point, CA, USA, October 2003.

[Gavrilovich]
C.D.Jr. Gavrilovich, *Broadband Communication on the Highways of Tomorrow,* IEEE Communications Magazine, vol. 39, issue 4, pp. 146-154, April 2001.

[Haas]
Z.J. Haas and M.R. Pearlman, *The Performance of Query Control for the Zone Routing Protocol,* ACM/IEEE Transactions on Networking, vol. 9, no. 4, pp. 427-438, Aug. 2001.

[Hong]
X. Hong, K. Xu, M. Gerla, *Scalable Routing Protocols for Mobile Ad Hoc Networks*, IEEE Network, vol. 16, issue 4, pp. 11-21, July-Aug. 2002.

[D.Johnson]
D. Johnson, C. Perkins, J. Arkko, *Mobility support in IPv6,* RFC 3775, June 2004.

[Johnson]
D.B. Johnson, D.A. Maltz and J. Broch, DSR: *The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks,* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 2001.

[Karp]
B. Karp and H.T. Kung, *Greedy Perimeter Stateless Routing for Wireless Networks,* Proceedings of the 6th ACM International Conference on Mobile Computing and Networking, MOBICOM '00, pp. 243-254, Boston, MA, USA, Aug. 2000.

[Ko]
Y.B. Ko and N.H. Vaidya, *Location-Aided Routing (LAR) in Mobile Ad Hoc Networks,* ACM Journal on Wireless Networks, vol.6, no. 4, pp. 307-321, 2000.

[Ko2]
Y.B. Ko and N.H Vaidya, *Geocasting in Mobile Ad Hoc Networks: Location-Based Multicast Algorithms,* Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications WMCSA99, pp. 101-110, New Orleans, LA, USA, Feb. 1999.

[Koodli]
R. Koodli (Ed.), *Fast Handovers for Mobile IPv6,* RFC 4068, July 2005.

[Kyriazakos]
S. Kyriazakos, D. Drakoulis and G. Karetsos, *Optimization of the Handover Algorithm based on the Position of the Mobile Terminals,* Proceedings of the IEEE Symposium on Communications and Vehicular Technology SCVT 2000, pp. 155-159, Leuven, Belgium, Okt. 2000.

[Li]
J. Li, J. Jannotti, D.S.J De Couto, D.R.Karger and R. Morris, *A Scalable Location Service for Geographic Ad Hoc Routing,* Proceedings of the 6th ACM International Conference on Mobile Computing and Networking, MOBICOM '00, pp. 120-130, Boston, MA, USA, Aug. 2000.

[Liao]
W.H. Liao, Y.C. Tseng, K.L Lo and J.P. Sheu, *GeoGRID: A Geocasting Protocol for Mobile Ad Hoc Networks based on GRID,* Journal of Internet Technology, vol. 1, no. 2, pp. 196-213, Dec. 2002.

[Lochert]
C. Lochert, H.Hartenstein, J.Tian, H. Füssler, D. Herrmann, M.Mauve, *A Routing Strategy for Vehicular Ad-Hoc Networks in City Environments*, Proceedings of the 58th IEEE Semiannual Vehicular Technology Conference VTC 2003-Fall, pp. 156-161, Orlando, FL, USA, October 2003.

[Maihöfer]
C. Maihöfer, R. Eberhardt, *Geocast in Vehicular Environments: Caching and Transmission Range Control for Improved Efficiency*, Proceedings of the IEEE Intelligent Vehicles Symposium IVS, pp. 951-856, Parma, Italy, June 2004.

[Maihöfer2]
C. Maihöfer, *A Survey of Geocast Routing Protocols,* IEEE Communications Surveys & Tutorials, vol. 6, no. 2, pp. 32-42, 2004.

[Markopoulos]
A. Markopoulos, P. Pissaris, S. Kyriazakos and E. Sykas, *Optimized Handover Procedure Based on Mobile Location in Cellular Systems,* Proceedings of the 14th IEEE Conference on Personal, Indoor and Mobile Radio Communications PIMRC 2003, vol. 3, pp. 2490-2494, Beijing, China, Sept. 2003.

[Mauve]
M. Mauve, A. Widmer, H. Hartenstein, *A Survey on Position-Based Routing in Mobile Ad Hoc Networks,* IEEE Network, vol. 15, issue 6, pp. 30-39, Nov.-Dec. 2001.

[Montavont]
N. Montavont, T. Noel, *Handover Management for Mobile Nodes in IPv6 Networks,* IEEE Communications Magazine, vol. 40, issue 8, pp. 38-43, Aug. 2002.

[Paik]
E.K. Paik and Y. Choi, *Seamless Mobility Support for Mobile Networks on Vehicle across Heterogeneous Wireless Access Networks,* Proceedings of the 57th Vehicular Technology Conference VTC 2003-Spring, vol. 4 pp. 2437-2441, Jeju, Korea, April 2003.

[Park]
V.D. Park, M.S. Corson, *A Highly Adaptive Routing Algorithm for Mobile Wireless Networks,* Proceedings of INFOCOM´97, pp. 1405-1413, April 1997.

[Pei]
G. Pei, M. Gerla, X. Hong and C.C. Chiang, *A Wireless Hierarchical Routing Protocol with Group Mobility,* Proceedings of the IEEE Wireless Communications and Networking Conference WNCN'99, vol. 3, pp. 1538-1542, New Orleans, LA, USA, Sept. 1999.

[Pérez-Costa]
X. Pérez-Costa, M. Torrent-Moreno, H. Hartenstein, *A Performance Comparison of Mobile IPv6, Hierarchical Mobile IPv6, Fast Handovers for Mobile IPv6 and their Combinations,* ACM SIGMOBILE Mobile Computing and Communications Review, vol. 7, issue 4, pp. 5-19, 2003.

[Perkins]
C.E. Perkins, E.M. Royer, *Ad-Hoc On-Demand Distance Vector Routing,* Proceedings of the 2nd IEEE Workshop on Mobile Computer Systems and Applications 1999, pp. 90-100, Feb. 1999.

[Royer]
E.M. Royer, C. Toh, *A review of current routing protocols for ad hoc mobile wireless networks,* IEEE Wireless Communications Magazine, vol. 6, issue 2. pp. 46-55, April 1999.

[Santos]
R.A. Santos, R.M. Edwards, L.N. Seed and A.E. Edwards, *A Location-Based Routing Algorithm for Vehicle to Vehicle Communication,* IEEE Communications Magazine, vol. 43, issue 3, pp. 101-106, March 2005.

[Santos2]
R.A. Santos, R.M. Edwards, A. Edwards, *Cluster-Based Location Routing Algorithm for Vehicle to Vehicle Communication,* Proceedings of the IEEE Radio and Wireless Conference 2004, pp. 39-42, Sept. 2004.

[Stallings]
W. Stallings, *Wireless Communications and Networks,* Upper Saddle River, NJ, USA: Prentice Hall, 2001.

[Stojmenovic]
I. Stojmenovic, *Position-Based Routing in Ad Hoc Networks,* IEEE Communications Magazine, vol. 40, issue 7, pp. 128-134, July 2002.

[Sun]
J.Z. Sun and J. Sauvola, *Mobility and Mobility Management: A Conceptual Framework,* Proceedings of the 10th IEEE International Conference on Networks ICON 2002, pp. 205-210, Singapore, Aug. 2002.

[Tian]
J. Tian, L. Han, K. Rothermel, *Spatially Aware Packet Routing for Mobile Ad Hoc Inter-Vehicle Radio Networks,* Proceedings of the IEEE Intelligent Transportation Systems Conference ITS 2003, pp. 1546-1551, Shanghai, China, Oct. 2003.