Optimal Geographic Routing for Wireless Networks with Near-Arbitrary Holes and Traffic

Sundar Subramanian and Sanjay Shakkottai
Department of ECE
The University of Texas at Austin
Email: {ssubrama, shakkott}@ece.utexas.edu

Piyush Gupta
Complex Systems Analysis & Optimization Department
Bell Labs, Alcatel-Lucent
Email: pgupta@research.bell-labs.com

***Abstract*—We consider the problem of throughput-optimal routing over large-scale wireless ad-hoc networks. Gupta and Kumar (2000) showed that a throughput capacity (a uniform rate over all source-destination pairs) of $\Theta(\frac{1}{\sqrt{n \log n}})$ is achievable in random planar networks, and the capacity is achieved by straight-line routes. In reality, both the network model and the traffic demands are likely to be highly non-uniform. In this paper, we first propose a randomized forwarding strategy based on geographic routing that achieves near-optimal throughput over random planar networks with an arbitrary number of routing holes (regions devoid of nodes) of varying sizes. Next, we study a random planar network with arbitrary source-destination pairs with arbitrary traffic demands. For such networks, we demonstrate a randomized local load-balancing algorithm that supports any traffic load that is within a poly-logarithmic factor of the throughput region. Our algorithms are based on geographic routing and hence inherit their advantageous properties of low-complexity, robustness and stability.**

I. Introduction

We study the problem of throughput-optimal routing in large wireless networks such as ad-hoc and sensor networks. In such large networks, there is a need for scalable, low-complexity and distributed routing algorithms that can provide good data rates for the traffic flows. The work in [8], [6] has shown that a throughput-capacity of $\Theta(\sqrt{\frac{1}{n}})$ is achievable in *uniform networks* with *uniform traffic* demands, and that the capacity achieving routes are straight-line paths. In many practical networks, both the network and the traffic distribution may be highly non uniform. Non-uniformities may arise due to factors such as network holes (regions devoid of any living nodes), the arbitrary locations of source-destination pairs or due to variations in the required data rates.

In recent studies [9], [10], [17], [11], [5], geographic forwarding based protocols have been suggested as a stable routing (providing fixed routes that do not flip) technique over large non-uniform networks as they are scalable, low-complexity and highly distributed. However, in recent work [19], it was demonstrated that network non-uniformities can cause significant losses in throughput (rates could be as low as $\Theta(1/n)$) while employing such schemes. A critical issue is that conventional "shortest-path" (such as straight-line) routes are oblivious to the distribution of other routes (between other source-destination pairs) and may cause heavy losses in throughput due to spatial congestion.

Typically, throughput optimal routing schemes over non-uniform networks and traffic demands are based on (i) solving a global optimization problem [1] (setting up routes such that the traffic balanced over the wireless links) or (ii) adaptive schemes [21], [18] that converge to an optimal set of routes (or a per-packet route) over time. While global optimization requires co-ordination and heavy computation by the nodes, adaptive schemes may take a long time to converge to good paths and also have issues of stability.

In this paper, we are interested in developing distributed routing algorithms that are "near-optimal" (close to the rates obtained by a global optimization) over non-uniform networks with arbitrary traffic demands, but are still low-complexity, distributed and stable.

A. Main Contributions

We consider a random planar network with n nodes arbitrarily distributed over a unit region, with each node having a uniform circular radio range of $M(n) = C\sqrt{\frac{\log n}{n}}$, for any $C > \frac{1}{\sqrt{\pi}}$. This scaling ensures that the resultant graph is connected [7].

(a) We first consider non-uniform networks with large number of routing holes and $n/2$ uniformly randomly distributed source-destination pairs. In contrast to earlier work [19], with finite number of holes of constant area, we allow for an arbitrary number of holes of varying sizes. Over such networks, we demonstrate that a near-optimal throughput capacity of $\Theta(\sqrt{\frac{1}{n}})$ is achievable (up to poly-logarithmic factors) by our algorithm RandHT(n). Unlike the RANDOMWAY algorithm [19], the new algorithm does not overload the network with increasing number of holes, and is also oblivious to the number of holes in the network.

(b) Next, we consider networks with an *arbitrary number* of source-destination pairs with *arbitrary locations* and *varying rate* requirements. We assume that the network however has no routing holes. Conventionally, cut-set bounds (amount of traffic that can enter/leave the boundary of any sub-region of the network) have been used to characterize upper bounds on network capacity [14], [12]. However, when sources and sinks can be arbitrarily close or far, and with widely varying traffic requirements, cut-set bounds alone are insufficient to characterize network loads. This is because traffic flows that never leave the region are unaccounted by such cuts. In this paper, we jointly utilize transport capacity bounds (bounds that arise due to the interference nature of the channel) along with cut-set based bounds to characterize the allowable set of source-destination pairs. That is, we demonstrate that *all* routing scheme have a local conservation property, where using less amount of the local cut capacity requires it to use more of the limited local transport-capacity, and vice-versa. Using this key property, we demonstrate that locally load balancing the traffic by means of a low-complexity randomized algorithm (RandLLB) is optimal up to a poly-logarithmic factor. Unlike RANDOMSPREAD [19], this algorithm does not assume that the source-destination pairs are $\Theta(1)$ away from each other, and distributes the traffic over an appropriate area rather than over the whole network as performed in RANDOMSPREAD.

Finally, we discuss some considerations when implementing these algorithms in practical deployments. These include combining RandHT and RandLLB algorithms over mixed networks and incorporating GPSR-like algorithms [9], [5] into RandHT in order to guarantee (low-rate) connectivity in "worst-case" network topologies.

B. Related Work

Routing in large wireless networks has been widely studied in the past decade (see [15] for an overview). Many of these algorithms are derived from Internet routing protocols, and do not scale well (in terms of route setup, routing table complexity) in large networks. Recently, geography based routing algorithms [9], [10], [17], [11], [5] have been investigated for providing low-complexity routing protocols that are scalable and stable. In these schemes, packets are greedily routed towards the location of the destination node and if the greedy routes are trapped in a routing "local minima", techniques such as planarization and face-traversal are used to route around these "holes". However, in recent work [19], we demonstrated that traditional shortest path schemes (such as DSDV or AODV) and greedy geographic schemes (with face-traversal) can cause heavy throughput losses in the presence of network non-uniformities or unbalanced traffic demands. As source-destination pairs setup routes without knowledge of other flows in the network, greedy or shortest path routing can cause spatial congestion. Certain randomized strategies to route around holes were suggested in [3]. However, such schemes may fail in networks with typical hole configurations, and even when working, may provide low throughput.

Traditionally, throughput-optimal schemes have either been based on a global optimization [1] (the routes are centrally chosen to balance the flows over the network) or on adaptive schemes where the packets are routed according to current queue/traffic states (e.g., back-pressure algorithms). Two main drawbacks of global schemes are the need for network-wide coordination and the high complexity of solving the optimization. On the other hand, while adaptive schemes that use only local coordination have been developed [21] (and more recent follow-ups [18], [13]), these algorithms have to contend with issues of stability, slow convergence to optimality (especially in large-scale networks) and long packet delays.

Randomized approximations to throughput optimal routing have been studied since the classical paper of [22], where the traffic flow is distributed equally over the entire network and recombined at the destination. We used an extension of this idea in [19] to provide near-optimal throughput in networks where the source-destination pairs were randomly chosen. Further, in [19] we assumed networks with a finite number of routing holes (whose size was comparable to the network size), and randomly distributed source-destination pairs (thus the typical distance between a source and its destination was comparable to the network diameter). For such networks, we proposed scalable and distributed algorithms based on geographic schemes that were near throughput-optimal.

In this paper, we allow for (i) a more complex network topology and (ii) an arbitrary number of arbitrarily located source-destination pairs with variable traffic requirements. We note that Valiant-like schemes (as used earlier in [19]) for distributing traffic over the whole network are provably sub-optimal as they require packets to be unnecessarily transported over long distances. Such networks require new routing algorithms as well as different proof techniques to demonstrate optimality, as we shall show in the rest of this paper.

II. System Description

We initially consider a random planar network where n nodes are randomly and uniformly thrown over a unit torus (a square region with wrap-around at the edges). We allow for a uniform circular radio range, $M(n) = C\sqrt{\log n/n}$, to ensure connectivity and a non-zero number of nodes in any tile of size $M(n) \times M(n)$.

A. Networks with Routing Holes

In the first part of the paper, we consider a random traffic pattern where $n/2$ source-destination pairs are chosen uniformly randomly from the torus. Further, we allow for an arbitrary number of holes to occur on the network. We ignore the traffic generated by any source or destination node that are removed by the occurrence of a hole. We assume the following conditions on the holes. (After the occurrence of holes with these assumptions, the connectivity and the non-zero nodes in any tile $M(n) \times M(n)$ outside the hole is preserved.)

Assumption 2.1: Hole placements: Let δ_r be the side of the smallest unique axis-parallel square that contains the hole r, and $\epsilon_r = \delta_r(1 + \Delta)$ be the side of a larger concentric square around the hole. Then, no other hole t can be placed such that its ϵ_t outer square can intersect with that of hole r. This ensures that each hole is separated from any other hole by a distance proportional to its diameter.

Assumption 2.2: Hole shapes: Consider the tiling of the unit region by square tiles of dimension $p \times p$ for some small

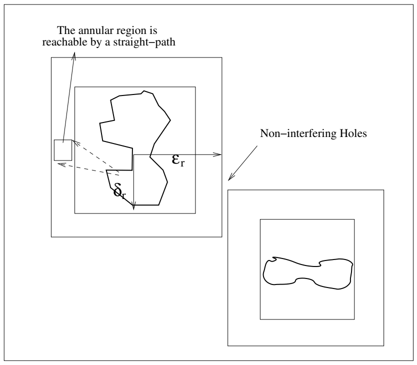

Fig. 1. Assumptions on holes in wireless networks.

$p > n^{\gamma-1/2}$ for some $0 < \gamma < 1/2$. Then the holes are composed by the union of contiguous tiles. Further, any node A in the interior of the δ-square can reach any point in the annular region between the ϵ and the δ-squares by straight line not intersecting the hole. For an illustration, see figure 1. Thus, any surviving nodes inside the δ-square are easily reachable from outside the hole. This assumption essentially disallows the formation of holes with complex topologies that house surviving nodes that are extremely hard to reach by local search methods. The holes can still be concave.

Holes can be chosen arbitrarily to occur over the network subject to the above assumptions. These assumptions are similar to [19] - however, we now allow for a large number of holes of varying sizes to occur on the network. Thus, the hole sizes are decoupled from the network size. We note that this allows for a significantly larger class of non-uniform network topologies. We also note that after the removal of nodes due to holes, the number of surviving source-destination pairs are $\Theta(n)$ (*w.h.p*).

B. Networks with Arbitrary Traffic

In the second part, we are concerned with the issue of traffic non-uniformity in networks. Here, the random planar network is without any routing holes, but with arbitrarily chosen source and destination pairs from the network.

Formally, we allow for H source-destination pairs, with $0 \leq H \leq n^2$ and with l-th source-destination pair ($l \in \{1, 2, \cdots, H\}$) at a distance $n^{\alpha_l - 1/2}$ away from each other, for $0 < \alpha^* \leq \alpha_l \leq \frac{1}{2}$. The algorithms and proofs described immediately extend to any constant scaling of the distance model described above. However for notational simplicity, we keep the constant as unity. The rate required by any flow is assumed to be from a finite set $\mathcal{R} = \{\frac{1}{\lceil n^{\gamma_1} \rceil}, \cdots, \frac{1}{\lceil n^{\gamma_{|\mathcal{R}|}} \rceil}\}$, with $0 \leq \gamma_i < \infty$. Thus, for a given source-destination configuration, a rate vector $\bar{r} = [r_1, \cdots, r_H]$, $r_l \in \mathcal{R}$ describes the traffic demand. In other words, the S-D pairs may be arbitrarily close to each other (compared to network diameter). An $\alpha = \frac{1}{2}$ signifies source-destinations that are a unit distance away from each other, and an $\alpha = 0$ a distance of $\frac{1}{\sqrt{n}}$, the average distance between nearest neighbors in a random planar network. Further, we assume the finite-level rate model only to illustrate our proof method clearly. We can extend it to any rate model by assuming a non-integral number of sources of the basic rate that are collocated, and our proof method can be used to show this result.

C. Interference Model and Standard Definitions

Definition 2.1: The throughput capacity $T(n)$ of a network is defined as the maximum data-rate that is simultaneously achievable by all surviving source-destination pairs.

Also, we assume the following to model the interference effects of simultaneously transmitting nodes which are within each other's radio range.

Definition 2.2 (Protocol Model, [8]): A transmission between a node A and its receiving node B is assumed to be successful if $d(A, B) \leq M(n)$ and $d(C, B) > (1 + d)M(n)$, for some $d > 0$, for all other transmitting nodes $C \neq A$.

This successful transmission occurs at rate '1' WLOG. We define the packet delay $D(n)$ as the maximum time taken by the routing algorithm to travel from the source to its destination over all source-destination pairs.

We define $f(n) = \tilde{\Theta}(g(n))$ if $f(n) = O(g(n)(\log n)^k)$ and $g(n) = O(f(n)(\log n)^{k_1})$ for some $k, k_1 < \infty$, and thus, a throughput $T(n)$ is near-optimal if it achieves $\tilde{\Theta}(T^*(n))$, where $T^*(n)$ is the optimal throughput.

III. Routing with Network Holes

In this section, we consider the problem of routing over a network with a large number of holes - in particular, we consider networks in which the number of holes may be comparable to the number of nodes in the network. An important question is to determine if geographic forwarding based schemes can provide routing strategies that are throughput and delay optimal.

Geography based routing schemes are preferred for routing over large networks predominantly for two reasons. Firstly, the routing information is scalable, i.e., the amount of routing information that a node needs to remember is proportional to the number of its neighbors and does not increase significantly with the network size. Secondly, the routing strategy is stable, low complexity and scalable - the routes are chosen in a greedy geographic manner, and hence the routes are easily computed and do not flip/switch due to the loss or the addition of a few extra nodes.

In earlier work [19], we studied a network with a finite number of holes, and demonstrated that pure greedy forwarding strategies such as GPSR can cause the throughput capacity of the network to be considerably reduced. We also proposed a randomized forwarding algorithm (RANDOMWAY) that was throughput optimal (while inheriting the nicer properties of geographic routing schemes) for networks with a finite number of constant area 'holes'. While the routing scheme was oblivious to the actual location of the holes, a drawback of the proposed scheme was (i) an exponential drop in throughput with increasing number of network holes, (ii) the algorithm required a knowledge of the number of holes in the network.

Field Name	Functionality
TOPOLOGY or DATA	Toggle bit - Topology information or Data Packet.
TOPOLOGY DATA	Information about Hole location and dimension
SRC-LOC	The ID and location of source
STAGE	The stage of routing
NEXT-DEST	Location of the next waypoint
SEC-DEST	Location of next+1 waypoint
FINAL-DEST	Location and ID of the original destination
DATA	Message to the destination node

TABLE I
FIELDS IN THE HEADER OF THE PACKET.

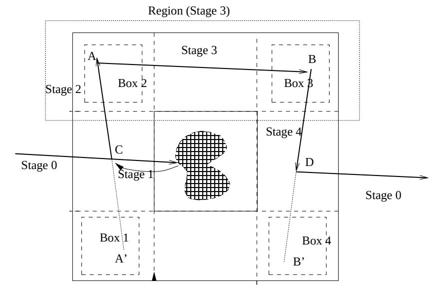

Fig. 2. RandHT algorithm - Routing around a hole.

In this section, we propose a randomized routing algorithm based on greedy forwarding that provides near-optimal throughput and delay even in the presence of a more complicated network topology (an arbitrary number of network holes), and operates without the knowledge of the number of holes. We also characterize the scaling laws for its throughput, delay and routing information at each node. The network model is as described in Section II-A.

A. The RandHT(n) Algorithm

We first define a packet structure to provide a common communication scheme between nodes. See Table I.

The source node while sending out a data packet sets the data flag bit, and sets its SRC-LOC and FINAL-DEST. It sets STAGE = 0, NEXT-DEST = FINAL-DEST and other fields to a NULL symbol.

We shall initially assume that the nodes that are on the boundary of a hole h know the dimensions and location of the smallest (up to an order) axis-parallel square that contains the hole h, i.e., they know the pair $\{xmin(h), ymax(h)\}$ which are the end points of the diagonal of the containing square. We denote this square as $Sq(h)$. We will shortly describe an update scheme by which the nodes on the hole boundary can obtain this data. The randomized hole traversing algorithm (RandHT(n)) is defined as follows (See Figure 2.)

Algorithm RandHT(n):

A node on receiving a packet with the data flag set (i.e., signifying that it is a data packet) checks if the FINAL-DEST id is identical to its own. If yes, it accepts the packet. Else it checks if it is on the boundary of a hole.

If the node is not on the boundary of a hole, it first checks if its node location 'matches' (within a radio-range hop) the NEXT-DEST. If that does not match its own location, it forwards the packet greedily towards NEXT-DEST. If it is the NEXT-DEST,

1) The node checks the STAGE to see what stage of routing the packet is in. If STAGE = 0, NEXT-DEST is always FINAL-DEST. The node would have already accepted the packet.
2) If STAGE = 1, it updates STAGE = 2, and sets NEXT-DEST = SEC-DEST and clears SEC-DEST to null, and forwards the packet to neighbor closest to the new NEXT-DEST.
3) If STAGE = 2, it picks a random location B from the $\Delta^2 Sq(h)$ Box 3 and sets NEXT-DEST = B, STAGE = 3 and forwards packet greedily towards B.
4) If STAGE = 3, it picks a random location B' from the $\Delta^2 Sq(h)$ Box 4 and sets NEXT-DEST = the intersection of BB' and the line joining source and destination, sets STAGE = 4, and greedily forwards towards NEXT-DEST.
5) If STAGE = 4, it sets NEXT-DEST = FINAL-DEST, STAGE = 0, and greedily forwards towards NEXT-DEST.

If it does lie on a hole,

1) it generates two random points A' and A from $\Delta^2 Sq(h)$ boxes 1 and 2 respectively, and sets NEXT-DEST =the intersection of AA' and the line joining source and destination, sets STAGE = 1, SEC-DEST = A and greedily forwards towards NEXT-DEST.
2) It also updates the TOPOLOGY DATA field to provide the $xmin(h), ymax(h)$ of the hole h that it is bordering.

This provides the nodes the information about the holes' dimensions to compute random points from appropriate boxes. Note that the SEC-DEST is modified only by a node that is on the boundary of a hole. **End of Algorithm**

Calculation of the Hole's dimensions:

A node on the hole perimeter (at location (x, y) receiving a packet with the topology flag set (i.e., signifying that it is a topology packet) computes $xmin(h) = min(xmin(h), x)$, $ymax = max(ymax(h), y)$ and passes it to the clock-wise closest neighbor that is on the hole boundary. A periodic update of such messages, along with their respective timeout mechanisms can be used to generate a knowledge of hole dimensions at the boundaries.

More informally, our algorithm constructs a random path (as shown in Figure 2) in the annular region around the hole, and then continues on in its straight-line path once it leaves the $(1+\Delta)Sq(h)$ region around the hole. In our algorithm, we choose the hole traversal algorithm to go above the hole for analytical simplicity. In practice one can randomize this choice (to go above or below) to perform better load balancing although the results would be order-wise the same. We note that the above algorithm can be either used to initialize static routes that can be remembered, or each packet can be independently routed.

For the following analysis, we assume that RandHT is run to setup static routes.

B. Analysis of RandHT(n) Algorithm

In this section, we provide a quantitative analysis of the throughput-capacity achievable in networks with holes (as defined in Section II-A) and random source-destination pairs. Before we begin our analysis, we show the following upper bound on the best achievable throughput capacity. In this section, we skip the proofs of the claims and refer to [20] for details.

Claim 1: In networks with holes and a random distribution of source-destination pairs the best achievable throughput-capacity $T(n) = O(\frac{1}{\sqrt{n}})$.

Theorem 3.1: Consider networks as defined in Section II-A. The simultaneously achievable throughput capacity $T(n) = \tilde{\Theta}(\frac{1}{\sqrt{n}})$. Further, the delay $D(n) = \tilde{\Theta}(nT(n))$.

Thus, we show that our routing scheme achieves near-optimal throughput & delay (at the maximum capacity), and the routing information at nodes does not grow significantly.

Proof: We shall make use of the following result (whose proof is similar to Lemma 4.13 of [8] and is skipped for brevity).

Result 3.1: Consider a torus of dimensions $n^{\gamma-1/2}$, with $0 < \gamma < 1/2$. We pick $Rn^{\gamma}\log n$ random source destination pairs and connect them with straight-lines. Then, in each tile of size $M(n) \times M(n)$, there are $O(\sqrt{n}\log n)$ lines through any tile, with high probability.

We begin by considering a tiling of the unit torus by square tiles of the size $M(n) \times M(n)$ and showing that the number of lines through any arbitrary tile chosen from the tiling is $\tilde{\Theta}(\sqrt{n})$ $w.h.p$. Note that a route may pass through the same tile more than once - each time using a different straight-line path. Then, based on standard coloring arguments in [8], [2], we can show that the constant bandwidth available at a tile can be uniformly split among all *lines* passing through it to provide a throughput $T(n) = \tilde{\Theta}(\frac{1}{\sqrt{n}})$ for all *routes*. There are three kinds of tiles: (i) Tiles that lie outside the $(1+\Delta)\,Sq(h)$ of all holes h, (ii) Tiles that lie within $Sq(h)$ for some h, (iii) Tiles that lie in the annular region $(1+\Delta)\,Sq(h) - Sq(h)$ of some h. We show the above bound for each of these possibilities.

CASE 1: If a tile is outside the annular region, the tile is exactly equivalent to a tile in a network without holes where $\Theta(n)$ random source-destination pairs are chosen. This is due to the fact that outside the annular regions, the number of lines that go through a tile is unchanged if the routing were according to our scheme or my a direct straight-line path - i.e., our stage 0 routes and the straight-line paths from source to destination are exactly the same on regions outside the region $(1+\Delta)\,Sq(h)$ of any h. From standard results on throwing $n/2$ random lines due to random choice of source-destination pairs on a unit torus (Lemma 4.13 of [8] or Claim 2 of [19]), we know that the maximally loaded tile is at most $\tilde{\Theta}(\sqrt{n})$ with probability $1 - \frac{1}{n^2}$.

CASE 2: If a tile is inside the square region $Sq(h)$, for some hole region h, it is clear that packets of only 2 stages pass through it. The stage 0 lines may pass through a tile in this region if a randomly chosen destination is on the other side of the hole. If a stage 0 packet hits a hole, it leaves the region by using the reverse path (Stage 1 packet) to a random point C in the annular region (Figure 2). Thus, for every stage 0 packet through a tile, there is at most one stage 1 packet passing through it. Since the stage 0 of any route is an exact subset of the straight-line between the random source-destination pair (stage 1 is a subset as well, but with flows in the opposite direction), the total load on a tile in the $Sq(h)$ region is again upper bounded by $\tilde{\Theta}(\sqrt{n})$.

CASE 3: Note that all stages of packets may pass through the annular region. But as the traffic due to stages 0 and 1 have been shown to be $\tilde{\Theta}(\sqrt{n})$, w.h.p, we restrict our attention to Stage 3 of any route. This is because, stage 2 routes are subsets of AA' and stage 4 routes are subsets of BB' and both AA' and BB' are symmetric to AB (in the sense that their distribution is identical to AB over the corresponding rectangular arm - Region (Stage 3)). Thus, if we show the load due to stage 3 of routes is no more than $\tilde{\Theta}(\sqrt{n})$ with probability $1 - \frac{1}{n^2}$, our claim on the achievable throughput capacity follows. First, we show a bound on the number of stage 3 routes that are generated for any hole h, and let $|h|$ be the side of the smallest square containing the hole.

Claim 2: The number of stage 3 routes around any hole h is $O(n|h|\log n)$.

Now, we consider a tile in the rectangular region where stage 3 routes are active (See Figure 2). The distribution of stage 3 routes over this region is not uniform for standard bounds to apply. We upper bound this system by the following uniform system.

Consider a toroidal region T_{bound} of side $2(1+\Delta)|h|$. In this region we throw $2(1+\Delta)^2|h| \times Kn(\log n)^2$ (we choose a sufficiently large K) random source-destination pairs. Noticing that this network is a smaller analog of the uniform network considered in the proof of Lemma 4.13 of [8], we apply our standard bounds on uniform networks to show the following claim.

Claim 3: The number of lines through any tile is no more than $\Theta(\sqrt{n}(\log n)^2)$ with probability at least $1 - \frac{1}{n^2}$.
In this toroidal region, we pick two boxes B_2, B_3 of size $(\Delta|h|)^2$ that are a distance $|h|$ apart from each other, i.e., a region similar to Region(stage 3) in Figure 2. We show that the number of source destination pairs such that the source lies in box B_2 and destination in box B_3 is greater than the number of stage 3 routes of the original network, and further as each of these lines are independently and identically distributed as the line segment AB. Now, as we throw $2(1+\Delta)^2|h| \times Kn\log n$ over $(1+1/\Delta)^2$ tiles of size $(\Delta|h|)^2$, the number of sources over box 2 is at least $\Theta(n|h|(\log n)^2)$ with probability $1 - \frac{1}{n^2}$. Further, each of these sources picks a random destination. We count the number of destinations that would fall in box 3. As we throw $\Theta(n|h|(\log n)^2)$ over $(1+1/\Delta)^2$ boxes, there exist at least $\Theta(n|h|(\log n))$ source-destination pairs that have a random source in box B_2 and a random destination in box B_3 (this is with probability at least $1 - \frac{1}{n^2}$). Let L be the number

of lines over a tile of size $M(n) \times M(n)$ in the Region (Stage 3), and let L^* be the number of lines passing through any $M(n) \times M(n)$ tile in toroidal region T_{bound}. Then, we can show that $\mathbb{P}(L > \Theta(\sqrt{n}(\log n)^2)) \leq 2/n^2$.

By our scheduling algorithm where each tile of size $M(n) \times M(n)$ can be allocated a constant fraction of a time-slot for collision-free transmissions (the interference graph is a finite degree graph that can be colored with finite colors [2]) it follows that every line through a tile can be provided an equal rate of $\tilde{\Theta}(\frac{1}{\sqrt{n}})$ thus providing the same throughput to all routes in the network. Further, the delay $D(n)$ is the sum of the time spent by a packet in each hop. Note that the number of hops is at most $3 \times dist(S-D)$, and the delay at each hop due to scheduling is no more than $\tilde{\Theta}(\sqrt{n})$. Thus, delays are no more than $\frac{3}{M(n)} \times \tilde{\Theta}(\sqrt{n}) = \tilde{\Theta}(nT(n))$. Note that this lies on the optimal throughput delay curve [4], [16]. ■

C. Scaling of Routing Information

A main motivation of geographic routing schemes is the minimal amount of routing information that each node has to store. Here, we discuss the scaling of routing information of our algorithm. The RandHT(n) algorithm can be used in two ways: (i) The route for each packet to its destination was setup independently and randomly according to RandHT, or (ii) the RandHT algorithm is run once initially to setup static routes (i.e. all packets from a $S-D$ pair follow the same route).

In case (i), the only routing information needed at any node is the locations of the neighboring nodes, which grows as $\Theta(\log n)$. This is due to the fact that the waypoint nodes are not required to remember the next waypoint, but generate it randomly, from the information available in the packet. In case (ii), the waypoint nodes are required to remember the next waypoint so that the packets are routed along the static routes. However, we note that a maximum of $n^{1/2-\gamma}$ holes can occur on the path between a source and its destination, and thus each path may have at least $3 \times n^{1/2-\gamma}$ waypoints, and with n routes, this implies that each node is a waypoint for $n^{1/2-\gamma}$ routes on an average.

We note that while our analysis for the throughput assumed static-routes for tractability, we strongly believe that the throughput achieved by per-packet routes would be order-wise unchanged.

IV. Networks with Arbitrary Traffic Patterns

In this section, we consider the problem of routing between arbitrarily chosen source-destination pairs, with arbitrary traffic demands. Thus, we consider a fairly general network and traffic model (cf. Section II-B for a description of the model). A critical issue is to determine if some form of randomized geographic routing can provide near-optimal throughput. Such a routing scheme would provide highly distributed networks (with low computational capabilities) to achieve high data rates without any route setup overheads. Also, geographic routing would converge immediately to the near-optimal routes.

In previous work [19], we studied networks with randomly chosen source-destination pairs (such that the source-destination pairs are $\Theta(1)$ distance away from each other, which corresponds to $\alpha = 1/2$) with a two-level traffic demand, and demonstrated a randomized routing algorithm RANDOMSPREAD that was near-optimal. Here, we generalize the model to allow arbitrary locations of source and destination (cf. Section II-B).

Typically, upper bounds on network capacity have utilized cut-set ideas to limit the traffic that can leave any set [12]. Essentially, if we consider any closed region of space, the amount of traffic that can enter or leave this area is bounded by the amount of radio resource along the boundary of the set. However, when sources and sinks can be arbitrarily close or far, and with widely varying traffic requirements, the cut-set bound alone is not sufficient to characterize network load distributions. (The traffic flows that never leave the region are unaccounted by such cuts.) In this paper, we jointly utilize transport capacity bounds (bounds that arise due the interference nature of the channel) along with cut-set based bounds to characterize the allowable set of source-destination pairs.

The joint approach is based on the following reasoning. Every sub-region of the geographic region contains two "resources": *(i)* the transport capacity of the sub-region, and *(ii)* the amount of traffic that can enter/leave the sub-region through its boundary (the perimeter cut-capacity). For each S-D pair, any routing algorithm "uses up" some amount of each of the two "resources". For instance, if the S-D pair lies completely within a sub-region, straight line routing uses up only the transport capacity within the region. On the other hand, if the S-D pair decides to route by spreading the load over the entire geographic region, it will use the perimeter cut-capacity of the sub-region along with some amount of the transport capacity of the sub-region. We demonstrate that any routing scheme has a local conservation property between these two resources, namely, that using less amount of the local transport capacity resource, requires it to use more of the local cut-capacity resource, and vice-versa.

We propose an algorithm RandLLB (Randomized Local Load Balancing) and demonstrate using the above property that it is 'near-optimal' for arbitrary traffic demands (a finite-level traffic model is considered for analytical tractability - this can be readily extended to an arbitrary traffic model). We describe the algorithm below.

Algorithm RandLLB(n) Consider a source-destination pair $l \in \{1, \cdots, H\}$ demanding a rate $r \in \mathcal{R}$, and whose destination is $n^{\alpha-1/2}$ away from its source [1] (with $0 < \alpha^* < \alpha \leq 1/2$).

1) The source node chooses n^α locations at random from within a circle of radius $n^{\alpha-1/2}$ about the source location for its first waypoint $S'(i)$, for $1 \leq i \leq n^\alpha$.
2) The source node then chooses n^α locations at random from within a circle of radius $n^{\alpha-1/2}$ about the destination location for its second waypoint $D'(i)$.

[1] For notational simplicity, we suppress the source-destination index l in α (i.e. α_l) with the understanding in the proof that each source-destination pair has potentially a different α_l.

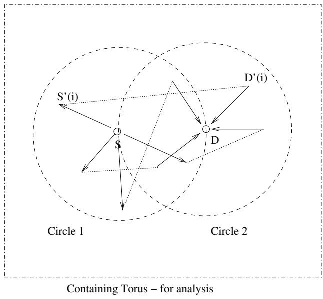

Fig. 3. Local load balancing with three-hop routing.

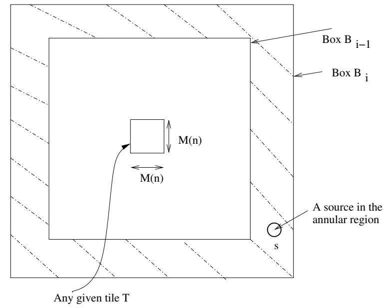

Fig. 4. The traffic load through any arbitrary tile.

3) Thus, each source constructs n^α paths from itself to the destination, and randomly distributes the traffic load over a region that is proportional to the square of the distance between the source and destination.
4) The source splits its rate uniformly over the n^α multi-path routes.

For an illustration of this process, see Figure 3. **End of Algorithm**

A. Analysis of RandLLB(n) Algorithm

We show that the above algorithm spreads the local traffic load in an appropriate manner such that the traffic through any tile is manageable for any configuration of source-destination pairs and their traffic demands if they are achievable by any other scheme. That is, we show the following theorem.

Theorem 4.1: Consider an arbitrary distribution of source-destination pairs over a random planar network (see Section II-B). Let a rate vector $\Lambda = [\lambda_{s,d} \in \mathcal{R}]$ be achievable by any scheme. Then, the RandLLB(n) achieves a throughput rate of $\tilde{\Theta}(\Lambda)$, i.e, the algorithm is throughput-optimal up to poly-logarithmic factors.

Proof: The main steps in the proof are as follows:

(i) We assume all the sources require a rate $r \in \mathcal{R}$.
(ii) We develop two basic spatial constraints (cut-set and transport capacity based bounds) on the positions of the S-D pairs that are necessary for all routing schemes.
(iii) We construct a bound on the total traffic that may pass through any given tile, given the constraints on the positions of S-D pairs.
(iv) We show that the traffic through any tile (of the size of the radio range) is no more than $\Theta(\log(n))$ for any achievable S-D pair configuration and since each tile of the radio range is capable of supporting a constant traffic (to its neighboring tiles), we can scale down the throughput of all sources by a log-factor to achieve near-optimal throughput.
(v) We repeat the argument for each traffic level in the finite set $\mathcal{R}$. [2]

[2]Alternately, we can sharpen this bound by considering all rate requirements that are a multiple of a basic rate $\hat{r}$. By showing achievability for this basic rate $\hat{r}$ with arbitrary S- D pairs, we can immediately generalize to rate requirements that are a multiple $\hat{r}$. This is because any multiple of $\hat{r}$ can be viewed as a group of sources (destinations) that are co-located in the same tile. However, we skip a formal proof of this due to space constraints.

Thus, we first assume that all S-D pairs require a rate r . Note that there are three types of traffic - the initial outward-star traffic, i.e., the routes between S and $S'_i, (1 < i < n^\alpha)$ ($*$-traffic), the traffic between the 1st and the 2nd waypoints of each route (routes between S'_i and D'_i for $1 < i < n^\alpha$) or (#-traffic), and the final inward-star traffic. By symmetry, the traffic load seen due to the inward-star traffic is same as the $*$-traffic. Consider any given tile (as in Figure 4) of size $M(n) \times M(n)$, and construct concentric squares B_i

1) Necessary conditions on source-destination pairs: The following are necessary conditions for any routing scheme:

Condition 4.1: (i) Traffic bound I (transport capacity): Since any tile can at most support a rate '1', the total traffic supported inside any box B_i is at most $(2i-1)^2$. (ii) Traffic Bound II (perimeter or cut-set): The total traffic leaving (or entering) box B_i is at most $4 \times (2i-1)$.

Now, consider a source whose destination is $n^{\alpha-1/2}$ away, or equivalently, for an $M(n) \times M(n)$ tiling of the space, $n^\alpha/\sqrt{\log n}$ boxes away. Let this source be in the square region B_i. Then the following holds:

(i) $Dist(S-D) < \frac{i}{3}$ boxes: In this case, we can show that an upper bound on the number of sources that can affect tile T within Box B_i and with distance to destination less than $i/3$ cannot be more than $32i/r$ (details in [20]).
(ii) $Dist(S-D) > 2\sqrt{2}i$ boxes: The destination lies outside the box B_i, and hence it uses up r units of capacity from the perimeter bound (allowable $4 \times (2i-1)$). Thus the total number of such sources is upper bounded by $4(2i-1) \times \frac{1}{r}$.
(iii) $Dist(S-D)$ is between $\frac{i}{3}$ and $2\sqrt{2}i$ boxes: In this case, the destination can either lie inside or outside the box. If the destination was outside, the source uses up r units from the allowable perimeter capacity of $4(2i-1)$. If the destination was inside, an arbitrary part r_{in} is supported completely inside the box, and $r - r_{in}$ leaves the box. Note that this is true for every routing scheme. Then the rate inside the box uses up at least $r_{in} \times \frac{i}{3}$ of the allowable transport capacity $(2i-1)^2$. The rate outside the box uses up $r - r_{in}$ of the perimeter bound $4 \times (2i-1)$. Thus the total number of such sources is upper bounded by $\left(\frac{(2i-1)^2}{i/3} + 4(2i-1)\right)\frac{1}{r}$.

Thus, a uniform bound on the number of sources in region B_i that can affect tile T is given by $\frac{32}{r}(2i-1)$, which is greater than $\left(\frac{(2i-1)^2}{i/3} + 4(2i-1)\right)\frac{1}{r} + 32\frac{i}{r}$.

2) Traffic through a tile due to a source: We now provide a bound on the traffic through tile T due to a source s in $B_i - B_{i-1}$. The bounds arise from two kinds of traffic:

The $*$-traffic: Note that the outward-star traffic is generated by choosing n^α points at random from a circle of radius $n^{\alpha-1/2}$ centered at the source node. If the source-destination separation was less than $i/3$ boxes, the $*$-traffic does not touch tile T. Else, the number of lines that can go through tile T (say $lines(T,s)$),

$$lines(T,s) \leq \frac{K\log n}{2i-1} \times n^\alpha$$

with probability $1 - \frac{1}{n^4}$ for some $K < \infty$. (The above bound can be obtained from standard results on throwing n^α lines randomly at $4(2i-1)$ boxes, when $4(2i-1) = O(n^\alpha)$). Thus, the traffic through a tile due to one source is $r.n^{-\alpha} \times lines(T,s) = r \times \frac{K\log n}{2i-1}$ with probability $1 - \frac{1}{n^4}$.

The $\#$-traffic: Again, if the source-destination separation was less than $i/3$ boxes, the $\#$-traffic does not touch tile T. If greater, note that the traffic is generated by picking n^α random lines that have a source in Circle 1 and destination in Circle 2 (of Figure 3).

Claim 4: The number of $\#$ lines through any given tile is $O(\log n)$ with probability at least $1 - \frac{1}{n^4}$.

Thus, the traffic through any "touchable" tile is (number of lines) $\times$ (traffic through each line). Since the rate of r was split uniformly among n^α routes, the $\#$ traffic through any tile is upper bounded by $K_1 \log n \times rn^{-\alpha}$. As n^α is at-least $i/3\sqrt{\log n}$ for any source whose $\#$ traffic can touch tile T, the total traffic

$$K_1 \log n \times rn^{-\alpha} \leq K_1 r \log n \frac{1}{2i-1}.$$

We note that the inward-star traffic is symmetric to the $*$-traffic.

3) Maximizing the traffic through any tile: Previously, we characterized the necessary conditions on the number of sources in any $B_i - B_{i-1}$ annular region, and also provided an upper bound (that holds with high probability) on the load seen on a tile due to any source in $B_i - B_{i-1}$. Thus, to demonstrate that our algorithm does not overload any tile, we maximize the traffic on any tile given the constraints on the source-destination pairs, and show that the maximum traffic is $\tilde{\Theta}(\log n)$, i.e., is near-optimal.

Let a_i be the number of sources in region $B_i - B_{i-1}$ that can affect tile T. Then, $\sum_{l=1}^{i} a_l \leq \frac{32}{r}(2i-1)$ for all $1 \leq i \leq \sqrt{n/\log n}$. Hence, the maximum traffic through a tile is upper bounded by the solution to the following optimization problem. Let $K_2 = K_1 + K$.

$$\max \sum_{i=1}^{\sqrt{n/\log n}} \left(a_i \times K_2 \frac{r\log n}{2i-1}\right) \quad \text{s.t} \qquad (1)$$

$$\sum_{l=1}^{i} a_l \leq \frac{32}{r}(2i-1), \quad \forall\{1 \leq i \leq \sqrt{n/\log n}\}.$$

Claim 5: $a^* = \frac{32}{r}[1,2,2,\cdots,2]$ maximizes the above problem.
The proof is available in [20].

The traffic through any tile is at most

$$\frac{10}{r} r K_2 \log n \sum_{i=1}^{\sqrt{n/\log n}} \frac{2}{2i-1} = \qquad \Theta(\log^2(n)). \qquad (2)$$

Thus, we show that for any allowable source-destination configuration, the traffic through any tile is at most $\Theta(\log^2(n))$. By means of a finite-coloring scheme for the tiles, we can provide a constant throughput for each tile, and hence, by reducing the throughput of each source by a poly-logarithmic factor, we can support $\tilde{\Theta}(\Lambda)$. We showed the above proof for a given rate r - the method can be similarly used for other rates r^* from $\mathcal{R}$. ■

V. Discussion and Conclusions

In the previous sections, we formally demonstrated that randomized geographic schemes can obtain near-optimal throughput performance, with low complexity and very little coordination. Here, we try to address some issues that may arise in practical networks.

A. Networks with Traffic and Node Non-uniformity

A key property of the RandLLB algorithm that allows it to achieve optimality is that a source-destination pair that is separated by a distance d spreads its traffic only over tiles that are of the same distance from either of them. However, in networks with holes, it is possible that the shortest path between a source and its destination is much larger than the Euclidean distance between them. In such situations, combining the RandLLB algorithm with the RandHT algorithm may be sub-optimal, as the hole traversing algorithm introduces a traffic demand of rate r over a region of much larger size than the source-destination separation. Consider networks with the following property (in addition to Section II).

Condition 5.1: Let $d(x,y)$ be the Euclidean distance between nodes x and y. Then, the shortest-distance path between nodes x and y in the network $\cong d_N(x,y) \leq K_3 d(x,y)\,\forall(x,y)$. For such networks, we propose the following scheme.

1) Each source x performs RandLLB(n) on its traffic to its destination y, spreading over an area $(2K_3 d(x,y))^2$.
2) A packet on hitting a hole h's boundary checks if the $|h|$ is greater than d(SRC-LOC, FINAL-DEST).
3) If greater, the packet is dropped at the hole boundary.
4) If the hole is smaller, it performs a RandHT(n) to traverse the hole.

Note that the above scheme has the following properties: (i) A source-destination pair (x, y) only loads tiles that are within $\Theta(d(x, y))$,and (ii) The shortest path has a tubular region of width at least $\Delta|h|/2$ around it that is not affected by holes. We expect that the above scheme "optimally" combines the two algorithms proposed in the paper, namely, RandHT and RandLLB. We plan to provide a formal proof of the above claim in a future work.

B. Networks with Arbitrarily Connected Graphs

While in many practical scenarios of ad-hoc wireless networks we may model the non-uniformity of the network as occurrence of holes, the actual network topology can be fairly complex and not satisfy the hole assumption in Section II-A. In such cases, constructing greedy routes in a throughput-optimal manner may require much more complex algorithms. Moreover, algorithms such as RandHT(n) may fail to construct a path to the destination in such complex networks.

Although GPSR-like algorithms provide low throughput even with minimal network non-uniformity, they are capable of constructing a path to the destination if one exists (however, with poor load-balancing). To overcome such pathological networks, practical algorithms could combine the randomized algorithms proposed in this paper along with deterministic GPSR-like algorithms to provide worst-case performance guarantees. For example, they could be combined in the following manner:

1) Each source tries to construct both a GPSR based route (green) and Randomized route (red) to the destination.
2) In each tile of size $M(n) \times M(n)$, the channel access time available at each tile is divided into a fraction β for randomized schemes and a fraction $1 - \beta$ for GPSR-like schemes.
3) Based on the fraction of red and green packets received at the destination, the β-factor can be updated (by some gossip mechanism) to utilize the more efficient of the two schemes.

This assures that if the network is complex, GPSR-like schemes guarantee a path to the destination, while if the network has "manageable" holes, the randomized algorithms provide much better throughput.

C. Practical Issues

We wish to emphasize here that the focus of these algorithms is on providing a near-optimal performance - there are some implementation issues that may arise in practical protocols:

1) Identification of hole perimeter - In our algorithms, we had assumed that the nodes have knowledge of whether they are on the hole boundary or not. In practice, techniques explored in [5] may be used by the nodes to learn of their membership on a hole-perimeter.
2) Stability of topology updates - With a changing topology, where nodes may move in and out of holes and hole shapes could change significantly over time, an important issue is if the hole update mechanisms can still provide a good path to the destination. As a part of future work, we will investigate such effects of node mobility.

VI. Acknowledgments

This research was supported in part by NSF Grants CNS-0325788, CNS-0347400, CNS-0519535, CCR-0325673, and Darpa CBMANET and ITMANET programs.

References

[1] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, Englewood Cliffs, NJ, 1987.

[2] J.A. Bondy and U.S.R. Murty. *Graph Theory with Applications*. Macmilliam Press, 1976.

[3] S. Douglas, J. De Couto, and R. Morris. Location proxies and intermediate node forwarding for practical geographic forwarding. *Technical Report MIT-LCS-TR-824, MIT Laboratory for Computer Science*, 2001.

[4] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput delay trade-off in wireless networks. In *Proc. of IEEE Infocom*, March 2004.

[5] Q. Fang, J. Gao, L. J. Guibas, V. deSilva, and L. Zhang. Glider: Gradient landmark-based distributed routing for sensor networks. In *Proc. of IEEE Infocom*, March, 2005.

[6] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran. Closing the gap in the capacity of random wireless networks. In *Proceedings of Intl Symp. on Info. Theory*, 2004.

[7] P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity in wireless networks. In *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming. Edited by W.M. McEneany, G. Yin, and Q. Zhang*, pages 547–566, Boston, 1998.

[8] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Trans. on Info. Theory*, IT-46(2):388–404, March 2000.

[9] B. Karp and H. T. Kung. GPSR: Greedy perimeter stateless routing for wireless networks. In *Proc. of the ACM/IEEE International Conf. on Mobile Comp. and Networking*, page 243, August 2000.

[10] E. Kranakis, H. Singh, and J. Urrutia. Compass routing on geometric networks. In *Proc. of the 11th Canadian Conf. on Comput. Geom.*, August 1999.

[11] F. Kuhn, R. Wattenhofer, and A. Zollinger. Worst-case optimal and average-case efficient geometric ad-hoc routing. In *Proc. of ACM MobiHoc*, 2003.

[12] R. Madan and D. Shah. Capacity-delay scaling in arbitrary wireless networks. *Allerton Conference on Comm., Cont., and Comp.*, 2005.

[13] M. J. Neely and E. Modiano. Capacity and delay tradeoffs for ad hoc mobile networks. *IEEE Transactions on Information Theory*, 51(6):1917–1937, 2005.

[14] C. Peraki and S. D. Servetto. Capacity, stability and flows in large scale random networks. In *Proc. IEEE Information Theory Workshop*, 2004.

[15] E. Royer and C. Toh. A review of current routing protocols for ad-hoc mobile wireless networks. *IEEE Personal Communications*, Apr 1999.

[16] G. Sharma, R. Mazumdar, and N. Shroff. Delay-capacity tradeoffs in mobile ad-hoc networks: A global perspective. In *Proc. of IEEE Infocom*, 2006.

[17] I. Stojmenovic and X. Lin. Gedir: Loop-free location based routing in wireless networks. *Proc. IASTED Int'l Conf. Parallel and Distributed Computing and Systems*, pages 1025–1028, 1999.

[18] A. L. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Syst.*, 50(4):401–457, 2005.

[19] S. Subramanian, S. Shakkottai, and P. Gupta. On optimal geographic routing for networks with holes and non-uniform traffic. In *Proceedings of IEEE Infocom*, 2007.

[20] S. Subramanian, S. Shakkottai, and P. Gupta. Optimal geographic routing for wireless networks with near-arbitrary holes and traffic. Bell Labs Technical Report, Jul 2007.

[21] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling for maximum throughput in multihop radio networks. *IEEE Trans. on Auto. Control*, 37(12):1936–1949, December 1992.

[22] L. G. Valiant and G. J. Brebner. Universal schemes for parallel communication. In *Proc. of the 13th annual ACM symp. on Theory of computing*, pages 263–277, New York, NY, USA, 1981. ACM Press.