SECOND QUARTER 2004, VOLUME 6, NO. 2

IEEE COMMUNICATIONS SURVEYS
The Electronic Magazine of Original Peer-Reviewed Survey Articles

www.comsoc.org/pubs/surveys

A SURVEY OF GEOCAST ROUTING PROTOCOLS

CHRISTIAN MAIHÖFER
DAIMLERCHRYSLER AG, RESEARCH & TECHNOLOGY (RIC/TC)

ABSTRACT

Geocasting is the delivery of a message to nodes within a geographical region. With geocast, new services and applications are feasible, such as finding friends who are nearby, geographic advertising, and accident or wrong-way driver warning on a motorway. In this article we present a survey on geocast routing protocols. The protocols mainly differ in whether they are based on flooding, directed flooding, or on routing without flooding, and whether they are suitable for ad hoc networks or for infrastructure networks. Based on these criteria we propose a classification of geocast protocols. Our protocol comparison includes message and memory complexity, robustness, and the ability to deliver geocast packets in partially partitioned networks. Finally, we present simulations to compare the approaches based on flooding, directed flooding, and routing without flooding.

The current Internet has no knowledge about the geographic location of the nodes it services. With cheap and small form factor Global Positioning System (GPS) receivers, and with evolving ad hoc networks and ubiquitous computing, it is likely that location information of network nodes or regions will become available. This is why recent proposals of point-to-point routing protocols for ad hoc networks make use of geographic knowledge to enhance their efficiency and scalability [1].

Besides using geographic information for unicast routing purposes, it is desirable to explicitly send a message selectively to a geographic region, which is the geocast problem. Geocast aims to send a message to some or all nodes within a geographic region. However, as geocast routing protocols similar to most other routing protocols do not guarantee reliability, not all nodes inside a geographic area may be reached. Furthermore, there are protocols that make it possible to refine the geographic region with a multicast group. This makes it possible, for example, to address only pedestrians or only vehicles inside a geographic region.

With geocast, new services and applications are feasible, such as finding friends who are nearby, geographic advertising, or more general position-based publish-and-subscribe services. Of particular interest is geocast in the automotive domain, making possible the implemention of virtual traffic signs, for example, for accident or wrong-way driver warning on a motorway. Indeed, geographic information and with it geocast, which uses this information, adds a new dimension to computer networks that will make possible the implemention of many promising applications.

In this article we present an overview of all known geocast routing protocols and discuss their characteristics, e.g, path strategy, scalability, message complexity, memory requirements, and robustness. The considered protocols are Location Based Multicast (LBM), Voronoi-based routing, Mesh, GeoGRID, GeoNode, and Temporally Ordered Routing Algorithm for Geocast (GeoTORA). Geocast protocols can be mainly categorized based on whether they are designed for infrastructure networks such as the current Internet or for ad hoc networks, and whether they are based on flooding the network or on forwarding a geocast packet on a particular routing path. An example of a taxonomy is given in Fig. 1.

None of the proposed protocols is based on naive flooding, that is, flooding of a whole network without trying to limit the flooding area. However, we briefly describe a naive flooding protocol in order to compare it with more complex solutions. We call such a naive protocol the simple flooding approach. Directed flooding tries to limit the message overhead and network congestion of naive flooding by defining a forwarding zone, which comprises a subset of all network nodes. The forwarding zone includes at least the sender of a geocast message and the destination region of the message and additionally should include a path between sender and destination region. If the last condition is not fulfilled, protocols either have to increase the forwarding zone or fall back on simple flooding. An intermediate node forwards a packet only if it belongs to the forwarding zone. Directed flooding protocols differ in the manner in which the forwarding zone is defined.

Non-flooding approaches do not use flooding to reach the destination region of a geocast but other routing protocols. Note that this refers only to the wide-area routing before the destination region of a geocast is reached. Inside the destination region, regional flooding may still be used even for protocols characterized as non-flooding.

This article is structured as follows. In the next section we discuss early work and background in introducing geographic information to networks. We describe geocast routing protocols in detail. Simulations to evaluate the most interesting protocols are presented. Comparative discussions of the protocols are provided. Finally we conclude with a brief summary.

An earlier overview of geocast protocols can be found in [2]. Our survey covers all currently known geocast protocols and provides, in contrast to this earlier work, a comparative simulation and comparative protocol discussion.

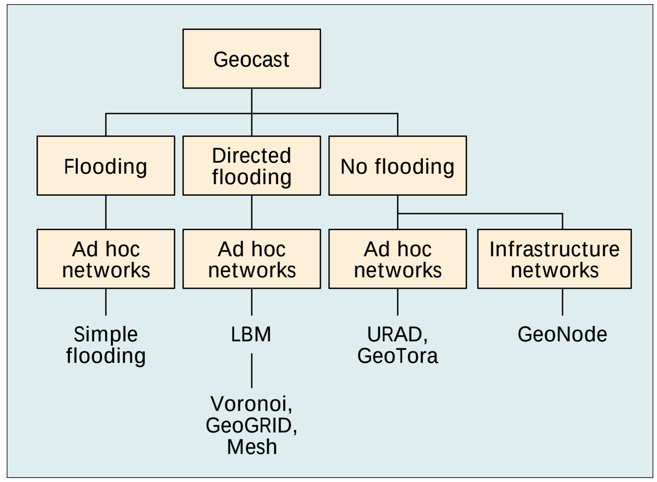


FIGURE 1. *Geocast taxonomy.*

EARLY WORK AND BACKGROUND OF GEOCAST

The first ideas for geocast go back to the attempt to relate IP addresses to geographic locations in the UUMAP project [3]. The project maintained a database in which the geographic locations of Internet hosts were stored. Later, two similar projects [4, 5] tried to relate DNS names to geographic locations. They extended the DNS data structure with geographic longitude and latitude information, which makes it possible to return the geographic location of a host based on its IP address or DNS name. However, both approaches were not able to support the reverse function, that is, they were not able to return the IP address or DNS name based on geographic information. Therefore, such systems made it possible to relate data flows with geographic areas, but they were unsuitable to direct data flow to a given geographic area.

Routing packets to a geographic destination location was first presented in [6] with Cartesian Routing. Cartesian Routing uses latitude-based and longitude-based addresses. Each network node, that is, source node, destination node, or intermediate node, knows its geographic address and the geographic addresses of its directly connected routers. Based on this information, geographic routing is possible where packets are forwarded to the neighboring node that is closer to the destination node than any other neighboring node or the forwarding node itself. If no neighboring node is closer to the destination than the forwarding node, the search space is enlarged by considering all nodes with n-hops distance to the forwarding node, using a flooding mechanism. Note that this approach is the basic algorithm for several later protocols. A restriction of Cartesian Routing is that only unicast is considered. An overview of position-based routing protocols can be found in [1].

GEOCAST ROUTING PROTOCOLS

In this section we present an overview of geocast routing protocols. All protocols have in common that they enable transmission of a packet to all nodes within a geographic region. In contrast to multicast, which enables a packet to be sent to an arbitrary group of nodes, for example, to all nodes that wish to subscribe to a news channel, a geocast group is only defined by a geographic region. Note that geocast is a subclass of multicast and can be implemented with a multicast service by simply defining the multicast group to be a certain geographic region as described later with the GeoNode approach. However, this leads in most cases to non-optimal protocols, especially in ad hoc networks, where geographic information can be used to make routing more efficient.

On the other hand, if a geographic region addressed by a geocast protocol also contains nodes that should not receive a geocast message, it is possible to refine a geocast region by a multicast group. For example, if only cars inside a geocast region rather than pedestrians should be addressed, a multicast group could refine the geocast region. However, the geocast routing protocol would still be the basic mechanism to deliver messages from a sender to a geographic region. Inside the addressed geocast region multicast filtering or a multicast protocol could then determine the final delivery.

For our descriptions the following two definitions are useful. We refer to the *destination region* of a geocast packet as the geographic area to which a packet has to be delivered. A *neighbor* is a node that can be reached without the help of intermediate nodes, that is, it is within the wireless transmission range of a node.

For most approaches we inherently assume that all nodes are aware of their own position. This can be achieved, for example, by using the Global Positioning System (GPS) or any other positioning technology.

ROUTING WITH FLOODING

Simple Flooding

Overview — Simple Flooding floods the whole network with a geocast packet irrespective of the geocast destination region. All receivers have to check whether they are within the destination area.

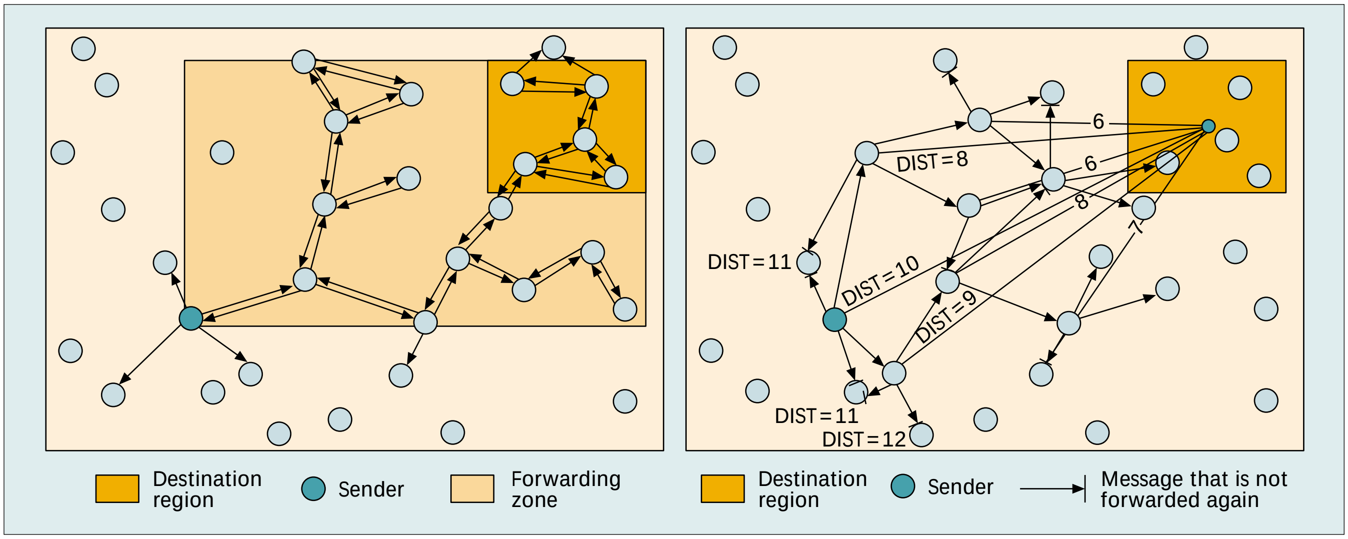
Description — Simple Flooding was not proposed as a geocast routing protocol, but it is useful for comparison with other geocast protocols and it is a building block for many of them.

A Simple Flooding geocast algorithm works as follows. A node broadcasts a received packet to all neighbors, provided that this packet was not already received before in order to avoid loops and endless flooding. A node delivers a packet if the own location is within the specified destination region, which is included in each geocast packet. This is a simple and robust but not efficient approach, since location information is not used for forwarding in order to reduce the number of packets.

ROUTING WITH DIRECTED FLOODING

LBM

Overview — Location Based Multicast (LBM) is based on flooding but avoids flooding the whole network by defining a forwarding zone. Outside the forwarding zone the packet is discarded.



■ **FIGURE 2.** *Geocast example with LBM: a) rectangular forwarding zone; b) distance–based forwarding zone.*

Description — A recent research area is geocast in ad hoc networks, that is, spontaneous constituting networks without a fixed infrastructure. In wireless ad hoc environments, Ko and Vaidya [7, 8] identified two approaches: modified multicast flooding and the modified multicast tree-based approach. For efficiency reasons, multicasting in traditional networks is mainly based on tree-based approaches. However, as tree-based approaches require frequent reconfigurations in ad hoc network environments, they are considered unsuitable by the authors to solve the geocast problem in ad hoc networks. Therefore, two schemes that improve multicast flooding with position information are presented, which are both derived from Location Aided Routing (LAR) [9], a protocol for unicast routing in ad hoc networks.

Simple Flooding as described above is modified by defining a forwarding zone that includes at least the destination region and a path between the sender and the destination region. An intermediate node forwards a packet only if it belongs to the forwarding zone. By increasing the forwarding zone, the probability for reception of a geocast packet at all destination nodes can be increased; however, overhead is also increased. Similar to their unicast routing protocol, the forwarding zone can be the smallest rectangular shape that includes the sender and the destination region possibly increased by a parameter δ to increase the probability for message reception. The forwarding zone is included in each geocast packet in order to allow each node to determine whether it belongs to the forwarding zone.

A second scheme defines the forwarding zone by the coordinates of the sender, the destination region, and the distance of a node to the center of the destination region. A node receiving a geocast packet determines whether it belongs to the forwarding zone by calculating its own geographic distance to the center of the destination region. If its distance decreased by δ is not larger than the distance stored in the geocast packet, which is initially the sender distance, the geocast packet is forwarded to all neighbors and the packet sender's distance is replaced by the calculated own distance. In other words, a node forwards a packet if it is not farther away from the destination region than the one-hop predecessor of the packet increased by δ. Finally, a geocast packet is forwarded to all neighbors if the one-hop predecessor is located inside the destination region.

Example — Figure 2 shows an example of LBM packet delivery with both schemes. With the first scheme (Fig. 2a) a rectangular forwarding zone is defined. Each node inside the forwarding zone forwards the packet by means of broadcast, that is, the packet may reach nodes outside the forwarding zone. However, these nodes drop the packet instead of rebroadcasting it.

The second scheme is shown in Fig. 2b. The initial sender has the distance 10 to the center of the destination region. The distance is included in the geocast packet. Nodes receiving the initial broadcast calculate their own distance to the destination region. If their distance is not larger than the distance included in the received packet (we assume δ = 0) the packet is rebroadcast. Otherwise, it is discarded.

Voronoi Diagrams

Overview — Voronoi-diagrams-based routing improves the LBM approach, which fails if the forwarding zone is empty or partitioned. A new definition of the forwarding zone is given which overcomes these problems.

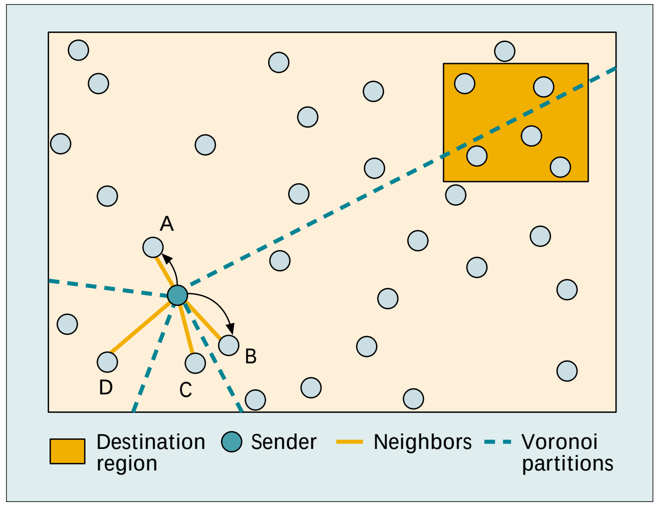
Description — The new definition of the forwarding zone is as follows [10]. A neighbor of a sender belongs to the forwarding zone if and only if it is closest in the direction of the destination. As the destination is not defined by a single position but by an area, all possible positions of destinations inside the geocast region are considered. This leads to having several neighbors belonging to the forwarding zone. Note that with this definition of a dynamic forwarding zone, which, in contrast to LBM, takes the current neighbor position into account, an empty forwarding zone is avoided. Another example of a protocol that uses a dynamic forwarding zone and that focuses mainly on avoiding unpredictable obstacles in it is given in [11]. In [12] a dynamic forwarding zone is defined in the context of inter-vehicle communication that takes vehicle velocities and driving direction into account.

The neighbors belonging to the forwarding zone can be determined using the concept of voronoi diagrams. A voronoi diagram partitions the network in *n* voronoi regions, where *n* is the number of neighbors. Each neighbor is associated with one voronoi region. The voronoi region of a neighbor consists of all nodes that are closer to this neighbor than to any other neighbor.

If a node holds a geocast packet, it starts with determining the voronoi diagram. The voronoi partitions intersecting with the geocast destination region belong to the forwarding zone and are selected for geocast forwarding. Inside the destination region, flooding can be used. In fact, any protocol can be used that can be independent of the protocol used outside the destination region.

We note that the major advantage of voronoi-diagrams-based routing is that empty forwarding zones are avoided. However, flooding overhead is still high and additional computation overhead is introduced by determining the voronoi partitions.

Example — Figure 3 shows an example of how voronoi diagrams are used to define the forwarding zone. The set of neighbors is shown with solid lines connected to the sender. In this example it is obvious that the rectangular forwarding zone of LBM would fail, since no neighbor of the sender is inside it. With voronoi diagrams, two partitions belong to the forwarding zone, that is, the partitions associated with neighbor *A* and *B*, since both intersect the destination region.



■ **FIGURE 3.** *Geocast example with voronoi diagrams.*

Mesh

Overview — Mesh uses directed flooding to discover redundant routing paths to the destination region. The actual payload geocast packet is sent on the discovered paths, called mesh, without network-wide flooding.

Description — Boleng *et al.* [13, 14] propose a simple geocast algorithm based on the mesh approach. The mesh provides redundant paths between the source and the destination region, in order to provide robustness against host mobility and link failures. Three approaches are proposed to create the mesh. They are based on simple flooding and on the forwarding zone approach of LBM [7].

One major difference is that this initial step is only used to create the mesh rather than for sending the actual geocast payload. After a node inside the destination region received the initial packet to join the mesh, a unicast reply is sent back to the sender on the reverse path and flooding is stopped. This requires that state information is maintained on each intermediate node or that the route is recorded in the flooded packet. Nodes that forward the unicast reply become part of the mesh. The mesh consists of all edge nodes of the geocast region and their paths to the sender. The mesh is henceforth used to deliver geocast packets to the edge nodes of the destination region by flooding the mesh. Inside the destination region Simple Flooding is applied.

Therefore, Mesh is a combination of directed flooding and establishing a routing path. Since directed flooding is mandatory to create the mesh and after creation periodically to adapt to network topology changes, we classify Mesh as a directed flooding protocol.

GeoGRID

Overview — GeoGRID partitions the network into logical grids, with a single elected gateway in each partition. Only gateways forward packets, which relieves other nodes from inefficient flooding.

Description — Liao *et al.* presented with GeoGRID [15] another geocast protocol for ad hoc networks. With GeoGRID, in contrast to the approach of [7], they avoided flooding of all nodes since they considered flooding as a costly operation, which is confirmed by [16]. Note that with flooding it is likely that a geocast packet will be received several times from several neighbors and that collisions occur frequently. Liao *et al.* also considered tree-based multicast approaches as unsuitable due to the high uncertainty of host mobility in an ad hoc network.

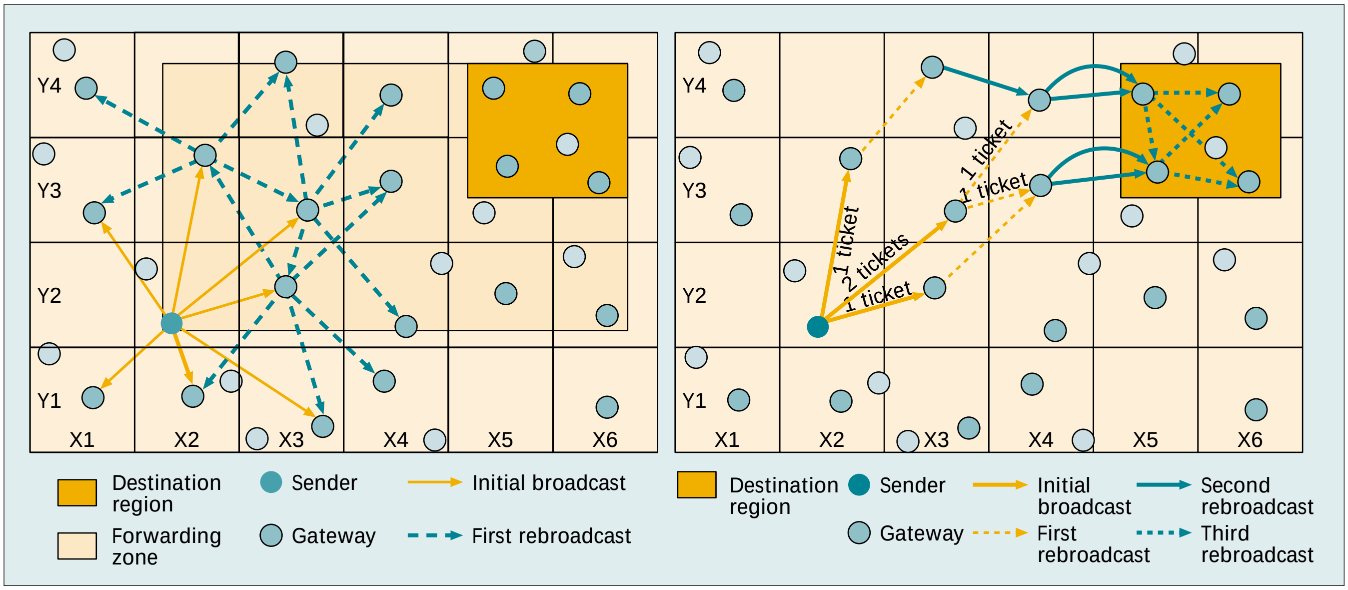
The basic idea of GRID [17], the unicast predecessor of GeoGRID, is to partition the network into logical grids. In each grid one host close to the grid center is elected to be the responsible node for propagating geocast packets to neighboring grids. These hosts are called gateways. Geocast packets are sent in a grid-by-grid manner through their gateways. Thus, gateways are responsible for forwarding geocast packets to neighboring grids, which decreases message overhead by relieving non-gateways from packet flooding. The assumption is that grid sizes are constructed such that a gateway is able to communicate with at least one or more other gateways without relaying.

Prior to sending a geocast packet no geocast-specific routes are established. When forwarding a geocast packet, the rectangular forwarding region introduced in [7] is used, in order to define a forwarding direction and to decrease message overhead. Outside the forwarding region a received packet is discarded. Otherwise, if a gateway inside the forwarding region receives a packet, it rebroadcasts the packet to its neighbor gateways provided that the packet is not a duplicate of a packet that was already broadcast. Inside the destination region a mobile node delivers a received packet to its upper application layers.

Besides this flooding-based GeoGRID the authors proposed ticket-based Geo-GRID. In this second scheme again a gateway within the forwarding region forwards geocast packets, but only a limited number of gateways will do this. To limit the number of gateways, a gateway forwarding a packet sends it to at most three neighbors rather than to every neighbor, and the initial sender limits the overall fan-out of the flooding by specifying a number of tickets. The idea is that each ticket is responsible for carrying one copy of the geocast packet to the destination region. Thus, by selecting a certain number of tickets the initial sender not only determines the overhead of geocast delivery but also the success probability of delivery.

If a gateway is not within the destination region, it will select up to three neighboring gateways whose grids are closer to the destination region and within the forwarding region. The geocast packet is then forwarded to the selected gateways and the tickets are evenly shared among them. If only one ticket is left, a packet is always forwarded to exactly one neighbor. Note that even if a gateway receives a duplicate message, that is, tickets from two different neighboring gateways, it will not discard the duplicate geocast message. Instead, the gateway will forward the duplicate message as described above since each ticket is responsible for carrying one copy of the geocast message to the destination region.

If a gateway receiving a geocast packet is within the destination region, it will rebroadcast the packet within that region to achieve a high arrival rate.



■ **FIGURE 4.** *Geocast example with GeoGRID: a) flooding–based; b) ticket–based.*

Example — Figure 4 shows an example of flooding-based and ticket-based GeoGRID. With flooding-based GeoGRID (Fig. 4a) the initial broadcast from the sender is sent to all neighbors. Neighbors outside the forwarding zone discard the packet, so that a geocast packet floods only the forwarding zone and its surrounding nodes that are within direct communication distance to a node inside the forwarding zone. Due to the large number of messages only the first two broadcast steps are shown.

Figure 4b shows ticket-based GeoGRID. Initially, the sender has four tickets, which are shared between its neighbors. The number of tickets is equal to the number of packets arriving in the destination region. Inside the destination region simple flooding is performed.

ROUTING WITHOUT FLOODING

URAD

Overview — Unicast Routing with Area Delivery (URAD) is a simple geocast routing protocol that uses a regular unicast routing protocol between the sender and the destination region. Inside the destination region, flooding can be used, as well as any other routing protocol that can be independent of the protocol used outside the destination region. For our classification it is important that no flooding is used outside the destination region.

Description — URAD cannot be found in the literature as a geocast proposal. It is a name we have chosen to identify a protocol class that is briefly described in [18]. This protocol encompass two phases: the unicast forwarding from the initial sender until the first node inside the destination region is reached, and the flooding inside the destination region. Unicast forwarding can be realized with any available unicast protocol. For example, in [18] a position-based greedy routing algorithm that was derived from GPSR [9] is used. The unicast destination is selected to be the center of the geocast destination region.

Each node on the unicast forwarding path checks whether it is already part of the destination region. If it is not, then unicast forwarding is continued. If it is part of the destination region, unicast forwarding is stopped and regional flooding inside the destination region is started. Each node inside the destination region sends a received geocast packet with a 1-hop broadcast to all neighbors. Sequence numbers of flooded geocast packets are stored to prevent a node from flooding the same packet more than once, which allows proper termination of the algorithm.

GeoNode

Overview — GeoNode requires an infrastructure network. Geocast routing is done with either the usual unicast, multicast, or with hierarchical GeoRouters.

Description — Imielinski and Navas [20–22] were the first to consider the problem of geographic multi-point to multi-point routing. In [20] they identified three solutions to integrate geographic addresses into the Internet design, which uses logical addresses so far:

- Unicast IP routing extended to deal with GPS addresses.
- GPS-Multicast.
- Application-layer solution using extended DNS.

Their assumption is that the network has a cellular architecture with a GeoNode (or mobile support station (MSS)) assigned to each cell, resulting in two-level routing, the first between sender and MSS and the second between MSS and destination region. They use GPS coordinates based on latitude and longitude information.

The DNS approach extends DNS servers and DNS entries with geographic information. A new first-level domain ".geo" is used for this purpose. Second-level domains represent states, third-level domains represent counties, and fourth-level domains represent polygons of geographic coordinates. In contrast to current DNS, a geographic address is now resolved to a set of IP addresses of GeoNodes covering the whole destination area. The packet is now sent with unicast to all resolved IP addresses of the GeoNodes or by multicast after all resolved GeoNodes are asked to temporarily join a multicast group for this purpose. Routing between GeoNode and destination nodes inside a cell is described later.

With the GPS-Multicast solution, geographic accuracy is limited by address space (especially with IPv4), that is, there is not enough addressing space available to address any and all arbitrarily small geographic areas. Instead, they introduced smallest addressable units, called atoms. Each atom and partition (a larger area of several atoms or partitions) is mapped to a multicast address, which is used for the first level of routing from the sender to the GeoNode. Each GeoNode joins all multicast groups for atoms and partitions that intersect its range. The sender determines the multicast address of the smallest partition that covers the original destination polygon to which he wants to send his message. He uses this multicast address as the address of the packet and puts the original polygon specification into the packet content. Later, the exact matching is done using the polygon specification in level two between GeoNode and the destination.

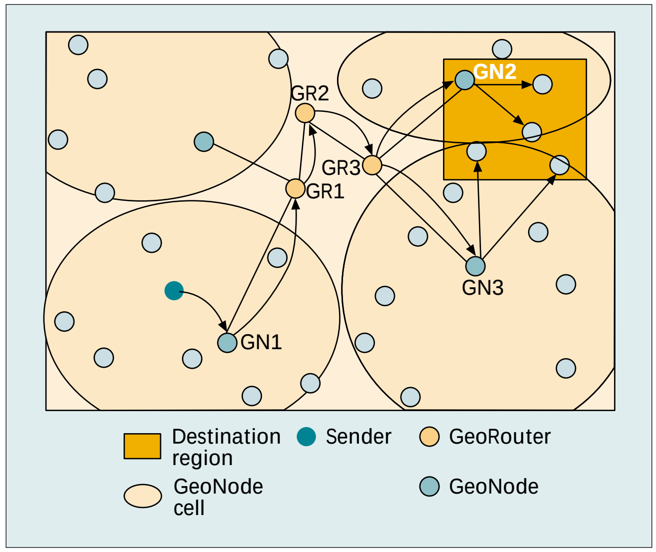
The last approach, integrating geographic addresses into routing decisions, is discussed in detail in [22, 23]. Three components are necessary for geographic routing: GeoHosts, GeoNodes, and GeoRouters. GeoRouters are in charge of moving a geographic packet from a sender to a receiver. They know their service area and exchange this information with other routers. To improve efficiency, they are arranged in a hierarchy with small service areas at the leaf nodes and merged service areas at non-leaf nodes.[1] Note that as an introduction scenario GeoRouters could be established on an overlay network with tunnels between GeoRouters that are not directly connected, similar to the current multicast backbone (MBone). GeoNodes store incoming geographic packets for the duration of their lifetime and periodically multicast them to their cell or service area. GeoHost is a daemon located on all hosts that is capable of receiving and sending geographic packets. It notifies client processes of geographic packets, current location, and the address of the GeoNode.

Routing works as follows. Sending a packet involves the following three steps: sending, shuttling between routers, and receiving. To send a packet, the GeoHost is queried for the GeoNode IP address. The packet is then forwarded to the GeoNode, which in turn forwards the packet to the local GeoRouter. A GeoRouter determines whether the destination polygon (carried inside the packet) and its own service area intersect each other. If they do not, the packet is forwarded to the parent router. If they partly intersect each other, a copy of the packet is sent to the parent router also. In case of any intersection, the GeoRouter checks each child node's service area and sends a copy of the packet if they intersect each other. In the last step GeoRouters deliver a packet to the responsible GeoNodes. Finally, the GeoNodes deliver a packet to all users of the destination area.

This second part of the routing between GeoNode and destination can be done in the same way for all three approaches. It can be based either on application-level filtering or on multicast filtering. With application-level filtering the GeoNode will use a multicast address (or several multicast addresses) to forward the packet, which additionally includes the GPS address. Matching will be performed on the application layer, that is, nodes will individually compare the destination polygon with their own geographic position and discard the packet if they do not match.

With multicast filtering, matching will be performed on the IP layer. The GeoNode sends out a list of all available packets, their geographic destination regions, and their assigned temporary multicast group addresses on a well-known multicast group address. All clients inside a destination region join the temporary multicast group on which the payload packet is later sent by the GeoNode. The address is cached for some time in case several payload packets are sent to the same destination.

Example — Figure 5 depicts the GeoRouter scheme. The sender forwards a geocast packet to its local Geo-Node GN1, which in turn forwards it to its GeoRouter GR1. Since GeoRouter GR1's service area does not intersect with the destination area, the geocast packet is forwarded to the parent GeoRouter GR2. GR2's service area encloses the destination area, hence the packet is forwarded to the child GeoRouter GR3 and from GR3 to the GeoNodes GN2 and GN3, since both cells intersect with the destination area. The local delivery within a cell is done with multicast.



■ **FIGURE 5.** *Geocast example with GeoNode.*

GeoTora

Overview — GeoTORA maintains for each geocast group a directed acyclic graph comprising all network nodes, which shows the routing direction to the destination region.

Description — GeoTORA [24, 25] is another geocast protocol for ad hoc networks. It is based on TORA (Temporally Ordered Routing Algorithm) [26, 27], which is a unicast routing algorithm for ad hoc networks. In TORA a directed acyclic graph (DAG) is maintained for each destination. The DAG shows for each node the direction to the destination node, hence it can be used for forwarding a packet to a destination starting at any node.

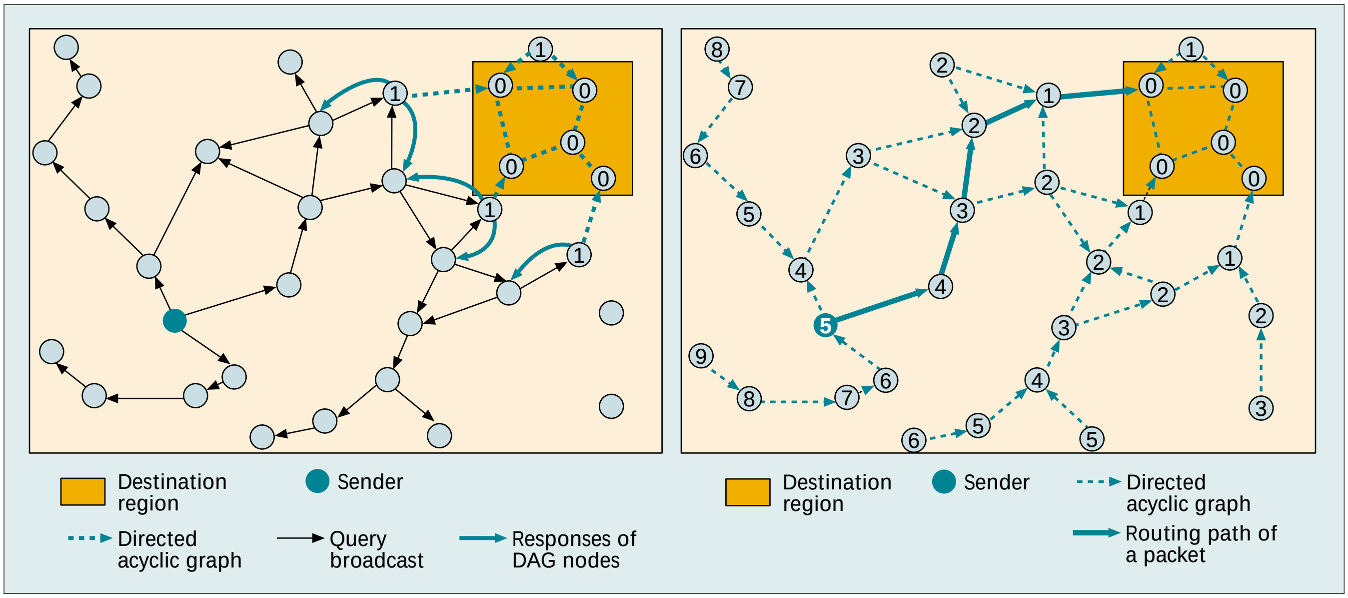
The GeoTORA algorithm is based on an anycast modification of TORA. First a DAG is maintained for each anycast group. Between members of the anycast group there is no direction in the DAG, that is, they are all possible destinations. The directions within the DAG are defined by assigning a height to each node. A packet is always forwarded to a neighbor with lower height. Basically, the height is the distance to the destination region.[2] Members of the geocast group are assigned height 0.

The initial DAG is created as follows. When a node first requires a route to a geocast group it broadcasts a query to all neighbors. The query is rebroadcast until a member of the DAG is found (the neighbor nodes of the destination region are already members). A flag on each node helps to identify duplicates, which are discarded. On receiving a query, a member of the DAG responds by broadcasting its height to its neighbors. A node that waits for a connection to the DAG sets its own height to the minimal height of all neighbors increased by one and broadcasts its height.

Since nodes are moving, the DAG is not stable. However,

[1] *Navas also proposed non-hierarchical approaches integrated in usual unicast routing protocols. They are not discussed in this article.*

[2] *Note that we give a simplified description of the exact GeoTORA protocol. A node's height is more than its distance from the destination region, which ensures that the height is unique and hence that a direction can always be derived from the height of two nodes. However, our description is accurate enough to provide a basic understanding.*



■ **FIGURE 6.** *Geocast example with GeoTORA: a) creating the DAG; b) forwarding of a geocast packet.*

maintaining the DAG is achieved without flooding, therefore we classify GeoTORA as a routing protocol without flooding. GeoTORA reacts to changes in the DAG if a node no longer has outgoing links. Then the direction of one or more links is changed, which is called link reversal. Neighbor nodes are only affected by this measure if their last outgoing link has changed to an ingoing link, which means that they have to repeat the link reversal process.

If the directed links of the DAG are followed to forward an anycast packet, it is finally delivered to a random node of the anycast group. Geocasting with this algorithm works as follows. It starts with an anycast to a random member of the geocast group using the approach described above. Upon receiving the first geocast packet the geocast group member floods the packet within the geocast region.

Example — An example of the GeoTORA protocol is shown in Fig. 6. The sender starts the DAG creation process by broadcasting a query, which is rebroadcast until the initial DAG is reached. Responses containing the node's height are replied (in Fig. 6a the first replies are shown) until the whole network is arranged in a DAG (Fig. 6b). Now a geocast packet is forwarded by following the DAG. Inside the destination region, flooding is performed on all geocast group members. Flooding stops on nodes outside the destination region or when a duplicate is received.

PROTOCOL SIMULATION

We performed simulations to compare some of the protocols described above. They include Simple Flooding, Unicast Routing with Area Delivery (URAD), LBM with a rectangular forwarding zone and, as an alternative, with a cone-based forwarding zone. The cone includes the sender as the peak of the cone and the destination region as the opposite end. With this selection of protocols we have one protocol from each class (flooding, directed flooding, no flooding).

SIMULATION SCENARIO

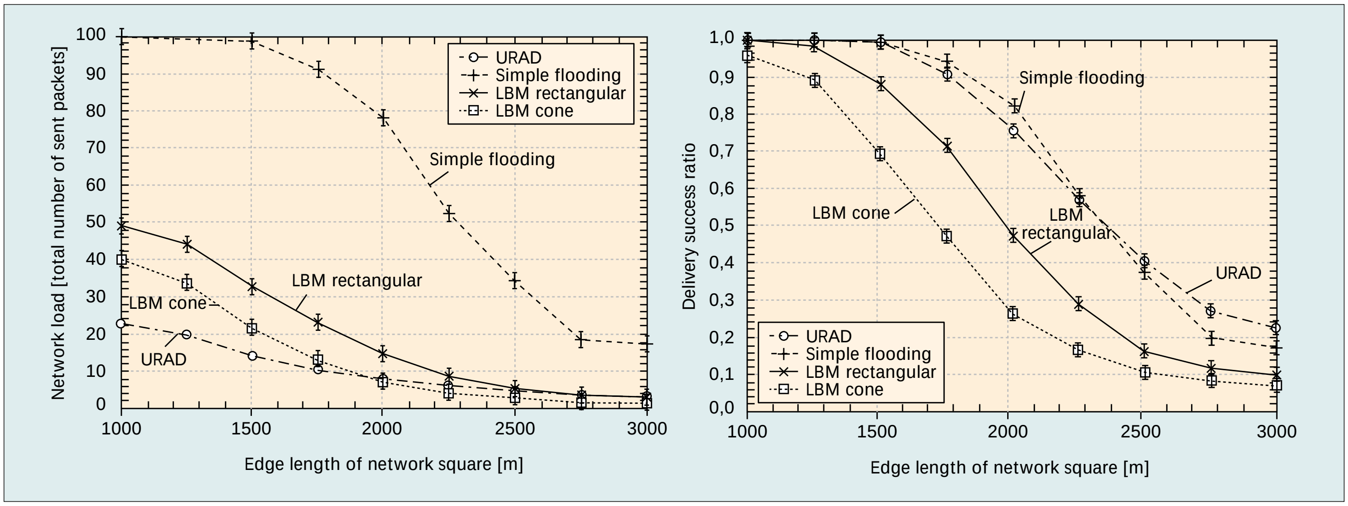
Our simulations are based on the network simulator NS-2 [28] with the CMU wireless extensions. This allows us to run simulations in a quite realistic scenario. The simulated IEEE 802.1 network was configured to have a 250m wireless transmission range and consists of 100 to 1000 nodes. It uses the RTS/CTS scheme preceding every unicast data packet exchange and ACKs to confirm successful unicast packet reception. For broadcasting no such scheme was used. Node movements follow the random waypoint model. Node velocities are up to 50m/s, which includes vehicular movements.

The load was induced with geocasts from random senders to circular destination regions with random centers. The diameters of the circular destination regions were randomly selected between 200m and 300m. The simulation time was 15s. The displayed results are the average of 75 simulation runs. The 95 percent confidence intervals are shown in all figures.

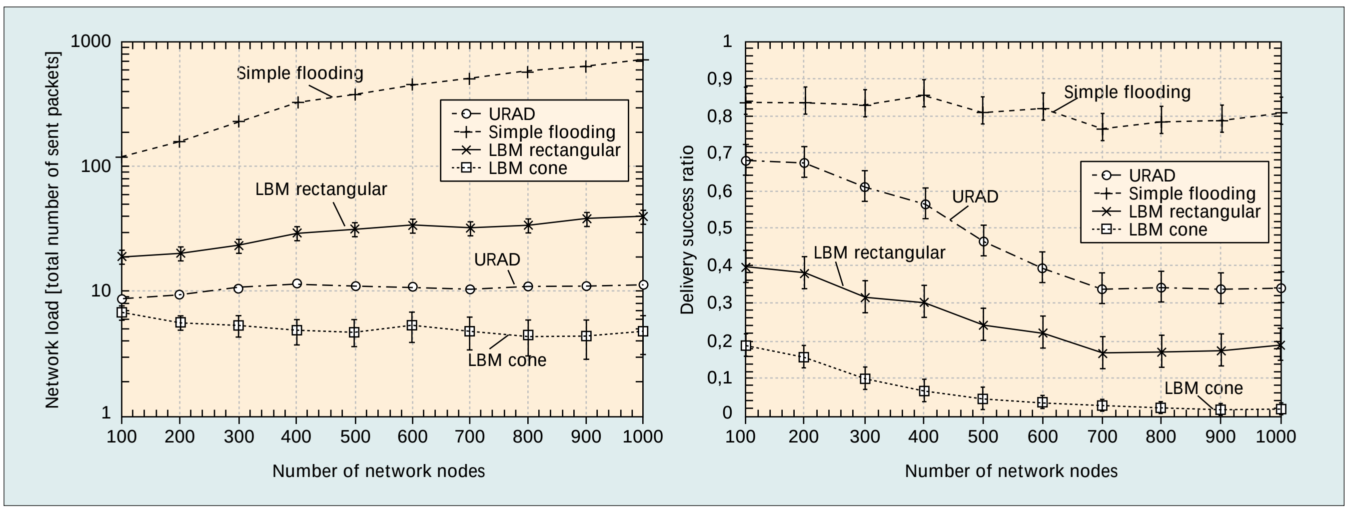
SIMULATION RESULTS

Figure 7 shows the simulation results for 100 nodes with varying network density. If the network is sparse, that is, if each node has only a few neighbors, the network load is decreased. Figure 7a shows that in particular the simple flooding approach benefits from less dense networks, which significantly reduces the network load. However, comparing the absolute results we see that the flooding approach overwhelms the network with the highest number of sent packets. Both directional flooding approaches, LBM with rectangular zone and LBM with cone forwarding zone, are able to reduce the network load and are superior to simple flooding. The URAD approach also results in a low message overhead since its unicast-based greedy forwarding scheme uses only a single path to the destination region.

In Fig. 7b we compare the delivery success ratios of the four approaches. A success ratio of 1 means that every node that is inside a geocast destination region has received the corresponding geocast packet. Obviously, the simple flooding approach, which has the most redundancy, achieves a high delivery success ratio. Note that the delivery success ratio may be worse for highly congested networks, especially for the simple flooding approach, since frequent broadcast collisions may occur. Surprisingly, the URAD approach, which has the least redundancy, also achieves a high delivery ratio. URAD benefits from unicast forwarding, which uses a RTS/CTS scheme preceding every unicast data packet exchange and ACKs with automatic retransmission in case of packet loss. It seems that



■ **FIGURE 7.** *Simulation results with varying network density, i.e., varying edge length of the network square and fixed number of nodes: a) total network load; b) delivery success ratio.*



■ **FIGURE 8.** *Simulation results with varying network size, i.e., varying number of network nodes and constant mean node density: a) total network load; b) delivery success ratio.*

this is more important than having multiple forwarding paths as in the LBM approach. In particular the cone-based LBM approach suffers from a low delivery success ratio, since its forwarding zone is too small to benefit from the possible redundancy.

Figure 8 shows simulation results for a varying number of network nodes. The network density was constant with an average of five neighbors per node, that is, the network area was increased with the number of nodes. For 100 nodes, the network size was $2000 \times 2000m^2$; for 100 nodes it was $6200 \times 6200m^2$.

The results in Fig. 8a show that simple flooding does not scale well with the number of network nodes. With more network nodes the overhead grows significantly (note the logarithmic scaling of the y-axis). The message complexity is indeed $O(n)$, where n is the number of network nodes, since every node in the network has to rebroadcast a geocast message. LBM rectangular and also URAD show increasing network load, but less severe than in simple flooding. LBM cone shows a decreasing network load. However, this is only caused by the low delivery success probability of that approach (Fig. 8b), that is, there is simply a high probability that geocast messages get lost. A closer look at the delivery success ratios shows that simple flooding is superior to the other approaches (assumed uncongested networks with only a few broadcast collisions on layer two). With longer routing paths, the probability for message loss increases if an approach has less redundancy than used by the flooding approach. However, we know from research in [16] that flooding can cause a high number of broadcast collisions which limits robustness. Thus, the results obtained here are only valid for uncongested networks.

DISCUSSION AND OPEN ISSUES

According to our taxonomy in Fig. 1 the protocols mainly differ in whether they are based on flooding, directed flooding, or on routing without flooding, and whether they are suitable for ad hoc networks or for infrastructure networks. Table 1 presents a comparison of all discussed protocols. Scalability is distinguished between sending a geocast packet only once or sending it several times, since protocols maintaining a routing path, for example, Mesh or GeoTORA, have a significantly reduced message complexity for the second geocast packet to a given destination region. However, note that Mesh requires a new flooding-based recreation of the routing paths if the topology changes. Since there is no signaling method for

Criterion	Flooding	LBM	Voronoi	Mesh	GeoGRID	URAD	GeoNode	GeoTORA
Fix network/ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Fix	Ad hoc
Path strategy	Flooding	Directed flooding	Directed flooding	Multipath routing	Directed flooding	Unicast	Multicast[9, 11], unicast[10, 11]	Unicast
Scalability/send once	Low	Medium	Medium	Low	Low-high[4]	Medium–high[12]	Medium[9, 11]–high[10]	Low
Scalability/send several times	Low	Medium	Medium	Medium–high	Low-high[4]	Medium–high[12]	High	High
Message complexity/ first time	O(n)	O(n)	O(n)	O(n)	O(n)[7], $O(\sqrt{n})$[8] +gw election	$O(\sqrt{n})$[12]	$O(\sqrt{n})$ + routing protocol	O(2n)
Message complexity/ second time	O(n)	O(n)	O(n)	$O(\sqrt{n})$ – O(n)[5]	O(n)[7], $O(\sqrt{n})$[8] +gw election	$O(\sqrt{n})$[12]	$O(\sqrt{n})$ + routing protocol	$O(\sqrt{n})$[3]
Memory requirements	No O(n)[1]	Low O(n)[1]	Low O(n)[1]	Medium O(ng)[6]	Low O(n)[7, 1], O(n)[8, 2]	Low O(n)[1, 12]	Low-medium O(ng)[9], O(n)[10], O(g)[11]	Medium O(ng)
Robustness	Medium–high[13]	Medium–high[13]	Medium–high[13]	Medium–high[13]	Medium	Medium	Medium	Medium
Cope with partial partitions	Yes	Limited	Yes	Limited	Limited	Limited[12]	Yes	Yes
Guaranteed delivery	No	No	No	No	No	No	No	No
Time stable	No	No	No	No	No	No	Yes	No
Multicast group refinements	No	No	No	No	No	No	Yes	No
Rely on other protocol	No	No	No	Yes	No	Yes	Yes	No

n = number of network nodes; g = number of geocast groups; j = number of joined geocast groups.
[1] Store last packets to detect duplicates. [2] Store neighbor information. [3] DAG maintenance not considered. [4] Depends on node mobility. [5] $O(\sqrt{n})$ assuming a two-dimensional regular distribution of nodes and no topology changes; worst case O(n). [6] O(ng) if state information is maintained on intermediate nodes, $O(nj\sqrt{n})$ if source routing is used. [7] Flooding-based. [8] Ticket-based. [9] GPS-Multicast. [10] GeoRouter. [11] DNS. [12] Depends on unicast routing protocol. [13] Depends on network congestion and other parameters (see [16]).

■ **Table 1.** *Comparison of geocast protocols.*

topology changes, a sender has to periodically recreate the mesh. As a result, Mesh is more efficient only if the geocast message frequency is higher than the frequency of topology changes. While maintaining a routing path can help reduce the message overhead, usually these protocols have higher memory requirements, since they have to maintain routing paths for every known geocast group. GeoGRID's scalability depends on the node mobility. With high mobility, frequent gateway elections decrease scalability.

The next table columns show the overall message complexity and memory requirement. O(n) is the message complexity for flooding, since each node has to rebroadcast a packet. GeoTORA also has a linear message complexity. However, we denote it as O(2n) since the network is flooded twice. The message complexity is $O(\sqrt{n})$ if a particular or several particular routing paths are followed. The memory requirements depend on whether each node has to store some recently delivered packets or neighbor information (O(n)), or geocast group information (O(ng)). The DNS approach of GeoNode requires that geocast group information be stored in the central DNS database rather than on every node, which results in memory requirements of O(g).

Robustness refers to the ability of a protocol to react to changing conditions and to cope with failures. For a flooding-based protocol, robustness is quite high since no state information is maintained and a single geocast packet is usually delivered several times to the destination region. However, broadcast collisions my decrease robustness [16].

For other protocols robustness is lower. GeoGRID needs to run a gateway election protocol when nodes are moving, which can limit the robustness with high node mobility. GeoNode has only a single routing path to a destination region. Finally, GeoTORA faces problems with possible oscillation and fast route adaptations when maintaining the directed acyclic graph with high node mobility. Note that there is a trade-off between robustness and communication complexity.

To be able to cope with partial network partitions is the next criterion addressed in the table. This refers to a protocol's ability to reach the destination region even if there are partial partitions in the network. This is a severe problem for the directed flooding approaches, since it is possible that they define a forwarding zone which provides no communication connection to the destination region.

None of the protocols provide guaranteed delivery, which is usually a service provided at a higher protocol layer. Only GeoNode is able to provide a time-stable geocast, that is, it periodically redelivers the geocast message for nodes joining a geocast destination region later. Among the discussed protocols only GeoNode provides, in addition to geocast, a refinement of the addressed destination nodes with a multicast function. Such a refinement or intersection of a geocast and multicast group is reasonable in many cases, for example, to address only cars in a geocast region rather than pedestrians also. Finally, only URAD, Mesh, and GeoNode rely on other unicast or multicast routing protocols that are not part of their specification.

From the overview table we can derive open issues for geocasting. Although we said that reliability is usually handled on a higher protocol layer, an open issue is to analyze the influence of retransmissions from a higher protocol layer on delivery delay, network load, and delivery success ratio. Even more important is to define the exact semantics of reliable geocast. We already know from other group communication research areas such as reliable multicast that it is not trivial or even impossible to define a semantics suitable for most or all applications. So it is possible this will remain an unsolved question. In [29] reliability is improved by caching of messages that are currently impossible to forward due to network partitioning. As soon as a suitable neighbor is within the transmission range again, forwarding of a cached message is continued. Although this does not strive for strict reliability, it can significantly improve delivery success, especially with high node velocities.

Besides GeoNode no protocol supports time stability, in particular none of the approaches for ad hoc networks. Many services and applications such as position-based advertising, position-based publish-and-subscribe, and in particular safety applications in the automotive domain would benefit from a time-stable geocast. For example, a time-stable geocast fixed to a certain geographic area could warn approaching vehicles about an icy road— not only once but every time a new node enters the area. Similar to reliability, time-stability may be better considered on a higher ISO OSI layer. In [18, 30] several proposals are made for achieving time stability with existing geocast routing protocols.

Finally, we identify an open issue in the addressing concept. In all cases geographic addressing is obviously part of the routing layer. But each approach has its own addressing concept, which means the routing layer is not transparent for higher layers. In the worst case, changing the routing layer will also require adaptations for all applications. A further open issue is that in many cases the addressed geocast region cannot be further refined by multicast groups, and that in some cases the addressing seems to be too coarse. A fine-grained addressing concept is, for example, proposed in [31].

SUMMARY

We have presented a survey of geocast routing protocols. The protocols mainly differ in whether they are based on flooding, directed flooding, or on routing without flooding, and whether they are suitable for ad hoc networks or for infrastructure networks, which is our proposed classification.

We have compared all described protocols. They differ in their message and memory complexity, in their robustness, and in their ability to deliver geocast packets in partially partitioned networks. Simulation results show that there are significant differences in a protocol's ability to successfully deliver the geocasts to their intended destination regions and in the network load induced.

Directions of future research should not only include other approaches for routing and forwarding geocast packets, as well as approaches to refine the addressing concept, to discuss and include reliability and time stability of geocast messages.

REFERENCES

[1] M. Mauve, J. Widmer, and H. Hartenstein, "A Survey on Position-Based Routing in Mobile Ad Hoc Networks," *IEEE Network*, vol. 1, no. 6, Dec. 2001, pp. 30–39.

[2] X. Jiang and T. Camp, "A Review of Geocasting Protocols for a Mobile Ad Hoc Network," *Proc. Grace Hopper Celebration (GHC)*, Vancouver, Canada, Oct. 2002.

[3] UUCP, "UUCP Mapping Project," software available at ftp.uu.net/uumap/, 1985.

[4] C. Farrell *et al.*, "DNS Encoding of Geographical Location," IETF, Network Working Group, Request for Comments, RFC 1712, URL: ftp://ftp.isi.edu/in-notes/rfc1712.txt, 1994.

[5] C. Davis *et al.*, "A Means for Expressing Location Information in the Domain Name System," IETF, Network Working Group, Request for Comments, RFC 1876, URL: ftp://ftp.isi.edu/in-notes/rfc1876.txt, 1996.

[6] G. Finn, "Routing and Addressing Problems in Large Metropolitan-Scale Internetworks," ISI Research Report ISI/RR-87-180, University of Southern California, Mar. 1987.

[7] Y.-B. Ko and N. H. Vaidya, "Geocasting in Mobile Ad Hoc Networks: Location-Based Multicast Algorithms," *Proc. 2nd Wksp. Mobile Comp. Sys. and Applications (WMCSA '99*), New Orleans, USA, Feb. 1999, pp. 101–10.

[8] Y.-B. Ko and N. H. Vaidya, "Flooding-Based Geocasting Protocols for Mobile Ad Hoc Networks," *Mobile Networks and Applications*, vol. 7, no. 6, 2002, pp. 471–80.

[9] Y.-B. Ko and N. H. Vaidya, "Location-Aided Routing (LAR) in Mobile Ad Hoc Networks," *Proc. 4th ACM/IEEE Int'l. Conf. Mobile Comp. and Net. (MobiCom '98)*, Dallas, USA, 1998.

[10] I. Stojmenovic, A. P. Ruhil, and D. K. Lobiyal, "Voronoi Diagram and Convex Hull-Based Geocasting and Routing in Wireless Networks," *Proc. 8th IEEE Symp. Comp. and Commun. ISCC*, Antalya, Tukey, July 2001, pp. 51–56.

[11] C.-Y. Chang, C.-T. Chang, and S.-C. Tu, "Obstacle-Free Geocasting Protocols for Single/Multi-Destination Short Message Services in Ad Hoc Networks," *Wireless Networks*, vol. 9, no. 2, 2001, pp. 143–155.

[12] A. Bachir and A. Benslimane, "A Multicast Protocol in Ad Hoc Networks Inter-Vehicle Geocast," *The 57th IEEE Semiannual Vehic. Tech. Conf. (VTC2003-Spring)*, Jeju, Korea, Apr. 2002, pp. 2456–60.

[13] J. Boleng, T. Camp, and V. Tolety, "Mesh-Based Geocast Routing Protocols in An Ad Hoc Network," *Proc. 15th Int'l. Parallel and Distributed Processing Symp. Issues in Wireless Networks and Mobile Computing (IPDPS 2001)*, Apr. 2001, pp. 184–93.

[14] T. Camp and Y. Liu, "An Adaptive Mesh-Based Protocol for Geocast Routing," *J. Parallel and Distributed Computing: Special Issue on Routing in Mobile and Wireless Ad Hoc Networks*, vol. 62, no. 2, Feb. 2003, pp. 196–213.

[15] W.-H. Liao *et al.*, "GeoGRID: A Geocasting Protocol for Mobile Ad Hoc Networks Based on GRID," *J. Internet Tech.*, vol. 1, no. 2, Dec. 2000, pp. 23–32.

[16] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network," *Proc. 5th ACM/IEEE Int'l. Conf. Mobile Comp. and Net. (MobiCom '99)*, Aug. 1999.

[17] W.-H. Liao, Y.-C. Tseng, and J.-P. Sheu, "GRID: A Fully Location-Aware Routing Protocol for Mobile Ad Hoc Networks," *Telecommun. Sys.*, vol. 18, no. 1, 2001, pp. 37–60.

[18] C. Maihöfer, W. Franz, and R. Eberhardt, "Stored Geocast," *Proc. Kommunikation in Verteilten Systemen (KiVS)*, Leipzig, Germany, Feb. 2003, Springer Verlag, pp. 257–68.

[19] B. Karp and H. T. Kung, "Greedy Perimeter Stateless Routing for Wireless Networks," *Proc. 6th ACM/IEEE Int'l. Conf. Mobile Comp. and Net. (MobiCom 2000)*, Boston, USA, Aug. 2000, pp. 241–54.

[20] T. Imielinski and J. Navas, "GPS-Based Addressing and Routing," IETF, Network Working Group, Request for Comments, RFC 2009, Nov. 1996.

[21] T. Imielinski and Julio C. Navas, "GPS-Based Geographic Addressing, Routing, and Resource Discovery," *Commun. ACM*, vol. 42, no. 4, Apr. 1999, pp. 86–92.

[22] J. C. Navas, "Geographic Routing in a Datagram Internetwork," Dissertation at Rutgers, The State University of New Jersey, May 2001.
[23] J. C. Navas and T. Imielinski, "Geographic Addressing and Routing," *Proc. 3rd ACM/IEEE Int'l. Conf. Mobile Comp. and Net. (MobiCom '97)*, Budapest, Hungary, Sept. 1997.
[24] Y.-B. Ko and N. H. Vaidya, "GeoTORA: A Protocol for Geocasting in Mobile Ad Hoc Networks," *Proc. 8th Int'l. Conf. Network Protocols (ICNP)*, Osaka, Japan, Nov. 2000, pp. 240–50.
[25] Y.-B. Ko and N. H. Vaidya, "Anycasting-Based Protocol for Geocast Service in Mobile Ad Hoc Networks," *Comp. Net. J.*, vol. 41, no. 6, Apr. 2003, pp. 743–60.
[26] V. D. Park and M. S. Corson, "A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks," *Proc. IEEE INFOCOM 1997*, Kobe, Japan, Apr. 1997, pp. 1405–13.
[27] V. Park and M. Corson, "A Performance Comparison of the Temporally-Ordered Routing Algorithm and Ideal Link-State Routing," *Proc. IEEE Symp. Comp. and Commun.*, Athens, Greece, June 1998.
[28] S. Bajaj *et al.*, "Improving Simulation for Network Research," Technical Report 99-702, University of Southern California, USA, 1999.
[29] C. Maihöfer, R. Eberhardt, and E. Schoch, "CGGC: Cached Greedy Geocast," *2nd Int'l. Conf. Wired/Wireless Internet Commun. (WWIC)*, Frankfurt (Oder), Germany, Feb. 2004, Springer Verlag.
[30] C. Maihöfer and R. Eberhardt, "Time-Stable Geocast for Ad Hoc Networks and its Application with Virtual Warning Signs," Special Issue of *Comp. Commun.*, to appear in 2004.
[31] F. Dürr and K. Rothermel, "On a Location Model for Fine-grained Geocast," *Proc. 5th Int'l. Conf. Ubiquitous Computing (UbiComp 2003)*, A. Dey, A. Schmidt, and J. McCarthy, Eds., Seattle, WA, USA, Oct. 2003; Lecture Notes in Computer Science 2864, Springer, pp. 18–35.

Biography

CHRISTIAN MAIHÖFER (christian.maihoefer@daimlerchrysler.com) received a diploma (Dipl.-Inf.) in computer science in 1997, and a Ph.D. from the University of Stuttgart, Germany in 2002, where he worked in the field of reliable multicast communication. He is currently working in the Daimler Chrysler AG Telematics Research Lab, where he heads the EU funded CarTalk2000 project. His current research interests include car-to-car communication, ad hoc networks, mobile and wireless communication, and geocast.