Spatial Distribution in Routing Table Design for Sensor Networks

Rik Sarkar Xianjin Zhu Jie Gao

Department of Computer Science, Stony Brook University. {rik, xjzhu, jgao}@cs.sunysb.edu

***Abstract*—We propose a generic routing table design principle for scalable routing on networks with bounded metric growth. Given an inaccurate distance oracle that estimates the graph distance of any two nodes up to constant factor upper and lower bound, we augment it by storing the routing paths of pairs of nodes, selected in a spatial distribution, and show that the routing table enables $1+\varepsilon$ stretch routing. In the wireless ad hoc and sensor network scenario, the geographic locations of the nodes serve as such an inaccurate distance oracle. Each node p selects $O(\log n \operatorname{loglog} n)$ other nodes from a distribution proportional to $1/r^2$ where r is the distance to p and the routing paths to these nodes are stored along the paths in the network. The routing algorithm selects links conforming to a set of sufficient conditions and guarantees with high probability $1+\varepsilon$ stretch routing with routing table size $O(\sqrt{n}\log n \operatorname{loglog} n)$ on average for each node. This scheme is favorable for its simplicity, generality and blindness to any global state. It is a good example that global routing properties emerge from purely distributed and uncoordinated routing table design.**

I. INTRODUCTION

The problem we study in this paper is a fundamental one: how to exploit the geometric properties in a sensor network for routing table design, in particular, what routing paths should be kept in the routing table, such that the average routing table size is small, the path stretch is close to optimal (1+ε for any given $\varepsilon > 0$), and both the preprocessing and the routing can be achieved by the nodes making decisions on their own, blind to any global state?

More precisely, given an inaccurate distance oracle $\mathcal{O}$, how to design compact routing tables to help deliver messages in a large scale sensor network.

For an example, in the sensor network setting, one can use the Euclidean distance to approximate the hop count distance of two nodes in the network. This is a reasonable model that does not assume unit disk graphs. It also allows for inaccurate localization.

We represent the oracle's distance estimate between nodes p and q by $d(p,q)$, and the true but possibly unknown graph distance between p and q by $\sigma(p,q)$. The oracle is then modeled by two positive constants δ_1 and δ_2 and the guarantee $\delta_1 d(p,q) \leq \sigma(p,q) \leq \delta_2 d(p,q)$.

With any inaccurate distance oracle, the solution we propose for routing is to augment in the routing table paths between pairs of nodes that are not immediate neighbors, these paths are called virtual links, or *long links*. In particular, for some selected pairs (u,v) a path between u, v, $P(u,v)$, is recorded in the routing table of all nodes on this path. Based on a set of sufficient conditions, we define a *forwarding region* (see Fig. 1) from which p selects the next hop in the path. If the selected node x is a neighbor through a long link, then the routing information stored on the path $P(p,x)$ is used to deliver the message to x. Node x then repeats an identical procedure to advance the message. Now the question is, what long links should each node build and which node should be selected in the routing stage, without knowing the global state, such that the routing table size is small, the path stretch is low, and delivery rate is high?

II. ROUTING WITH SPATIAL DISTRIBUTIONS

In this section we describe the idea of using spatial distribution to route with $1+\varepsilon$ stretch in a suitable metric space $\mathcal{M}$. We assume that a node is able to get the approximate distance $d(p,q)$ from just the names of p, q. The implementation of this distance oracle is beyond the scope of this short abstract. When the long links are carefully chosen the delivery rate is high and the routing stretch is low. Due to lack of space, we can only present the core intuitions and theorems here. A full version containing all proofs is available online [5].

Accurate distance oracle. To demonstrate the basic concept, we first consider the case in which the oracle is in fact accurate, that is, $d = \sigma$. The objective is to recursively build a route from s to t with the help of the long links. Suppose s takes a long link to node p, then we want $\sigma(s,p) + \sigma(p,t)$ to be not very large compared to $\sigma(s,t)$:

$$\sigma(s,p) + \sigma(p,t) \leq \gamma \cdot \sigma(s,t), \tag{1}$$

Where $\gamma \geq 1$ is a parameter depending on ε. Observe that inequality (1) defines an ellipse in $\mathbb{R}^2$ with s and t at foci. Now we impose an additional restriction that moving from s to p implies a certain progress in direction of t. In particular, p is closer to t by a factor of at least $0 \leq \beta \leq 1$:

$$\sigma(p,t) \leq \beta \cdot \sigma(s,t). \tag{2}$$

This describes a disk centered at t.

Next, we select γ and β such that the selection procedure enforced by inequalities (1) and (2) when applied recursively, produces a path of stretch at most $1+\varepsilon$:

$$R(s,t) \leq (1+\varepsilon) \cdot \sigma(s,t), \tag{3}$$

where R gives the length of the path created recursively.

A *forwarding region* $F_\varepsilon(s,t)$ is a set of points p in $\mathcal{M}$ from which s can select p satisfying the relations above. The following lemma gives a precise relation:

Lemma 2.1. *Values of γ and β satisfying $\gamma + \varepsilon\beta \leq 1+\varepsilon$ constitute the forwarding region, with the equality corresponding to the region boundary.*

It is easy to see that γ must lie in the interval $[1, \frac{2+3\varepsilon}{2+\varepsilon}]$ for a given ε. For each value of γ, we have a region $H_{\gamma,\varepsilon}(s,t) \subseteq \mathcal{M}$ which is the intersection of the ellipse bounded region and the disk. Thus, formally, the forwarding region is the union: $F_\varepsilon(s,t) = \cup_\gamma H_{\gamma,\varepsilon}(s,t)$. See Figure 1.

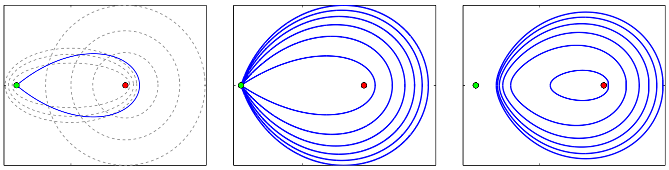

Fig. 1. (i) Boundary of F_ε as intersection of ellipses and circles. (ii) Forwarding regions for different values of ε from 0.2 to 2. (iii) Forwarding regions for different values of ε from 0.2 to 2 for approximate oracle. Observe that in this case the forwarding regions are smaller and source s is not in the forwarding region. This is due inaccurate distance estimates and necessitates the use of *long links* - without which s cannot access the forwarding region.

Approximate distance oracle. For approximate distance oracles, it would be sufficient to guarantee the following inequalities (corresponding to relations (1)-(2) respectively):

$$\begin{aligned} \delta_2 d(s,p) + \delta_2 d(p,t) &\leq \gamma \delta_1 d(s,t) \\ \delta_2 d(p,t) &\leq \beta \delta_1 d(s,t) \end{aligned} \quad (4)$$

It can be verified that lemma 2.1, with the same boundary condition holds here as well. Figure 1 shows the forwarding regions for the euclidean case, with accurate and inaccurate oracles. Note that the shape of the forwarding regions is independent of the distance between the two nodes in question.

In a sensor graph setting, we use the (upper and lower) bounded growth rate model. If we place at most a constant number of sensor nodes inside any unit disk and the holes in the sensor networks are not very fragmenting, the number of nodes at k hops from a node p will be around $\Theta(k)$.

In general, we consider a graph such that number of nodes in an r neighborhood $|N_r(p)| = \Theta(r^\rho)$. Note that the diameter D of such a graph is bounded by $\Theta(n^{1/\rho})$.

A. Routing table construction

To build the routing table, we use a spatial distribution (see [3]) of directed links. In particular, for nodes p and q separated by a distance r, the probability of a directed link pq being built is proportional to $1/r^\rho$.

Theorem 2.2. *From each node it is sufficient to select* $O\left(\left(\frac{2}{\varepsilon}\right)^{O(\rho)} \ln n \ln \ln n\right)$ *links, to guarantee a link in the forwarding region for every possible destination with probability* $1 - 1/\log^{\Theta(1)} n$.

The theorem above describes a guarantee for a suitable link to a forwarding region to exist. However, we still need to prove the existence of a path of $(1+\varepsilon)$ stretch for a given routing request, that will take us to within a small constant distance of the destination. This can in fact be done :

Theorem 2.3. *It is sufficient to select* $O\left(\left(\frac{2}{\varepsilon}\right)^{O(\rho)} \ln n \ln \ln n\right)$ *long links per node to guarantee a path of stretch at most* $1+\varepsilon$ *with probability at least* $1 - \frac{1}{\rho \log^{\Theta(1)} n}$.

And the routing table size is not too large.

Theorem 2.4. *The average routing table size of the scheme is bounded by* $O\left(\left(\frac{2}{\varepsilon}\right)^{O(\rho)} n^{1/\rho} \ln n \ln \ln n\right)$.

In the case of sensor networks in a plane ($\rho \approx 2$), for a given stretch ε, this amounts to a table size of $O\left(\sqrt{n} \ln n \ln \ln n\right)$ per node.

III. Implementation and Simulations

We simulated our scheme using Euclidean distance as a distance oracle, and compared simulations with VRR [1] and S4 [4] on different topologies with varying node density and hole distributions. We found that in networks of 1000-2000 nodes, the scheme achieves a delivery rate of 99% or more with $6-7$ long links per node. To achieve this delivery rate, it requires a smaller routing table than VRR and S4 in all cases. It achieves stretch equal or comparable to S4 and better than VRR in all cases.

For networks that are do not satisfy suitable deployment criteria, Euclidean distance is not a very good oracle. In such cases, it may be necessary to *build* an oracle in the network itself. We simulated an oracle based on random landmarks in the network. It can be shown that flooding from $O(\ln n \ln \ln n)$ landmarks suffice to select long links for each node with the correct spatial distribution. Also, the shortest path tree from the flood automatically creates the routes for the long links at no extra cost, implying very small routing tables per node.

In simulations, we used the centered distance metric described in [2]. The results show that the scheme achieves a very small stretch and good delivery rates with very small routing tables compared to other schemes.

IV. Conclusion

This scheme is not specific to sensor networks. It does not rely on graphs being unit-disk or quasi-unit-disk in a euclidean space. The concept of forwarding region applies to any metric space, while the bounds for number and storage costs of long links applies to any suitably growth bounded metric space. As such, these concepts can be applied to relatively large classes of graphs to produce self organizing distributed routing schemes with strong guarantees. The design of approximate distance oracles for the landmark scheme remains a challenge and will be addressed in future work on this topic.

References

[1] M. Caesar, M. Castro, E. B. Nightingale, G. O'Shea, and A. Rowstron. Virtual ring routing: network routing inspired by DHTs. In *SIGCOMM '06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 351–362, 2006.

[2] Q. Fang, J. Gao, L. Guibas, V. de Silva, and L. Zhang. GLIDER: Gradient landmark-based distributed routing for sensor networks. In *Proc. of the 24th Conference of the IEEE Communication Society (INFOCOM)*, volume 1, pages 339–350, March 2005.

[3] J. Kleinberg. The small-world phenomenon: an algorithm perspective. In *STOC '00: Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 163–170, 2000.

[4] Y. Mao, F. Wang, L. Qiu, S. S. Lam, and J. M. Smith. S4: Small state and small stretch routing protocol for large wireless sensor networks. In *Proceedings of the 4th USENIX Symposium on Networked System Design and Implementation (NSDI 2007)*, April 2007.

[5] R. Sarkar, X. Zhu, and J. Gao. Spatial distribution in routing table design for sensor networks. http://www.cs.sunysb.edu/∼rik/routing/spatial-routing-full.pdf, 2008.