

Power-Efficient Geographic Routing for MANETs*

LAN LUAN, WEN-JING HSU AND RUI ZHANG

*School of Computer Engineering
Nanyang Technological University
639798 Singapore*

We present a location-aware routing protocol called MGPSR (Modified Greedy Perimeter Stateless Routing) for Mobile Ad Hoc Networks. MGPSR offers the crucial correctness guarantee of the well known Greedy Perimeter Stateless Routing (GPSR) protocol; moreover, it possesses two additional attractive properties: (1) the modified Greedy forwarding scheme balances between the transmission power consumption and the transmission latency, and (2) the modified perimeter forwarding makes use of the localized delaunay graph which offers higher connectivity and provides a path with fewer hops. In contrast to pure topology-based protocols, MGPSR does not drain the network bandwidth by imposing large amount of protocol traffic. Extensive simulations have demonstrated that the MGPSR outperforms the GPSR protocol in terms of planar graph connectivity and energy consumption.

Keywords: mobile ad hoc networks, planarization, location-aware routing, topology construction, delaunay triangulation, GPSR

1. INTRODUCTION

1.1 MANETs and Routing Protocols

In places where there is little or no pre-established communication infrastructure, the technology of Mobile Ad Hoc Networks (MANETs) allows mobile applications to maintain dynamic connections. Applications of such a network include mobile conferencing, emergency services, personal area networks, sensor dust and military communications [12].

In MANETs, autonomous wireless mobile hosts double as a router to transport information collaboratively. Since the topology in an ad hoc network changes unpredictably and frequently, an efficient routing protocol needs to determine high quality routes whilst holding the maintenance overhead to a minimum. Many protocols have been proposed, which can be broadly classified into *topology-based routing protocols* and *position-based routing protocols*. Metrics for evaluating the quality of a protocol are transmission latency, delivery energy, success rate and so forth. The reader is referred to [12] for an overview of the major approaches.

Received January 31, 2003; accepted July 4, 2003.

Communicated by Ruay-Shiung Chang.

* A preliminary version of this paper has been presented at the 2002 International Computer Symposium, 2002.

Position-Based Protocols and GPSR

Position-based protocols or location-aware protocols exploit the location information to facilitate routing. The routing decision relies on the destination's position to select the next forwarding host among the sender's one-hop neighbors. Routing a packet typically comprises two distinct phases: (1) discovering the position of the destination, and (2) the actual forwarding, based on the location information.

Each node determines its own location through positioning services like GPS [2] and obtains the location of the destination through a location service (see [7] for a survey), such as GLS with Geographic Forwarding [9].

A prime example of effective yet localized location-aware protocol is proposed by Karp and Kung in [8]: the *Greedy Perimeter Stateless Routing (GPSR)* protocol, which is nearly stateless and requires propagation of topology information for only a single hop. The proposal proves to be very elegant and practical.

Our research focuses on improving GPSR in the following aspects: (a) routing optimization in greedy mode to strike a balance between energy consumption and transmission delay, and (b) the construction and maintenance of dynamic connectivity information for better route quality in perimeter mode. The second aspect has been studied in [5], while the first has received little attention. Herein we present a novel scheme called the Modified Greedy Perimeter Stateless Routing (MGPSR) protocol. The quality of routing is secured by inheriting the correctness guarantee of GPSR; the density-based greedy forwarding redefines the route selection criteria to conserve energy; the protocol does not drain the network bandwidth by imposing large amount of protocol traffic.

1.2 Notations

The following conventions are adopted throughout the paper:

$d(x, y)$: The Euclidean distance between nodes x and y .

\overline{xy} : The edge connecting nodes x and y .

$Circle(O)$: The circle that has O as its center.

$Circle(x, y)$: The circle that has \overline{xy} as its diameter.

R : The one-hop broadcast radio range initially fixed for all nodes, which may not necessarily equal to the actual transmission radius.

$Nghd(x, R)$: The neighborhood of node x , i.e. the circular region within radius R with x at the center, which is alternatively written as $Nghd(x)$ if R is understood.

1.3 Organization

The rest of the paper is organized as follows. Section 2 surveys closely related work. Section 3 presents our Localized Delaunay Diagram. It uses Gabriel's Graph as its starting point. Section 4 details the density-based greedy forwarding strategy. Section 5 examines various properties of the MGPSR and compares it with GPSR via simulations. Section 6 highlights the main features of our approach and points out useful directions for future research.

2. PRELIMINARIES

2.1 GPSR Protocol in Detail

In the GPSR protocol [8], the sender of a packet incorporates the approximate position of the recipient into the packet. Whenever possible, a message will be routed nearer to its destination by *greedy forwarding*, i.e. forwarding the packet to the single-hop neighbor that makes the most progress towards the destination. The greedy routing may fail to find a path, even though one does exist. At any node, where none of its neighbors is closer to the destination, the perimeter forwarding will be applied, which essentially moves the packet around the void area by stipulating the *right-hand rule* (viz., rotate counterclockwise to get the first neighbor as the next hop) and the *face change rule*. A packet enters this recovery mode when arriving at a local maximum and returns to the greedy mode when it reaches a node closer to the destination than the perimeter mode entrance point. Karp and Kung demonstrated that these two methods can ensure the delivery of a packet in static network if there is an existing route.

Note that, to avoid circular tours caused by crossing edges, the underlying graph must be planar for perimeter forwarding. Karp and Kung have proposed to use a *planarization* scheme that results in a special type of planar graph named Relative Neighborhood graph (RNG) [14]. An alternative called Gabriel's Graph (GG) [6] was also suggested. Assuming uniform distribution, the computational overhead for adding or eliminating one edge per node is $O(1)$ for both graphs (cf. [10]).

RNG Definition Given two (arbitrary) nodes x and y , the edge \overline{xy} is in RNG if the distance between vertices x and y is less than or equal to the distance between any other vertex z , and whichever of x and y is farther from z . In other words, for any vertex z other than x and y , if $d(x, y) \leq \max[d(x, z), d(y, z)]$, \overline{xy} is an RNG edge.

GG Definition Given two arbitrary nodes x and y , an edge \overline{xy} is in GG if there exists no other vertex z such that $d(m, z) < 0.5d(x, y)$, where m denotes the mid-point of \overline{xy} .

2.2 Delaunay Triangulation (DT)

Both RNG and GG are rather sparse graphs. In [4], it was shown that, these graphs are more vulnerable to disconnection or partition in the presence of link failures. Therefore, it is desirable to design a denser planar graph which ideally (a) incurs low protocol overhead during construction and maintenance, and (b) incurs low operational cost in perimeter forwarding.

2.2.1 Delaunay triangulation (Global DT)

Global DT Definitions

Any circle in the plane is said to be *empty* if it encloses no vertex of a given set of vertices V (vertices are permitted on the circle). The *circumcircle* of a triangle is the unique circle that passes through its three vertices. A triangle is said to be *Delaunay* if and only if its circumcircle is empty.

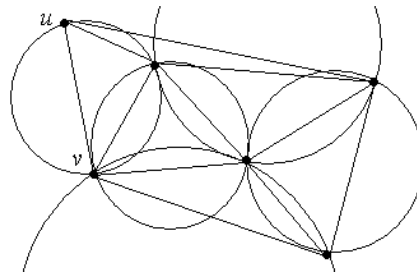


Fig. 1. Global delaunay triangulation.

Let u and v be any two vertices in V . A *circumcircle* of the edge \overline{uv} is any circle that passes through u and v . The edge \overline{uv} is in *Global DT* if and only if there exists an empty circumcircle of \overline{uv} . (Fig. 1)

The Delaunay Triangulation (DT) has long been regarded as a good spanner of a given set of nodes (refer to [10] for definition). It is also well known that DT is a superset of GG which in turn, is a superset of RNG [14]. The traditional algorithms for the construction of DT are not suitable for distributed environment because they either require prior knowledge of the *entire* set of vertices or employ incremental edge-flipping strategies that propagate an *unbounded* number of hops. With power constraints, it is impractical for all nodes to obtain and maintain positional information over a large area.

2.2.2 Localized delaunay diagram (LDD)

We developed an approximation to DT, which is feasible for local construction while ensuring the planarity, reachability and scalability.

LDD Definition Given a set of nodes V and a length R , a *1-hop Localized Delaunay Diagram (LDD)* of V contains all Delaunay edges that have a *length* $\leq R$.

A recent study by Gao et al. [5] described similar concepts and brought forward a distributed algorithm to construct such a graph. It works as follows. Each node acquires the position of its neighbors and computes the Delaunay triangulation in the one-hop circle. Since the local construction at different nodes could be inconsistent, additional information propagation is performed. Each node broadcasts its local Delaunay triangulation result to its immediate neighbors. A local DT edge \overline{uv} is deleted if it does not belong to the local Delaunay graph of any mutual neighbor w of u and v .

One of the drawbacks lies in the fact that, to validate the edge \overline{uv} when node v has just moved to u 's one-hop area, node u must wait until all common neighbors firstly become aware of v and recalculated local Delaunay Diagrams from them are received in the worst case. During that particular interval, the network topology may become invalid and crossing edges are temporarily possible. The problem of slow response to decommissioning of old links, or instatement of new ones would be serious especially under high mobility. Such drawbacks are eliminated from our protocol. In [5], transmission power issues are not addressed either.

3. LDD CONSTRUCTION

The local construction could be generalized from one-hop to k -hops. Intuitively the larger value k is, the better approximation to DT will result with correspondingly larger overhead. However, as we will demonstrate shortly, k need only be a small value for the nodes to obtain the necessary information for the local construction of 1-hop LDD.

3.1 One-Hop Local Delaunay Diagram Based on GG with Additional Crossing Edge Elimination (1^+ -GLDD)

Definition Let u and v be any two nodes in a set of given nodes V such that $d(u, v) \leq R$. Define the union of $\text{Nghd}(u)$ and $\text{Nghd}(v)$ as $\text{JointNghd}(u, v)$. A triangle that includes u and v in its set of vertices is said to be a 1-hop Local Delaunay Triangle (1-LDT) if and only if its circumcircle contains no other nodes than its vertices within $\text{JointNghd}(u, v)$.

Definition An edge \overline{uv} is in 1-GLDD if (a) it is in GG, or (b) there is a node w in $\text{Circle}(u, v)$ such that Δuvw is a 1-LDT.

Two mobile hosts u and v may obtain the knowledge about $\text{JointNghd}(u, v)$ through beaconing and exchange of neighborhood tables. An illustration is given in Fig. 2.

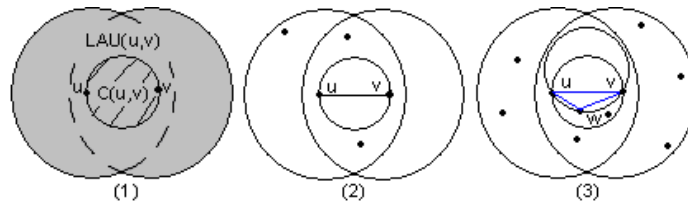


Fig. 2. 1-GLDD graph (1) The inner circle is $\text{Circle}(u, v)$ with diameter \overline{uv} . The shaded area (including $\text{Circle}(u, v)$) is $\text{JointNghd}(u, v)$, union of u ' and v 's one-hop radio areas. (2) \overline{uv} is in 1-GLDD since it is in GG. (3) \overline{uv} is not in GG ($\text{Circle}(u, v)$ is not empty). However, because the circumcircle of Δuvw does not contain any node from $\text{JointNghd}(u, v)$, \overline{uv} is in 1-GLDD.

Unfortunately, the simple-minded planarization method described in the preceding paragraph may result in crossing edges. Fig. 3 shows that, when there is a point such as node D , lying just beyond on the border (in the shaded region outside $\text{JointNghd}(A, B)$), the node may be still be able to set up a link with another node, such as node C , inside $\text{JointNghd}(A, B)$. In which case, neither node A nor B can detect the existence of node D , hence they will not be able to detect the edge between C and D , while C and D will maintain the edge CD as the triangle ACD contains no other vertices.

It can be shown that the scenario shown in Fig. 3 is the only possibility that causes edge-crossing in 1-GLDD. Please refer to Appendix A for the proof.

These rare cases are eliminated through an algorithm called 1^+ -GLDD, where the “+” sign denotes the additional crossing edge elimination step. We give the following variations to achieve this purpose.

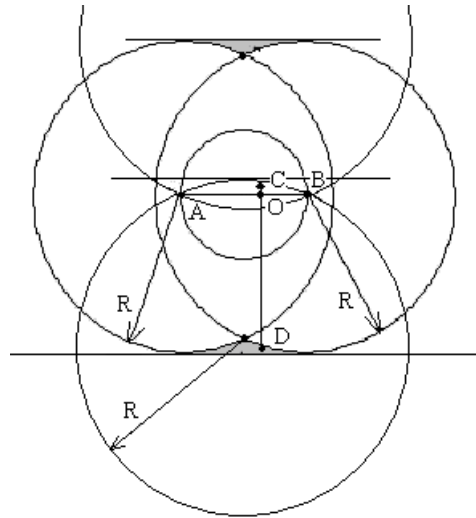


Fig. 3. The only possible case of crossing edges in 1-GLDD.

1. Each node extends the broadcasting radius to $\sqrt{5}R/2$ so that A and B can detect the presence of nodes in the shaded area, as shown in Fig. 4. Each node distinguishes neighbors within broadcast radius R and the extended area with a tag. Nodes from extended area are not used for routing, but only for checking of crossing edges.

The beaconing area is increased by $\frac{(\frac{\sqrt{5}}{2}R)^2 - R^2}{R^2} = 25\%$

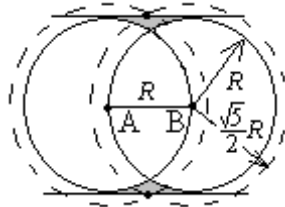


Fig. 4. Extended radius to cover D's possible location.

2. If a common neighbor C is found, nodes A and B will send a query packet to C independently. Upon receiving this packet, C checks its own neighbors, and reply to A and B whether there is a possible crossing edge.

The first scheme avoids network traffic produced by query packets at the expense of broadcasting energy, while the second consumes less transmit power in each round but induces extra rounds of transmission, a burden on the network bandwidth. The choice could be made according to actual requirements.

Time Complexity The computation and communication overhead incurred at each node is linear to the size of its neighborhood, which is bounded by a constant if the nodes are uniformly distributed; the overhead over the entire network is linear to the total number of nodes.

3.2 1⁺-GLDD Construction

An edge \overline{AB} is established immediately if in GG, otherwise, each node carries out 1-LDT test independently. Node A (B takes the same steps) first enumerates its 1-hop neighbors to find a node C such that the angle $\angle ACB$ is the largest among all (whose justification will be given in Claim 3.3). Then, it checks whether the circumcircle of ΔABC is empty of nodes in A or B's neighborhood. Neighbor information of B is obtained by neighbor table exchanges. If ΔABC is found to be a valid 1-LDT, the last step is the additional crossing edge checking.

Claim 3.2

- Assume that edge \overline{AB} is not in GG but is in 1-GLDD. Let ΔABD be any triangle formed by A, B, and any point D that lies inside Circle(A, B). If all such triangles like ΔABD are not 1-LDT, then none of the triangles in JointNghd(A, B) could be. (Fig. 5)
- For a non-GG edge \overline{CD} , if there exists a third point B such that $\angle CBD$ is the maximum angle among all common neighbors in circle Circle(C, D) and if ΔCBD is not 1-LDT, then none of the other triangles inside Circle(C, D) could be an 1-LDT. (Fig. 6)

If $\angle CBD > \angle CAD$ and ΔCBD is not 1-LDT, neither is ΔCAD .

Proof: Please refer to Appendix B.

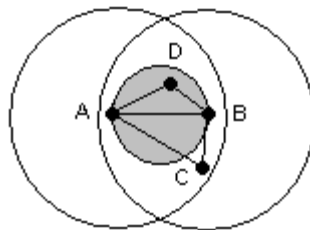


Fig. 5. Illustration of Claim 3.2a.

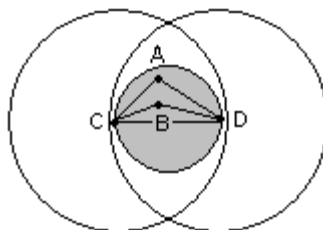


Fig. 6. Illustration of Claim 3.2b.

3.3 Properties of 1^+ -GLDD

It is known that the Delaunay Triangulation is both a Euclidean and topological spanner graph with constant *stretch factor*[♦] [5] in the sense that the distance between any pair of nodes on the graph is no more than constant times longer than the straight-line (Euclidean) distance. Let R denote a fixed radius. For a given set of nodes, the *unit-disc graph* is the graph resulting from allowing links between any two nodes, as long as the distance between them is within R .

The graphs constructed by the 1^+ -GLDD algorithm can be shown to be planar and free of unidirectional edges, moreover, it retains connection of the underlying unit-disc graph (Claim 3.3). As illustrated in section 5.2, the algorithm runs effectively in a dynamic and distributed fashion at a reasonably low cost, and the resulting graphs (cf. Figs. 12 and 13) are densely connected.

Claim 3.3

a) The graph 1^+ -GLDD retains the connectivity of the unit-disc graph.

Proof: The 1^+ -GLDD is trivially connected: it is a super graph of the GG, and the GG retains the connectivity of the underlying unit-disc graph [15]. \square

b) There is no unidirectional edge in 1^+ -GLDD.

Proof: Refer to Fig. 7. Assume that \overrightarrow{xy} is a unidirectional edge from x to y , i.e. x has an outgoing link to y , but y does not have a link to x . \overrightarrow{xy} could either be in GG or not in GG.

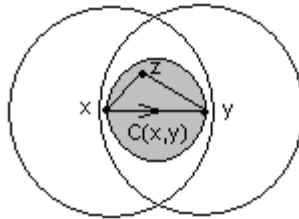


Fig. 7. Unidirectional edge is impossible in 1^+ -GLDD.

Case 1. \overrightarrow{xy} is in GG. In this case, the inner (shaded) circle $\text{Circle}(x, y)$ is empty. Since $\text{Circle}(x, y)$ is inside x and y 's 1-hop radio range, both of them should detect the emptiness of $\text{Circle}(x, y)$ and add each other as a neighbor after GG test.

Case 2. \overrightarrow{xy} is not in GG. In this case, x must have found a node z within $\text{Circle}(x, y)$, such that Δxyz is a valid 1-LDT. Since x and y share common knowledge of all

[♦] *Stretch factor* of a sub-graph G' of a graph G is the worst-case ratio of the length of a shortest path in G' to the length of the shortest path with the same endpoints in G .

nodes in their 1-hop neighborhood, it is impossible that only one of them knows Δ_{xyz} while the other does not. \square

c) There are no crossing edges in 1^+ -GLDD.

Proof: The only case where crossing edges are possible has been eliminated by 1^+ -GLDD algorithm. \square

4. DENSITY-BASED GREEDY FORWARDING

In most ad hoc networks, the life time of mobile nodes like hand-held devices depends on limited battery power supply. Given this, how to lengthen the lifetime of batteries while still maintaining a high packet success rate and low packet delay?

Energy is mainly consumed in (a) transmitting or forwarding data to the recipient or (b) maintaining the topological information in response to the changing network. The transmission success rate is largely dependent on the freshness of topological knowledge kept by nodes, and is thus constrained by the limited power capability. Given the energy budget, the node behavior could also affect the transmission latency in two ways: First, the more link information is made available, the better quality a route can be established; the routes stored are also more likely to be valid with frequent updates. Second, among routing options, forwarding to the nearest hop would reduce the transmission distance and thereby conserve power. Nevertheless, this may involve more intermediate hops, and thus the nodal delay, a major contributor of latency.

The tradeoff between battery power utilization and transmission latency is one major factor affecting the performance of MANET protocols. Our approach attempts to optimize and balance these two fundamental metrics.

Power Consumption Model The power E required in a transmission will be based on the relation $E \propto R^2$, where R is the transmission distance.

Beaconing range is unified among all nodes to assure the correctness of planarization algorithms. To obtain more topological information, the beaconing range should be larger; while to reduce power, smaller beaconing range is preferred. We also notice that the beaconing radius is closely related to the packet transmission distance. If it is always better to send the packet towards neighbors located at some distance, beacons need only cover this distance in order to gather enough routing information.

Greedy forwarding always chooses the neighbor closest to the destination within a fixed radius R , which is known as *most forward within R* or MFR in [13]. MFR minimizes the number of hops a packet has to traverse to reach the destination; hence, the nodal delay would be greatly reduced. It could be a good methodology when senders *cannot* adjust the signal strength to the transmission distance. However, the packet transmission range can be easily made adaptive. If the packet is sent to a nearest neighbor that is closer to the destination than the sender itself, the transmit power and the probability of packet collisions will be reduced significantly.

Consider an extreme case whereby all $N + 1$ nodes (inclusive of both the sender and destination) participating in the packet transmission are located evenly on a line. The

total energy E required for sending one packet is proportional to D^2/N , where D denotes the overall distance. For general cases under mobility, we will rely on experiments.

In previous studies of GPSR, neither beaconing range nor transmission range is adaptive. Karp and Kung simulated networks with a nominal 250-meter radio range in regions of density 1 node/9000m² [8]; Gao et al. monitored 300 nodes with fixed transmission radius 2 in a square of side length 24 [5]. To develop an energy-efficient scheme with a properly predefined beaconing radius and adjustable transmission distance, the following approach is adopted:

- 1) Hypothesis: there is a correlation between the optimal transmission range and the network density. The validity of this hypothesis is confirmed empirically (see section 5.3).
- 2) Based on the hypothesis, we define a one-hop radio radius R_{scg} that is just large enough to build a *Strongly Connected Graph (SCG)*[♦] for uniformly distributed networks.
- 3) Density-Based Greedy forwarding: node u always sends the packet to its neighbor v such that v is nearer to destination w and closest to the *energy-latency efficient point*, defined as the point lying on the line \overline{uw} and $d(u, v) = f \times R_{scg}$. Here, f denotes the factor to be determined through experiments and the beaconing radius R is slightly larger than $f \times R_{scg}$.
- 4) Measure the transmission delay and energy consumption for a wide range of f .

5. SIMULATION RESULTS AND DISCUSSION

5.1 Simulation Environment

Network In the simulation model, the initial node locations are generated by a random, uniform distribution over the plane. The experiments in sections 5.2 and 5.3 cover both sparsely- and densely-populated networks (Table 1). For comparisons with GPSR, the settings are exactly the same as those specified in [8].

Table 1. Network parameters.

Sec	Region (m ²)	Nodes	Density
5.2	800×800	20,	1node/32000 m ²
5.3		40,	1node/16000 m ²
		60,	1node/10667 m ²
		300	1 node /2133 m ²
5.4	1500×300	50	1 node/9000m ²
	2250×450	112	
	3000×600	200	

[♦] A graph in which it is possible to reach any node starting from any other node by traversing edges is defined as *Strongly Connected Graph (SCG)*.

Node Communication Live nodes periodically send out beacons to the neighborhood and upon receiving acknowledgement from neighbors, they append the information (address/ID and position) to the neighborhood table. The beacon interval is set to 1.5 seconds, and the neighbor information becomes stale (invalid) after a time out of 6 seconds. These are roughly the same as in GPSR [8]. For planarization, each node will determine locally whether a link is valid. To achieve maximal freshness, the planar edges are re-validated whenever an exchange of beacons takes place. Each node x will only be examining links within the circle centered at x of radius R . However, nodes could get information about nodes further away through exchange of neighbor tables.

All nodes share a fixed one-hop radio radius R for beacon broadcasting. However, unlike [8] and [5], we allow the packet transmission radius to vary with the approximate distance of the next hop to save power.

Movement Model Each node's motion obeys *random waypoint* model [1], i.e. it chooses a destination uniformly and randomly in the simulated region, chooses a velocity uniformly at random from a configurable range and then moves towards that direction. Upon arriving at the designated point, the node pause (dwells) for some time, and then repeats the same process. The mobility is affected by the speed as well as the pause time. Settings for experiment are shown in Table 2. (the settings for section 5.4 follows [8]).

Table 2. Movement parameters.

Sec	Speed	Pause Interval
5.3	0/10-20/30-40/50-60/ 70-80/90-100 m/s	0-60s
5.4	1-20 m/s	0/30/60/120s

Address Resolution In location-aware protocols, each node determines its own position through GPS or some other types of positioning service; the position of a recipient is provided by a location service whose design is beyond the scope of this paper.

In the current implementation, a central administrator tracks all node positions to simulate the location service. It is responsible to resolve location queries from node entities.

Computational Methodology After a packet is initiated; an internal mechanism of the simulator accumulates the hop numbers, transmission time and estimated energy at each intermediate transmission step. The packets are dropped when a timeout is encountered. The energy is attained by summing up the squared transmission distances. The number of hops and the success rate (delivery ratio) are computed for each simulation. The final result is averaged over 30 independent samples.

Other characteristics cf. [11].

5.2 Efficiency of 1⁺-GLDD Algorithm

For the completeness, we will cite a few relevant experimental results from [10].

5.2.1 Computation cost

1^+ -GLDD algorithm consumes about 10% more time on average than RNG and GG [10]. Compared with GG, additional time is spent on validating 1-hop Local DT edges. However, since the computation is distributed to individual nodes, the cost is acceptable and compensated by the much higher connectivity.

5.2.2 Connectivity

RNG, GG vs. 1^+ -GLDD

As shown in Figs. 8-13 (graphs for other densities are omitted), all three schemes are free of crossing edges.

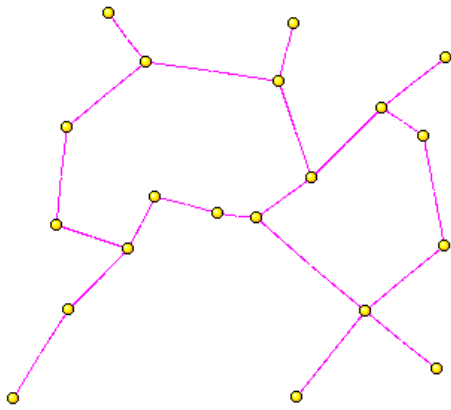


Fig. 8. RNG-20 nodes.

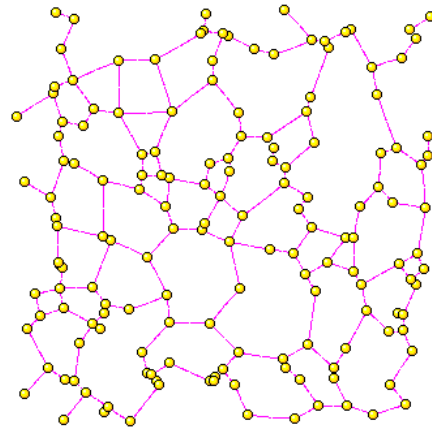


Fig. 9. RNG-160 nodes.

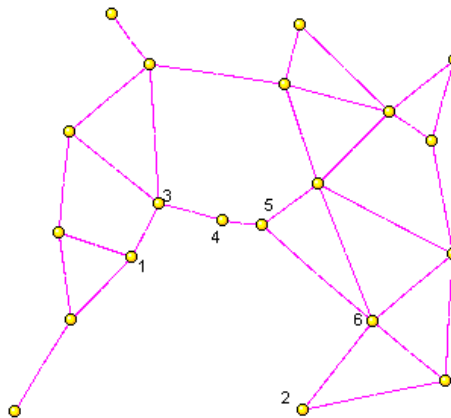


Fig. 10. GG-20 nodes.

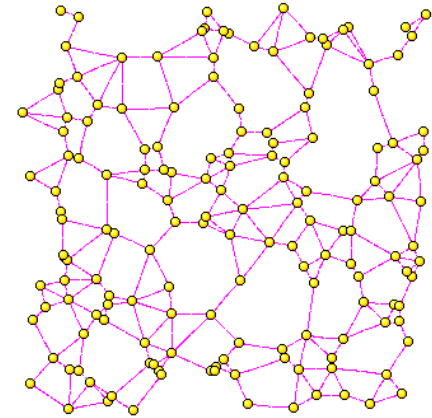
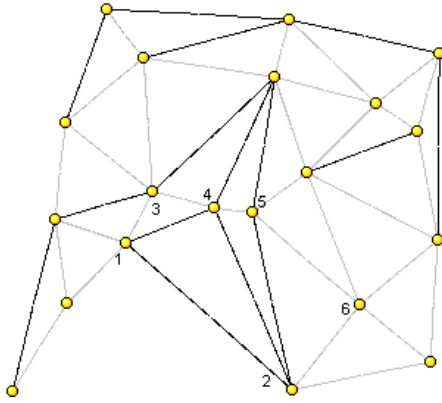
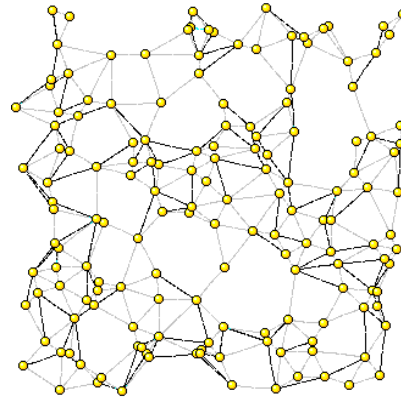


Fig. 11. GG-160 nodes.

Fig. 12. 1⁺-GLDD-20 nodes.Fig. 13. 1⁺-GLDD-160 nodes.

RNG offers the least connectivity in all settings of density: it could only build a very sparse graph and hence can easily become partitioned. More edges are present in GG. In Figs. 12 and 13 for 1⁺-GLDD, the gray edges are in GG, while darkened edges are those added after the 1-hop local Delaunay Triangle test.

The average number of links per node is listed in Fig. 14. 1⁺-GLDD graph, on average, contains 45% more edges than GG, and 104% more than RNG.

With much higher connectivity, mobile network topology built on 1⁺-GLDD is more robust and less prone to congestion. In perimeter mode, these connections could eliminate unnecessary intermediate stations. For instance, as Figs. 10 and 12 show, if node 1 is the perimeter mode entrance point and 2 is the greedy mode recovery point, routing on RNG or GG results in extra hops at nodes 3, 4, 5 and 6.

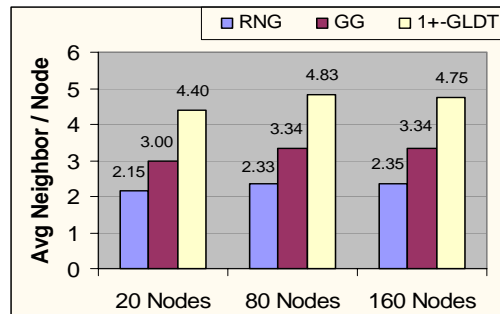


Fig. 14. Connectivity.

Global Delaunay Triangulation vs. 1⁺-GLDD

In addition to RNG, GG, and 1⁺-GLDD, Global Delaunay Triangulation is also constructed for each density by a modified Java program from [3]. It has been verified in [10] that *all* edges in 1⁺-GLDD can be found in the Global DT and except for edges longer than R , *all* edges in Global DT also appear in 1⁺-GLDD. This confirms that 1⁺-GLDD is a good approximation of the Global DT.

5.3 Density-based Greedy Forwarding

5.3.1 R_{scg}

Fig. 15 illustrates the relationship between physical connectivity and radio range (radius) for different numbers of evenly distributed nodes on the $800m \times 800m$ region. By inspecting this figure and constructing connectivity graphs, we obtained R_{scg} values listed in Table 3. Fig. 16 indicates that R_{scg} is inversely proportional to the square-root of the network density.

Table 3. R_{scg} values (Area = $800m \times 800m$).

Node Number	R_{scg} (m)	Node Number	R_{scg} (m)
20	380	140	130
40	240	160	110
60	190	180	110
80	170	200	100
100	170	300	83.5
120	150		

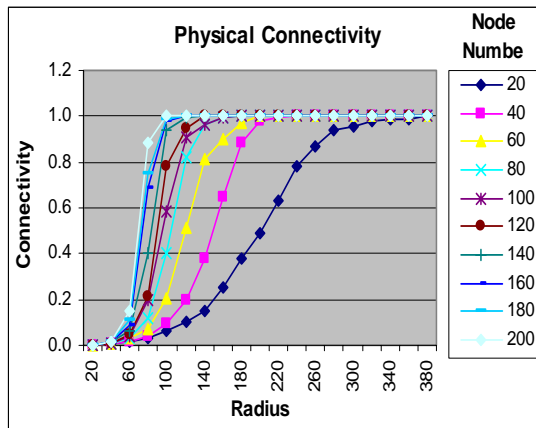


Fig. 15. Connectivity vs. radio range.

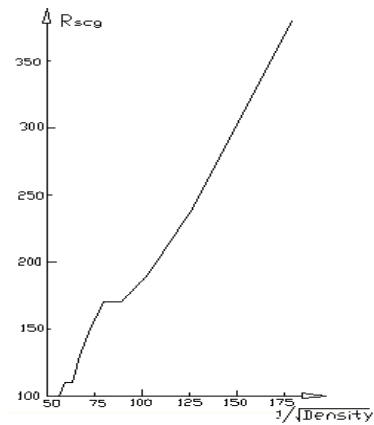


Fig. 16. R_{scg} vs. $\frac{1}{\sqrt{\text{density}}}$.

5.3.2 $f \times R_{scg}$ vs. energy and latency

In section 4, we designed an approach to locate the optimal transmission distance that could achieve a good balance between the delivery energy and latency. Figs. 17(static) and 18(mobile) reveal the relationship between these two metrics and the transmission radius (refer to [11] for more details).

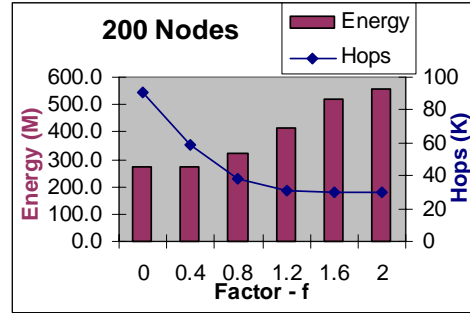
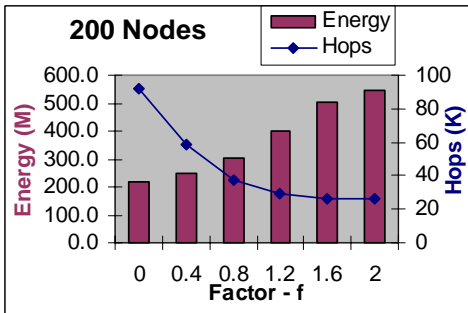
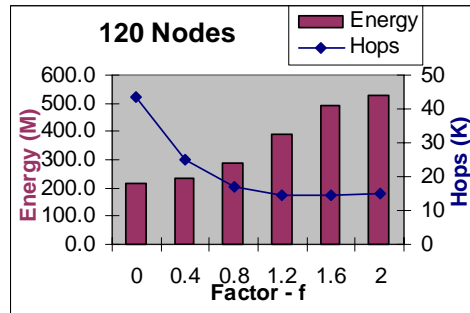
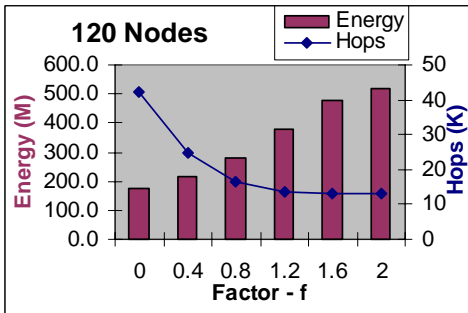
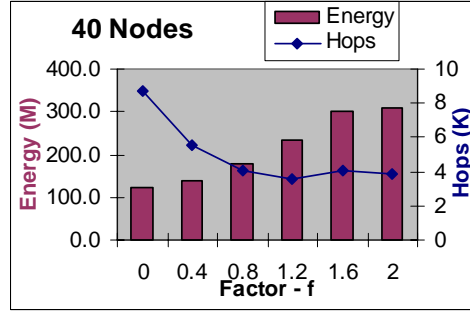
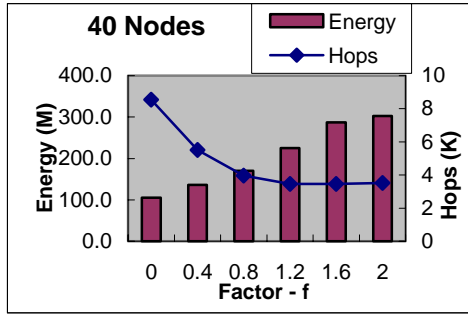


Fig. 17. Static Mode – Speed: 0.

Fig. 18. Mobile Mode – Speed: 50-60.

As expected, the energy consumption rises with ascending packet transmission radius. The shorter the distance between each pair of sender and receiver, the lower the energy cost. In contrast, the downward sloping curve of hop numbers suggests that routing time decrements fast with increasing factor, especially at small values. The existentially optimal balancing point is within the range of $0.6 - 0.8R_{scg}$.

One reasonable inference is that forwarding packets to neighbors at a distance of $0.6 - 0.8R_{scg}$ usually can alleviate the energy wastage with little increment in routing delay. Another noticeable fact is that a value slightly larger than $0.6 - 0.8R_{scg}$ could be set as the upper bound of the beaconing range. Furthermore, this range is closely related to the network density via R_{scg} . Consequently, the transmission range can be preset on mobile devices when the network density can be estimated in advance. Or, by analyzing the changes in the distribution, nodes may automatically regulate the transmission range.

5.4 MGPSR vs. GPSR

Most simulation parameters used in this section are consistent with GPSR. The value of R_{scg} is set at 210m and the beacon range is set at 250m to hold the beaconing energy at the same level as GPSR. Measurements in experiment (Fig. 19) show the energy efficiency has been achieved by MGPSR. However, it is worth noticing that because the vertical dimension (300, 450, and 600m) of the region is not much longer than the 250m radio range, very few voids could exist. That means, perimeter mode is rarely entered. The advantage of Delaunay-based graph in reducing nodal delays becomes insignificant, which probably accounts for the longer transmission latency. It was further shown that with such settings, the relationship between energy/hop number and the factor f exhibits similar properties as before [11].

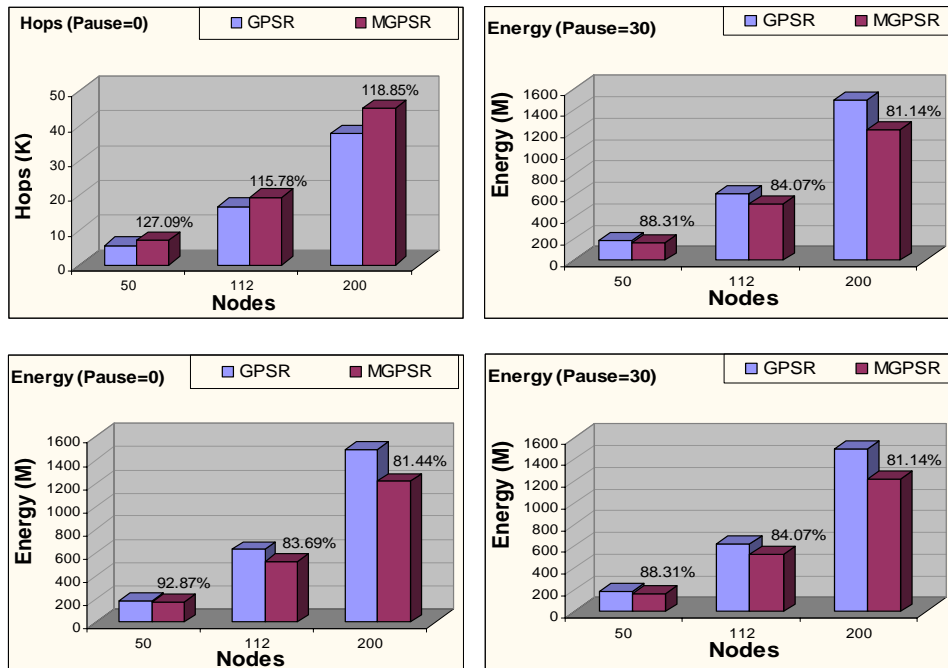


Fig. 19. GPSR vs. MGPSR (Factor = 0.7).

6. CONCLUSIONS

We have presented a new ad hoc network routing protocol, namely, the Modified Greedy Perimeter Stateless Routing (MGPSR) protocol. We have also compared this scheme with the original GPSR protocol on quite a number of networks with uniform distribution and simulation results quantitatively verified the merits of MGPSR in balancing energy consumption and delivery time.

In addition, the findings confirm that 1^+ -GLDD leads to considerably higher connectivity and, as such, it supports more efficient and robust routing. This higher connectivity comes at minimal additional cost. In fact, because of the planarity of the graph, the overall space required to maintain the graph on all nodes can be bounded by a linear function of the number of neighboring nodes. The time complexity for the construction and maintenance of the graph is also well within an acceptable bound. The instantaneous response to link-state changes fits well in mobile networks.

Extensive simulations demonstrated that the proposed density-based forwarding algorithm could minimize the amount of energy with little performance degradation in terms of delivery time. Severe constraint on battery resources heightened the need for such a power-efficient routing strategy. The idea behind this solution could be adopted in various routing protocols besides MGPSR.

In conclusion, MGPSR overcomes the present limitations of GPSR and it also widens its scope. The new protocol has been theoretically and empirically proved to be quite promising in the realm of mobile ad hoc networks. For future work, because the protocol relies on uniformity of node distribution, it will be of great interest to remove this assumption and further address the open problem of adaptive forwarding radius in non-uniform, clustered networks.

REFERENCES

1. J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM '98)*, 1998, pp. 85-97.
2. S. Capkun, M. Hamdi, and J. Hubaux, "GPS-free positioning in mobile ad-hoc networks," in *Proceeding of Hawaii International Conference on System Science*, 2001, pp. 3481-3490.
3. C. Denis, *Delaunay Triangulation Java Program*; <http://cage.rug.ac.be/~dc/allhtml/Delaunay.Html>.
4. D. Eppstein, "Spanning trees and spanners," in J.-R. Sack and J. Urrutia, ed., *Handbook of Computational Geometry*, Elsevier Science Publishers, B.C. North Holland, Amsterdam, 2000, pp. 425-461.
5. J. Gao, L. J. Guibas, J. Hershberger, L. Zhang, and A. Zhu, "Geometric spanner for routing in mobile networks," in *Proceeding of MobiHoc*, 2001, pp. 45-55.
6. K. Garbiel and R. Sokal, "A new statistical approach to geographic variation analysis," *Systematic Zoology*, Vol. 18, 1969, pp. 259-278.
7. J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *Computer*, Vol. 34, 2001, pp. 57-66.
8. B. Karp and H. Kung, "Geographic routing for wireless networks," in *Proceedings of the ACM/IEEE International Conference on Mobile Computer and Networking (MOBICOM)*, 2000, pp. 243-254.
9. J. Li, J. Jannotti, D. de Couto, D. Karger, and R. Morris, "A scalable location service for geographic ad-hoc routing," in *Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM 2000)*, 2000, pp. 120-130.

10. L. Luan and W. J. Hsu, "Localized delaunay triangulation for topological construction and routing on MANETs," Technical Report: CAIS-TR-02-46, Centre for Advanced Information Systems, School of Computer Engineering, Nanyang Technological University, Singapore, July 2002. Available from: http://www.cais.ntu.edu.sg:8000/Research_Projects/Technical_Reports/technical_reports.html.
11. L. Luan and W. J. Hsu, "Power-efficient geographic routing for MANETs," Technical Report: CAIS-TR-02-47, Centre for Advanced Information Systems, School of Computer Engineering, Nanyang Technological University, Singapore, July 2002. Available from: http://www.cais.ntu.edu.sg:8000/Research_Projects/Technical_Reports/technical_reports.html.
12. C. Perkins, *Ad Hoc Networking*, Addison Wesley, 2000.
13. G. Takagi and L. Klenrock, "Optional transmission ranges for randomly distributed packet radio terminals," *IEEE Transactions on Communications*, Vol. 32, 1984, pp. 246-257.
14. G. Toussaint, "The relative neighborhood graph of a finite planar set," *Pattern Recognition*, Vol. 12, 1980, pp. 261-268.
15. I. Stojmenovic and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks," *Wireless Networks*, Vol. 7, 2001, pp. 609-616.

APPENDIX A – PROOF FOR SECTION 3.1 (ONLY POSSIBLE CROSSING EDGE)

Refer to Fig. 1. Write circles centered at A, B with radius R as Circle(A) and Circle(B). Circle(A), Circle(B) intersect at points E and J. Draw the tangents \overline{FG} , \overline{HI} of Circle(A) and Circle(B). Denote the area enclosed by arcs EH, EI and line \overline{HI} as A(E, HI), by JF, JG and line \overline{FG} as A(J, FG).

Draw circle Circle(E) centered at point E with radius R. Draw a tangent \overline{MN} of Circle(E), \overline{MN} intersects Circle(E) at P, and $\overline{MN} \parallel \overline{AB}$. Define area enclosed by arc APB and line \overline{AB} as A(P, AB). Similarly, draw circle Circle(J), and define the area by arc AQB and line \overline{AB} as A(Q, AB).

___ If point D is in A(E, HI) and C is in A(P, AB) such that $d(C, D) \leq R$, \overline{CD} intersects \overline{AB} at O as shown.

Proof: Assume that there are two crossing edges \overline{AB} and \overline{CD} in 1-GLDD. $\overline{AB} \leq R$ and $\overline{CD} \leq R$ (Fig. 2). We divide it into two possible cases and prove each as follows.

Case 1: Pairs {A, B} {C, D} know about each other

Node B could know C by (1) beacon exchange if $d(C, B) \leq R$; (2) neighbor table exchange: B is able to obtain C's position through A, provided $d(C, A) \leq R$. The same holds for each of the pairs {B, D}, {A, C}, and {A, D}.

Since \overline{AB} is in 1-GLDD, there exists an empty circle O1 that circumscribes it. Similarly, empty circle O2 circumscribes \overline{CD} . There are three possible relative positions of O1 and O2 (Fig. 3). We will show that no crossing could exist in each case.

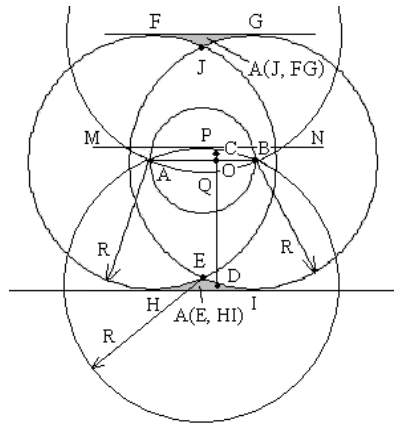


Fig. 1. The only possible case of crossing edges in 1-GLDD.

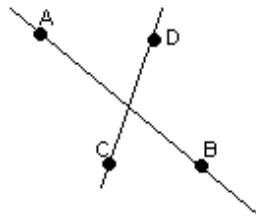


Fig. 2. Crossing edges.

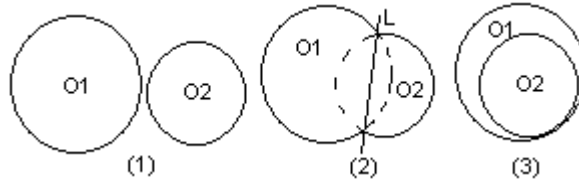


Fig. 3. Three possible positions of O_1 and O_2 , (1) there is no common intersection between O_1 and O_2 . Clearly, AB and CD cannot cross each other, (2) draw line L through the intersection points of O_1 and O_2 . If O_1 is empty and the pair $\{A, B\}$ knows about $\{C, D\}$, neither C nor D could sit on the dotted arc of O_2 at the left-hand side of line L . Similarly, neither A nor B could sit on the dotted arc of O_1 . It is obvious that no intersection could occur since AB and CD are on different sides of line L , (3) nodes C and D are on O_2 , thus inside O_1 , i.e. O_1 is not empty, which contradicts to our assumption.

Case 2: Nodes $\{A, B\}$ know one of $\{C, D\}$ or none of them

Refer to Fig. 4 which is a portion of Fig. 1. Define area below line HI as Region I. Define area above HI but outside $Circle(A)$, $Circle(B)$, $A(E, HI)$ and $A(J, FG)$ as Region II.

First, we prove if C or D is not inside A and B 's neighborhood, i.e. $Circle(A)$ and $Circle(B)$, it must locate in either $A(E, HI)$ or $A(J, FG)$. Here we choose D for proof, and the same argument holds for C . Assume that D is below Line AB (the case when D is above the line AB can be proved by symmetry). CD intersects AB at O .

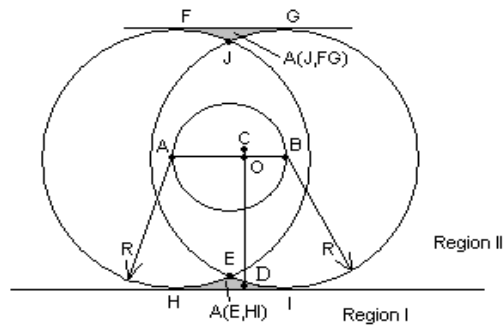


Fig. 4. D's location.

We prove that D must be in A(E, HI) when it is below Line AB by following two arguments.

1. D is not in Region I.

$d(O, D) < R$, since $d(C, D) \leq R$. Notice the distance between two parallel lines AB and HI is R. Therefore, distance from O to any point in Region I must be greater than R. Hence, D could not be in Region I.

2. D is not in Region II.

Draw a circle O' about O with radius R. Clearly, region II is beyond the circles with radius R about A and B. O' should never reach Region II since O lies in between nodes A and B.

Refer to Fig. 5. Draw circle Circle(E) about point E with radius R. Draw the tangent MN of Circle(E), MN intersects Circle(E) at P, and $MN \parallel AB$. Draw line ST through E, such that $ST \parallel AB$. Define the area enclosed by arc APB and line AB as A(P, AB). We prove if CD and AB cross each other and D is in A(E, HI), C must be in A(P, AB).

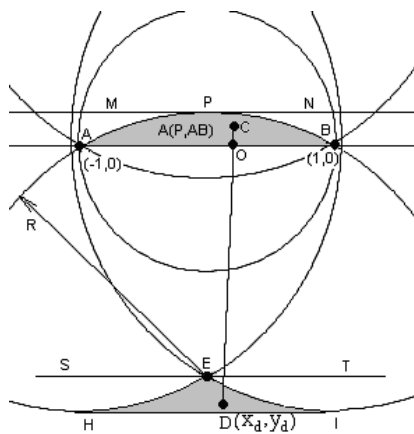


Fig. 5. C's possible locations.

Let the coordinates of A, B and D be $(-1, 0)$, $(1, 0)$, and (x_d, y_d) .

$$(x_d - 1)^2 + y_d^2 > R^2 \quad (1)$$

$$(x_d + 1)^2 + y_d^2 > R^2 \quad (2)$$

$$-R < y_d < 0, \quad (3)$$

$$\text{and, } -1 < x_d < 1 \quad (4)$$

Write the equation of circle Circle(D) about D with radius R,

$$(x - x_d)^2 + (y - y_d)^2 = R^2$$

If Circle(D) intersects \overline{AB} at point $(x_0, 0)$,

$$x_0 = x_d \pm \sqrt{R^2 - y_d^2} \quad (5)$$

From (1), (2), (4) and (5),

$$-1 < x_0 < 1$$

When D is at E, x_0 has maximum values -1 and 1 , and the area of Circle(D) above line AB is A(P, AB).

Since point C is inside circle Circle(D), it must be in A(P, AB).

Therefore, crossing is only possible when both conditions (D in A(E, HI) and C in A(P, AB)) are met. Furthermore, the smaller $d(A, B)$, the smaller A(E, HI), and the possibility of crossing edges gets even lower (Comparison of the two cases with different $d(A, B)$ is shown in Fig. 6).

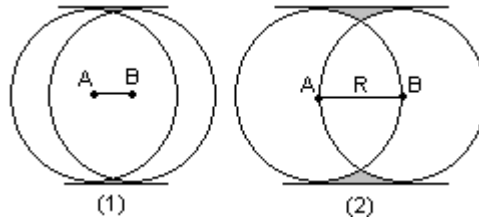


Fig. 6. Comparison between two cases. When the distance between A, B becomes smaller, the chance of crossing edges is even less since the area in which D could locate becomes smaller.

APPENDIX B – PROOFS FOR SECTION 3.2

Proof: A lemma is needed for the proof of claim 3.3a and 3.3b:

Refer to Fig. 7. $Circle(O_1)$ and $Circle(O_2)$ are two circles passing through points $A(-1, 0)$ and $B(1, 0)$, with centers $O_1(0, y_{o1})$ and $O_2(0, y_{o2})$ ($y_{o1} > y_{o2}$). Point (x_1, y_1) is either on or inside $Circle(O_1)$, and point (x_2, y_2) is either on or inside $Circle(O_2)$.

Lemma If $y_2 > 0$, then (x_2, y_2) is in circle Circle(O_1); if $y_1 > 0$, then (x_1, y_1) is in circle Circle(O_2).

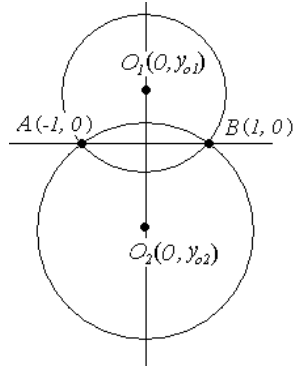


Fig. 7. Lemma.

Proof of Lemma: Write the equations of circle Circle(O_1) and Circle(O_2) as

$$x^2 + y^2 - 2y_{o1}y = 1 \quad \text{and} \quad x^2 + y^2 - 2y_{o2}y = 1$$

If $y_2 > 0$ and $y_{o1} > y_{o2}$

$$1 \geq x_2^2 + y_2^2 - 2y_{o2}y_2 > x_2^2 + y_2^2 - 2y_{o1}y_2$$

Hence, (x_2, y_2) is in circle Circle(O_1).

Similarly, (x_1, y_1) is in circle Circle(O_2). □

Proof of Claim 3.1b: For a non-GG edge \overline{AB} , there must be a third point $M (M \in C(A, B))$ such that $\angle AMB$ is the maximum angle among all common neighbors in circle Circle(A, B). If $\triangle AMB$ is not 1-LDT, we need to prove none of other triangles inside Circle(A, B) could be.

As shown in Fig. 8, O is the midpoint of line \overline{AB} . $\overline{CD} \perp \overline{AB}$ at O . Draw circum-circle Circle(A, M, B) of $\triangle AMB$, Circle(A, M, B) intersects CD at M' . Write point M'' such that $d(M'', O) = d(M', O)$.

It is clear that the area enclosed by arc $AM'B$ and arc $AM''B$ is empty of nodes. Otherwise, if there is a point N in this area, $\angle ANB$ must be greater than $\angle AMB$, which violates the condition that $\angle AMB$ is maximum.

If $\triangle AMB$ is not 1-LDT, we now prove there is no point P , which is in the area enclosed by arc ACB and arc $AM'B$, such that $\triangle APB$ is 1-LDT.

Since the circumcircle Circle(A, P, B) of $\triangle APB$ intersects with Circle(A, M, B) at points A and B . By the lemma, Circle(A, P, B) must at least contain point M . Hence, $\triangle APB$ is not 1-LDT.

Similarly, there does not exist any point in the area enclosed by arc ADB and arc $AM''B$ that can form a 1-LDT with points A and B .

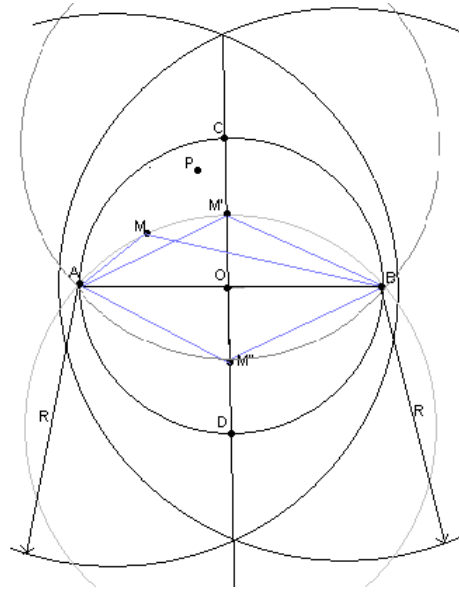


Fig. 8. Claim 3.1b.

Proof of Claim 3.1a: Refer to Fig. 9.

Let Circle(A, B) be a unit circle, with A located at $(-1, 0)$ and B at $(1, 0)$. If a node $C(x_c, y_c)$ ($y_c < 0$) outside Circle(A, B) exists such that ΔABC is a valid 1-hop Local Delaunay Triangle, circumcircle of ΔABC must be empty.

Define $D(x_d, y_d)$ ($y_d > 0$ since the circumcircle of ΔABC is empty) as a node in Circle(A, B) such that $\angle ADB \geq \angle AEB$, where E is any node other than D in Circle(A, B).

If the circumcircle of ΔABC is empty, the same holds for the circumcircle of ΔABD .

Denote circumscribing circles of ΔABC and ΔABD as Circle(O_c) and Circle(O_d), whose centers are at $(0, y_{oc})$ and $(0, y_{od})$ respectively.

We will show that ΔABD must also be a 1-LDT. The claim follows.

For Circle(O_d) we have: $x^2 + (y - y_{od})^2 = 1 + y_{od}^2$.

Since $D(x_d, y_d)$ is on Circle(O_d),

$$y_{od} = (x_d^2 + y_d^2 - 1) / 2 y_d \tag{1}$$

Since Circle(O_c) is empty, $D(x_d, y_d)$ must be outside of it, hence,

$$\begin{aligned} x_d^2 + (y_d - y_{oc})^2 &> 1 + y_{oc}^2 \\ \text{i.e. } x_d^2 + y_d^2 - 1 &> 2 y_{oc} y_d \end{aligned} \tag{2}$$

$$\text{From (1) and (2), } y_{od} > \frac{2 y_{oc} y_d}{2 y_d} = y_{oc} \quad (y_d > 0)$$

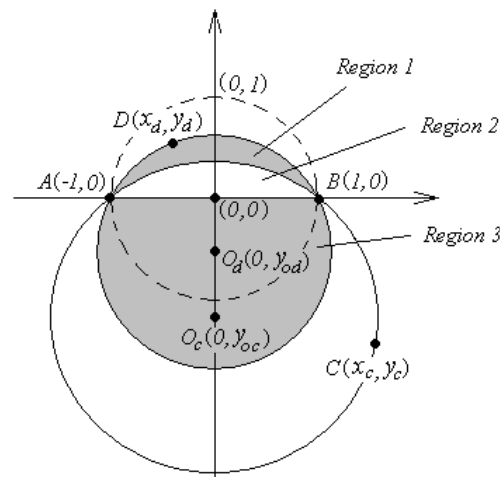


Fig. 9. Claim 3.1a.

Define Region 2 as the area of $Circle(O_c)$ above x-axis, Region 3 as the area of $Circle(O_d)$ below x-axis. From Lemma 3.1, Region 2 is contained in $Circle(O_d)$ and Region 3 is in $Circle(O_c)$. Since $Circle(O_c)$ is empty from the assumption, these two portions of $Circle(O_d)$ are also empty.

Define Region 1 as the part inside unit circle $Circle(A, B)$ and $Circle(O_d)$ but outside $Circle(O_c)$. By Claim 3.1b and assumption placed on point D, it is also void of nodes.

Since all three regions are empty, $Circle(O_d)$ is empty and thus, $\triangle ABD$ is 1-LDT. The claim is true. \square

Luan Lan is presently a graduate student in Syracuse University, N.Y., U.S.A. She graduated from the School of Computer Engineering, Nanyang Technological University, Singapore. She was the recipient of the prestigious Lee Kuan Yew Gold Medal and the Institution of Engineers Singapore Gold Medal as well as the Sony Book Prize for the best Final Year Project. The paper is part of her work during her final year project.

Wen-Jing Hsu is presently a faculty member of the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. Dr. Hsu was the Director of Center of Advanced Information Systems, Deputy Director of Financial Engineering, and is currently the Deputy Director of the Maritime Research Center, all of NTU. During his tenure of the directorship, he has helped getting several multi-disciplinary research projects. He is a senior member of IEEE, a Faculty Fellow of the Singapore-MIT Alliance Program.

Rui Zhang, a former graduate student from Nanyang Technological University, was from China.