A Novel Paradigm for Geographic Routing in Ad Hoc Networks: Comparison of Geograms and Geocircuits

Sophia Fotopoulou-Prigipa and A. Bruce McDonald
Northeastern University, Department of Electrical and Computer Engineering

Abstract—**Geographic routing exploits information specific to relative node locations to make routing decisions. This contrasts to traditional techniques that are based on graph-theoretic network models and algorithms. The strategy is well adapted to mobile networks, however, it is subject to increased overhead due to its inherent susceptibility to local minimums. The problem is a consequence of locally greedy forwarding decisions based on partial information. This paper introduces a novel strategy intended to reduce this overhead by caching path information to avoid repeated recovery from the same local minimum. Using "geocircuits" packets associated with a given traffic flow are efficiently detoured, thus preventing costly repeated recovery operations. Simulation results demonstrate significant performance gains relative to conventional "geogram' routing over a wide range of operating conditions. The advantage is shown to be greater as network parameters are increased towards their limiting values, namely, under heavy traffic load and high mobility.**

I. Introduction

Recent years have witnessed phenomenal growth in the deployment and application of wireless communications networks and services. The vast majority of these systems rely on centralized control and a sophisticated mix of wired and wireless infrastructure, wherein, typically only the "last hop" is wireless. Recent world events and increased interest in very large scale sensor networks that can be instantly deployed, support sensor mobility and operate in harsh terrains have motivated the development of more flexible, robust and rapidly deployable alternatives. As such, substantial work has been focused on developing "reconfigurable wireless networks" (RWN) that require little or no fixed infrastructure, namely, wireless ad hoc networks. The potential benefits of ad hoc networking include rapid deployment, greater flexibility and increased cost effectiveness relative to conventional wireless network architectures.

The most salient feature of the ad hoc network paradigm is that the end nodes, or hosts, are themselves responsible for cooperating in order to provide basic network services. Routing is one of the most important services provided by any communications network. Routing in ad hoc networks requires adapting to the dynamically varying topology without over-utilizing scarce resources, namely, energy, bandwidth and processing capacity. The problem remains open and is difficult due to the time-varying and uncertain nature of the network connectivity. However, a more fundamental roadblock to success can be tied to the shortcomings of graph-theoretic approaches. It can be argued that modeling an ad hoc network as a time varying random graph is inherently flawed due to the unique characteristics of the wireless channel. As such, the design of routing algorithms stands to benefit from re-evaluating well-established beliefs including the basic network model.

The class of routing algorithm that comes closest to challenging traditional models are the location-based routing schemes. To a large extent location-based routing abandons the more conventional topological approach to routing by viewing the network in terms of physical locality and relative position. Some location-based schemes utilize location information to narrow the scope of flooding during the route establishment phase of reactive routing protocols. However, the sub-class that represents a true paradigm shift are the geographic or geodesic routing algorithms.

Two underlying services are required to support geographic routing: (1) localization; and; (2) location management. Localization utilizes physical-layer techniques to enable each station to estimate its current position relative to a reference coordinate system. Location management consists of the databases and protocols implemented to discover the location of other nodes. For example, the identity and location of adjacent nodes can be reported in periodic "hello" messages, whereas, the location of a remote destination node generally requires distributed position update and query mechanisms. Based on these two services each node acquires a restricted view of the network based only on a limited set of node locations.

Packet forwarding decisions in geographic routing are made in a fully distributed manner. Greedy forwarding decisions are made at each hop based on localized, partial information. As a result, the technique is susceptible to local minimums, wherein, a node is reached that is 'optimal' with respect to its own neighbors. For example, it may be the closest node to the desired destination. Various approaches have been proposed for recovering from these situations, however, 'void' recovery may result in substantial overhead on the network. Furthermore, proposed techniques involve searches that can extend over large regions and result in significant packet looping, thus, leading to excessive end-to-end delays. Moreover, without closed-loop control the search must be bounded to prevent potential instability. Consequently, in a connected network it may not be possible to guarantee delivery.

This paper introduces a novel strategy intended to reduce the overhead inherent to geographic routing. This is achieved by caching path information to avoid repeated recovery from the same local minimum. Using "geocircuits" packets associated with a given traffic flow are efficiently detoured, thus preventing costly repeated recovery operations. The analysis also includes a void recovery algorithm based on a bounded depth-first intended to balance stability against the probability of success. Simulation results are presented that demonstrate significant performance gains relative to conventional "geogram' routing over a wide range of operating conditions. Geocircuits are shown to reduce end-to-end delay and routing overhead while achieving comparable throughput. Contrary to initial expectations the advantage is shown to be more dominant as network parameters are increased towards their limiting values, namely, under heavy traffic load and high mobility.

The remainder of this paper is organized as follows: Section II. characterizes and compares our model for geograms and geocircuits. The simulation model, experiments and system parameter selection process are described in Section III.. Simulation results and analysis are presented in Section IV., and, finally, conclusions are presented in Section V..

II. Geograms and Geocircuits

Geographic routing belongs to the class of location-based routing schemes. These schemes exploit location information to reduce the overhead of routing in ad hoc networks. This objective is pursued by reducing the scope of searches in the network for a route, in a manner similar to selective paging in cellular networks. For this purpose, all network stations maintain a localized view of the network topology, involving specific location information. Alternatively, location information can be used more directly in lieu of an address for geographic routing. In this case packets may be forwarded through any existing neighbor that meets a given optimization criteria based on the available local information.

As discussed in the Introduction geographic routing can lead to local minimums that require specific algorithms to recover from and continue forward progress towards the destination. Given that realistic traffic scenarios involve strong correlations among packets flowing through portions of a network at a given time it seems evident that the recovery process will be repeated many times over short intervals. Since void recovery requires "extra" hops it increases the effective network load, end-to-end packet latency and the rate of energy dissipation in the network. Consequently, it would be globally and session-wise advantageous to avoid unnecessary repetition of the recovery process.

We observe that once a packet has "discovered" its way around a local minimum it is theoretically possible to "patch" the hole, that is, to effectively create a detour for subsequent packets that would normally follow the same course, including the costly search and recovery process. This can be achieved by setting up temporary virtual circuit for subsequent packets associated with a given flow (or destination). Although virtual circuits are familiar from conventional networks, and indeed the technique of path caching is widely used by topology based ad hoc routing protocols, the application to this particular problem space and geographic routing in general is a novel contribution made in this work. We name the proposed scheme *geocircuit routing* to emphasize the coupling of virtual circuit and geographic routing paradigms. In contrast we refer to conventional geographical routing as *geogram routing* to emphasize that packets are routing independently, like datagrams, yet based on geographic techniques.

A. Geogram Routing

The restricted information maintained at each station in geographical routing concerns explicit geographical location of stations in the vicinity. In general, local topology information is more precise, while any potential knowledge about distant stations is vague. Typically, a station maintains the geographical positions of its direct neighbors.

In geogram routing each packet is routed independently according to the following basic procedures: Upon originating a packet, a source must acquire location information for the destination utilizing an underlying location management scheme as described in the sequel. Once acquired, the location information is included in each geogram (packet) header and use like an address. Each node (including the source node) selects the neighbor to forward a packet based solely on (imprecise) knowledge of the destination node location and the local view of the network it maintains. The most commonly used criteria for making the forwarding decision are direction, distance, and forward progress. The next hop may be selected so as to be the closest to the direction of the line connecting the forwarding node to the destination. Alternatively, the neighbor that minimizes the Euclidean distance to the destination, or the one that maximizes the progress toward it may be considered the locally optimal choice.

B. Recovery from Local Minimums

Locally greedy geographical routing cannot guarantee globally optimal paths. Worse yet, it cannot guarantee xa given forwarding decision is minimally acceptable. The quality of a given forwarding decisions depends strongly upon the network connectivity, mobility patterns, precision of the location estimates and the forwarding criteria. There is currently no systematic or analytical basis for making globally "good" decisions.

For a given topology the best, or perhaps the only route to a destination requires temporary movement against the forwarding criterion. Using only local information there is no known technique for assessing this in advance. As such, local minimums are common even in highly connected networks. Under this scenario the packet is said to be stuck at this node [], which is termed concave node []. Assume that a packet destined to node D is stuck at node X. Then the intersection of the transmission range of X with the area of better potential neighbor locations with respect to D (according to the forwarding criterion) is empty of nodes. When the criterion is distance minimization, this area has been termed a void [] or hole [].

Recovery from a void requires an algorithm that must locate a better node according to the forwarding criteria. A search must be initiated that continues until a better node is found at which time geographic forwarding may continue, or the search fails. Theoretically it is possible to implement a search algorithms that can guarantee that a path will be discovered if the the network is not partitioned. However, an unbounded search can lead to instability if the true path is "difficult" to discover or, indeed, if the network is partitioned. Hence, in general one cannot assume that void recovery can be guaranteed.

Algorithms that have been proposed to recover from local minimums include flooding, depth first search [], breadth first search and greedy perimeter search that adopts the right hand rule [] for traversing faces of a planar subgraph of the graph representing the network.

C. Geocircuit Routing

To achieve the objective of avoiding redundant recovery operations from local minimums we propose the geocircuit routing strategy. The first packet forwarded towards a given destination is a special geogram in that it not only proceeds according to geographic routing criteria, but, it also initiates the caching of path information as it proceeds. When a local minimum is encountered the recovery algorithm is executed prior to path caching. As such, a detour can be established through which future packets not only avoid the search algorithm, but the follow a more direct route that cuts out the unsucessful portions of the search.

To each geocircuit corresponds a unique geocircuit number on each link along the path. Each node maintains a table that holds an entry for each established geocircuit traversing it, named the *GC-number translation table*. Each entry in the table maps an incoming GC to an outgoing GC and desintation node. Subsequent packets associated with a destination for which a geocircuit has already been established are routed according to GC-number mappings. The GC-number corresponding to the next-hop to which a packet is forwarded, is recorded on the packet's header, on the *GC-number field*. To account for the dynamically changing topology of ad hoc networks a time stamp is associated with each GC table entry. This is initialized to the time when the entry was created, and is reset whenever the entry is utilized. When the topology changes and the geocircuit is broken packets will cease to flow on it, hence, it will eventually time out and be flushed from the network.

D. Geograms versus Geocircuits

Geocircuit routing is proposed to detour expensive local optimum recovery procedure. However, geogram routing may achieve better forwarding decisions for individual packets, exploiting more recent information. Hence, geocircuits may become suboptimal over time relative to geograms. Another anticipated shortcoming of geocircuit routing is that they are subject to failure due to mobility and varying channel characteristics. As such, they require measures for path repair. The algorithm is simple and efficient—if the next hop of a geocircuit is not available then re-initiate geocircuit construction from the break point. The downside to the approach is that is can lead to very inefficient paths over time in highly mobile environments.

III. Simulation Model

The main objective of the simulation model developed is to enable an unbiased, systematic analysis of the relative performance of geograms and geocircuits, and to gain deeper insight into the inherent characteristics of geographic routing. This section presents an overview of our simulation environment, the main elements of the simulation model and discussion of the system parameters selected to achieve the objectives of the analysis.

A. Mobility Model

Mobility represents a crucial performance factor in the present analysis. This subsection describes the entity mobility model [] implemented in the simulation. The entity mobility model introduces correlation between the trajectory of a node in successive time intervals, as well as between the mean and the variance of node velocity. Time is divided to consecutive time intervals called *epochs* according to a Poisson stochastic process of rate λ_M. During each epoch, the movement of a node is characterized by a velocity vector. The value of the velocity is constrained within the interval $[0, v_{MAX}]$. Specifically, the velocity value during consecutive epochs is an independent identically distributed random process, following a normal distribution of mean $\bar{v} = v_{MAX}/2$ and standard deviation σ^2_{vel}. Values falling out of the predefined range are bounced back within the range repeatedly until they fall within it. The velocity direction is dependent to that of the preceding epoch, following a Markov stochastic process. Specifically, the direction of movement is normally distributed with mean its value in the previous epoch and standard deviation σ^2_{DIR}. The direction of each initial epoch is selected uniformly within $[0, 2\pi]$. All nodes move independently according to the described model within the network deployment area. If a node hits a boundary of the area it bounces back in.

B. Traffic Model

The network traffic load represents another crucial performance factor. A session based traffic model was implemented to capture the locality of reference correlations in traffic flows. It is assumed that all packets originate at each node in sessions. The source and destination nodes of each session are uniformly selected. Sessions originate according to a Poisson process of rate λ_S, while the duration of each session is an exponential random variable of mean μ_S. Packets have a constant length L and arrive in a constant rate λ_P (CBR) within each session.

C. Comparative Routing Model

Upon originating a packet to a new destination, a node employs the underlying mobility management scheme to obtain the destinations location. Specifically, it queries the location database for the destination's position. If no response is received after a predefined time interval, a new query is issued. A total of three queries are transmitted with increasing timeout intervals. If the destination location is resolved, the session begins and packets are generated at a constant rate. Each node caches the location coordinates of its current session in order to avoid querying for each packet during a session (similar to ARP).

D. Geographical Forwarding

The local information maintained in each node consists of its own position and the the positions of its one-hop neighbors. This information is recorded in the *neighbor table*. Each node conveys this information by broadcasting "hello" messages bearing its node identifier and position. The hello process generates messages according to a Poisson process of rate λ_H. The rate and neighbor table flush timer were determined experimentally to reflect the mean and variance in the acquisition of new neighbors and the loss of existing neighbors under moderate mobility levels.

Additionally, every node that forwards a data packet piggybacks updated location information on each packet. Apart from a neighbor's identifier and location, a neighbor table entry also includes a time stamp, recording the time when this information was obtained. A predefined parameter l_N determines the lifetime of the entries. If a node does not hear from a neighbor for this time interval, it assumes that the corresponding association no longer exists and deletes the entry.

Euclidean distance minimization is employed as the forwarding decision criterion. Distance based schemes are the most widely employed, while they were demonstrated to have a slight advantage in the performance of static networks [].

E. Local Minimum Recovery

A restricted depth first search is proposed and implemented to recover from local minimums when a node realizes that it lies closer to the destination than all its neighbors. The neighboring nodes are ordered according to their distance from the destination, and are visited in this order. No node is visited more than once. A node may know if it has previously seen a packet by recording the packet's identifier.

In order to reduce the overhead of the recovery process and reduce the chance of searching for a non-existent path, a predefined parameter D determines the maximum depth of the search. The value of D is proportional to the some fraction of the maximum network diameter. To further reduce the overhead of this search, we switch back to geographical forwarding when a node is reached that lies closer to the destination than the node where the packet was stuck. For this reason, the node that initiated a local minimum recovery records its geographical position on the data packet.

F. Mobility Management

We assume that each node may obtain its current geographical position by an underlying localization scheme, like GPS or some equivalent.

Studying the location service scheme is out of the scope of this paper, as it consists a major problem in its own right. However, we acknowledge the important role location service plays in the performance of the routing algorithm. Location service both provides service to geographical routing as well as operates using it. For the above reasons, we choose to implement a simplified location service scheme, in order to account for it within the scope of geographical routing, while not aiming to optimize it. We impose a geographical partition of the network deployment area in equally sized regions of edge G. Each region has a fixed component in its center that serves as a location server for a number of network nodes. The network nodes are evenly distributed among the regions. A deterministic many-to-one mapping maps a node's identifier into a specific region, known as its *home region*. Each location server is equipped with a transceiver of the same power with the ones borne by mobile network nodes. Location updates, queries and replies all employ geogram routing. A node initiates an update in a static manner, whenever it enters a different region.

G. Geocircuit Repair

To account for the dynamic nature of ad hoc networks, we allow for geocircuit table entries to be removed if they are no longer in use. A predefined lifetime l_{GC} determines whether a recorded entry may be considered reliable or stale. Entries that haven't been utilized for as long as this lifetime are erased, so that memory space is preserved.

Due to the dynamic topology of ad hoc networks, the lifetime of a geocircuit is restricted and unpredictable. An approach to rebuild a broken geocircuit is examined to anticipate for such an event. The packet that realizes that its subsequent node in the circuit is no longer a neighbor employs geogram routing and at the same time rebuilds the broken part of the geocircuit. Entries in the disconnected part of the old circuit are soon deleted, since they are no longer utilized.

H. Experimental Design

We employ a graphical technique to determine the transient period of our system, and observe the output only when steady state has been reached. Statistical independence of 10 observations is ensured by the method of batch means. 95% confidence intervals are utilized to indicate statistical significance of output mean estimations.

System parameters (Table 1) were selected so as to provide a realistic, unbiased model. We denote the transmission rate as C, the processing delay as d_{proc}, the propagation speed as s, and the mean number of neighbors as d. We select the hello rate λ_H and the neighbor and geocircuit table lifetime l_N and l_{GC} by experimentally estimating the rate of change in the one-hop neighborhood of nodes.

We choose the network deployment area E to be a dependent parameter, in order to preserve a constant network density while varying the number of nodes. Since there are $d+1$ nodes within a πR^2 area, the total number of nodes N in the network are $(d+1)E^2/(\pi R^2)$. Solving for E, we obtain Equation 3. We select the maximum depth D in the local optimum recovery process to be one third of the network diameter. Approximating the diameter as the network area diagonal divided by the transmission range [], we obtain Equation **??**. We select the velocity standard deviation σ_V^2 as a linear function of the maximum velocity v_{MAX}, so that the velocity is with probability $1/2$ within the range $[\bar{v} - v_{max}/8, \bar{v} + v_{max}/8]$.

Table 1: System Parameters.

C	$11\ Mbits/sec$
L	$512\ bytes$
d_{proc}	$0\ sec$
s	$infinite$
R	$250\ m$
G	$800\ m$
d	$6.5\ neighbors$
μ_S	$30\ sec$
λ_P	$31.25\ packets/sec$
λ_M	$0.0333\ epochs/sec$
σ^2_{DIR}	$\pi/4\ rad$
λ_H	$3.8387\ message/sec$
l_N	$12.0982\ sec$
l_{GC}	$12.0982\ sec$

$$E = R \cdot \sqrt{\pi N/(d+1)} \tag{1}$$

$$D = \frac{E \cdot \sqrt{2}}{3R} \tag{2}$$

$$\sigma^2_V = v_{max}/8 \tag{3}$$

IV. Simulation Results and Analysis

In order to quantitatively compare the two routing strategies, we vary the maximum velocity v_{MAX}, total number of nodes N, and traffic session rate λ_S.

A. Average End to End Delay

The effect of network mobility on the average end to end delay is shown in Figure **??**, for a medium network size of $N = 100$ nodes and session rate of $\lambda_S = 0.0150$ sessions per second. We observe that geocircuit routing exhibits a lower delay, and this advantage increases as node mobility increases. This contradicts our intuition that the performance of geocircuits would degrade in highly mobile environments. It is however evident, that even maximal velocities of $v_{MAX} = 40$ m/sec are not high enough to lead to performance loss by shortened geocircuit lifetime. Moreover, as mobility increases, the overhead of the local optimum recovery process increases. From the increase of the advantage of geocircuit routing, we infer that this strategy succeeds in detouring this expensive process.

The effect of network size on the average end to end delay is depicted in Figure 2, for a moderate network mobility and traffic load of $v_{MAX} = 10$ m/sec and $\lambda_S = 0.0150$ sessions/sec respectively. We observe an advantage in geocircuit routing, that becomes more evident in large networks, where the corresponding 95% confidence intervals are totally disjoint. In such networks, the average path length is larger, hence more local minimums are encountered. Furthermore, since the maximum depth D in the local optimum recovery is proportional to the network diameter, the overhead of the recovery process increases as the network size increases.

In a network of moderate mobility and size, the advantage of geocircuits exhibits a slight increase as the session rate increases (Figure 3).

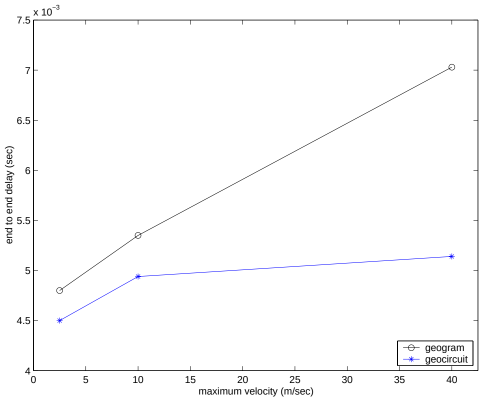

Figure 1: End to End Delay vs. Network Mobility.

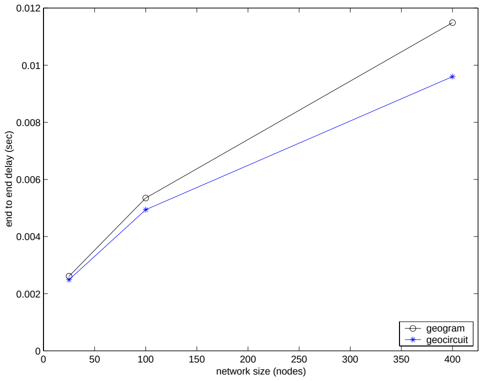

Figure 2: Average End to End Delay vs. Network Size.

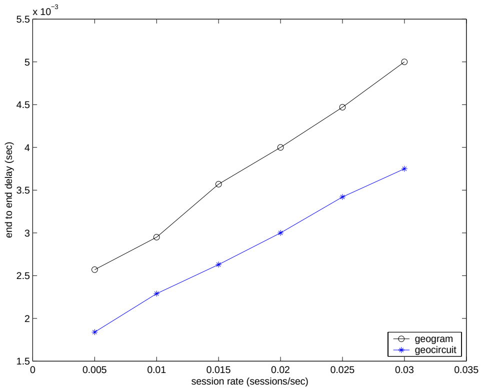

Figure 3: Average End to End Delay vs. Traffic Load.

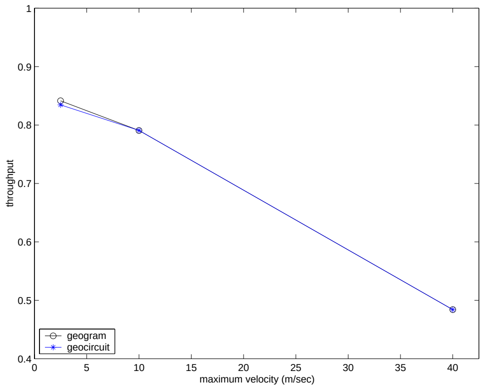

Figure 4: End to End Throughput vs. Network Mobility.

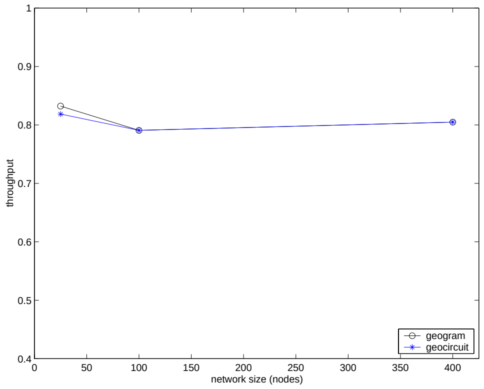

Figure 5: End to End Throughput vs. Network Size.

B. End to End Throughput

We observe that the average end to end throughput of the two routing strategies remains comparable, as limiting network parameters are increased (Figures **??**). Hence geocircuit routing reduces the overhead in terms of average end to end delay, while maintaining comparable routing performance.

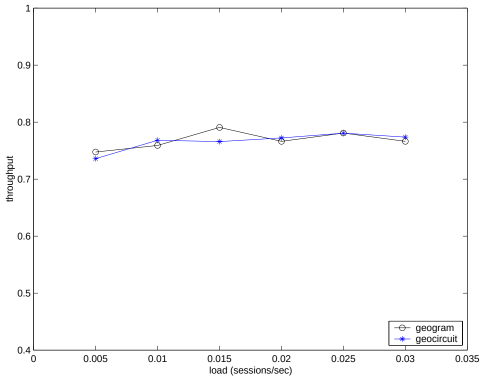

Figure 6: End to End Throughput vs. Traffic Load.

V. Conclusions

The local optimum recovery process of geographical routing induces a substantial overhead in the already scarce in resources ad hoc networks. Since geographical routing is overall a promising approach to reduce the routing overhead, techniques to mitigate such negative effects are necessary. We have applied a strategy borrowed from conventional wired networks, introducing a new scheme addressed as geogram routing. In a realistic traffic scenario where packets are issued in bursts, we exploit already discovered paths to route succeeding packets. We introduce advances in the local optimum recovery scheme and our mobility model, and devise techniques to address network dynamicity. We compare the proposed scheme with conventional geographical routing, addressed as geogram routing. We have demonstrated through simulations that geocircuit routing may reduce the end to end delay up to 30%, while maintaining comparable performance in terms of throughput. Investigation of adaptive schemes, aiming to combine advantages of both presented strategies, as well as verification of our results in different mobility patterns, location service schemes, or imprecise location information, consist interesting future research directions.

References