# Regional Gossip Routing for Wireless Ad Hoc Networks

Xiang-Yang Li*     Kousha Moaveninejad*     Ophir Frieder*

## I. Introduction

One of the key challenges in *ad hoc* networks is the development of dynamic and efficient routing protocols. Some routing protocols use nodes' positions, while some protocols are based on flooding essentially. Gossip based routing method had been used and re-investigated [1] to reduce the number of messages: Every node will forward a message based on a uniform probability $p$. The gossiping exhibits a bi-modal behavior [2]. Then, there is a threshold value $p_0$ such that, in sufficiently large random networks, the gossip message quickly dies out if $p < p_0$ and the gossip message spreads to all network if $p > p_0$. So ideally, we would set the gossiping probability close to $p_0$, thus save about $1 - p_0$ message overhead compared with the flooding.

Gossiping still produces lots of unnecessary messages in regions that are far way from some line connecting the source to target nodes. We propose regional gossiping, in which essentially only nodes inside some region (derived from the source and target) will execute the gossiping protocol, and nodes outside the region will not participate in the gossiping at all. The region we select in our simulations are some ellipses using the source and target as foci. Our results show that, by using appropriate optimization heuristics, we can save up to $94\%$ messages even compared with the global gossiping method.

## II. Regional Gossip

Assume that wireless mobile hosts are a set $V$ of $n$ points distributed in a two-dimensional space, such as a unit area square (or disk) according to random uniform distribution $\mathcal{X}_n$, or Poisson process $\mathcal{P}_n$. Assume that all nodes have the same maximum transmission range equal to $r$. Let $G(V, r)$ be the graph defined on $V$ with edges $uv \in E$ if and only if $\|uv\| \leq r$. Let $\mathcal{G}_\Omega(\mathcal{X}_n, r_n)$ be the set of graphs $G(V, r_n)$ for $n$ nodes $V$ that are uniformly and independently distributed in a two-dimensional region $\Omega$. Gupta and Kumar [3] showed that a random graph in $\mathcal{G}_\Omega(\mathcal{X}_n, r_n)$ is asymptotically connected with probability one as $n$ goes infinity, when $\Omega$ is a unit disk.

In this paper, a *broadcast* by a node $u$ means $u$ sends the message to all nodes within its transmission range. The main communication cost in wireless networks is to send out the signal while the receiving cost is neglected here. We assume that the source node knows the target position, the global ellipse factor $\ell$, in addition to its own position through a low-cost GPS. In many applications such as data-centric sensor network, there is only a fixed number of sinks, which are often static, thus every node knows the positions of these possible target nodes. Otherwise, location service is needed for finding the location of the destination node. The geometry information of the source node and the destination node is piggybacked along with the message packet.

When a node, say $v$, receives a message, it retrieves the geometry position of the source $s$ node and the target node $t$. Node $v$ then checks if it is inside the ellipse defined by using $s$ and $t$ as foci, i.e., whether $\|vs\| + \|vt\| \leq \ell \cdot \|st\|$. When a node is not inside the ellipse, the node will just simply discard this message. Otherwise, with a fixed probability $p$, the node forwards this message to all nodes within its transmission range. Hereafter, we call $p$ the *relay probability* and $\ell$ the *ellipse factor* of our regional gossiping method. Obviously, the

Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616. Emails: xli@cs.iit.edu, moavkoo@iit.edu, ophir@cs.iit.edu.

probability that the destination node receives the message depends on the relay probability $p$, the ellipse factor $\ell$, the number of nodes $n$, and the transmission range $r$.

We then estimate the relay probability. Consider the network of $n$ nodes distributed in a square region with side length $a$. It was shown in [4] that if each node has transmission range $r_n$ and being off or fault with probability $p$, the network is connected with high probability if $p \cdot n\pi r_n^2 \simeq 2 \ln n$. Assume that the source and target distance is $d$ and the ellipse factor is $\ell$. The number of nodes inside the ellipse is then about $N_d = \frac{n}{a^2} \cdot \frac{\pi \ell \sqrt{\ell^2 - 1}}{4} d^2$. Since each node inside the ellipse forwards the message with probability $p$ after it receives the message, to let the target receive the message almost surely, the subnetwork composed of the nodes inside the ellipse with fault probability $p$ must be connected. The relay probability in our regional gossiping is thus at least $p \geq \frac{\ln N_d + c(N_d)}{N_d \cdot \pi (r/a)^2}$. Here $c(N_d)$ is a number going to $\infty$ when $N_d$ goes to $\infty$. The probability that the network (each node is chosen with probability $p$) is connected is $e^{-e^{-c(N_d)}}$. Substituting in $N_d$, we have

$$p \geq \frac{4a^4 \ln(n\pi\ell\sqrt{\ell^2 - 1}\, d^2 / (4a^2))}{\pi^2 d^2 r^2 \ell \sqrt{\ell^2 - 1} \cdot n} = \frac{\ln(n\pi\widetilde{\ell}^2 \widetilde{d}^2 / 4)}{n\pi^2 \widetilde{\ell}^2 \widetilde{d}^2 \widetilde{r}^2 / 4}$$

Here $\widetilde{\ell}^2 = \ell\sqrt{\ell^2 - 1}$, $\widetilde{d} = d/a$, and $\widetilde{r} = r/a$. Since for a random pair of source and target nodes, $d \leq \sqrt{2}a$, we have $p \simeq \frac{\ln(n\pi\widetilde{\ell}^2/4)}{n\pi^2\widetilde{\ell}^2\widetilde{r}^2/4}$. For example, consider a network with $n = 1000$, $a = 15$, and $r = 1$. For ellipse factor $\ell = 1.2$, we can estimate the relay probability $p$ such that the regional gossiping routing can deliver the packets almost surely as $p \simeq \frac{\ln(n\pi\widetilde{\ell}^2/4)}{n\pi^2\widetilde{\ell}^2\widetilde{r}^2/4} = 0.74$. The actual relay probability should be larger since we omit the number $c(N_d)$ here, which actually decides the success probability of the regional gossiping. The percentage of all vertices involved is at most $p \cdot N_d / n = \frac{\ln(n\pi\widetilde{\ell}^2\widetilde{d}^2/4)}{\pi\widetilde{r}^2 \cdot n} \simeq 0.46$. Since most pairs of source and target have distance $d$ much smaller than $a$, the actual number of involved vertices is much smaller. Let $P_d$ be the probability that a pair of source and target has distance $d$. The average percentage of number of vertices (for all source and target pairs) is actually $\int_{x=0}^{a} p \cdot N_x P_x / n\, dx$. It is not difficult to show that the percentage of vertices involved in regional gossiping is at most $p \cdot N_d / 2n = 0.23$. When the ellipse factor $\ell = \infty$, we can estimate the relay probability of the regional gossiping as $p \simeq \frac{\ln n}{n\pi\widetilde{r}^2} = 0.495$. The actual relay probability should be larger, so do the percentage of vertices involved in global gossiping. The experiments discussed in the following sections verify the above study.

## III. Experimental Studies

We tried unit disk graphs with different number of vertices that are randomly placed in a $15 \times 15$ square. We choose $n$ among $1000$, $1500$ and $2000$. The ellipse factor $\ell$ is chosen from $1.2$, $1.4$, $1.6$, $1.8$ and $2$. We also consider the case where the ellipse factor $\ell$ is $\infty$ which is just the traditional global gossiping method. The transmission range is chosen from $1$, $1.5$, $2$, $2.5$ and $3$. From [3], given $1000$ nodes in a $15 \times 15$ square, we need $r > 0.7$ to guarantee a connected network theoretically. We use the relay probabilities $p$ from $0.1$ to $1.0$ with step $0.1$ To study this transmission phenomenon in detail, we conduct further simulations using relay probabilities from $0.02$ to $0.30$ with step

$0.02$. We randomly select 100 pairs for each graph and conduct regional gossiping routing for each pair. To compute the delivery rate, we tried sending the message 1000 times for each pair. The algorithm works better for dense graphs than sparse graphs with the same parameters $p$ and $\ell$. Here we want to involve as little nodes as possible, i.e., to minimize the number of involved nodes that relay the message. The ellipse factor, the relay probability, and graph density are the major factors determining the number of involved vertices.

We first conduct extensive simulations to study the effect of the relay probability on the message delivery rate. Intuitively, if we increase the relay probability, the message delivery rate increases. Besides the relay probability, we vary either $\ell$, $n$, or $r$. As can be seen in Figure 1, when the probability $p$ exceeds some threshold the delivery rate jumps from near $0\%$ to near $100\%$. This threshold decreases as $\ell$ increases.
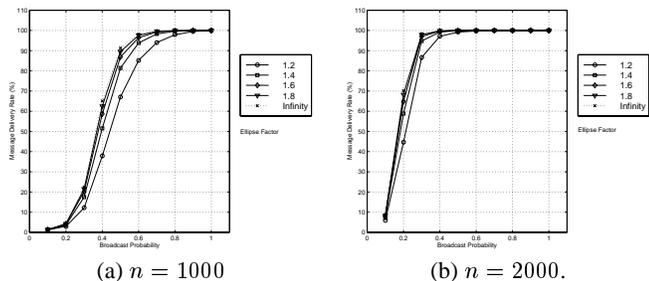


(a) $n = 1000$        (b) $n = 2000$.

Fig. 1. Message delivery rate v.s. relay probability. Here $r = 1$.

From Figure 1, we observed that when the graph becomes denser the curve jumps earlier, since each time when a vertex relays the message, more nodes will get the message so the probability that the message reaches the target increases. Another important observation is as we increase the ellipse factor, the message delivery rate does not increase proportionally. Surprisingly, when the ellipse factor is around $1.8$, the message delivery rate is almost as good as the one using global gossiping (i.e., $\ell = \infty$). Intuitively, the vertices far away from the line connecting the source and target do not help improving the message delivery rate. We also observe that, for dense graphs, even a narrow ellipse could guarantee a good message delivery rate. On the other hand, the ellipse factor does not compensate the the graph density.
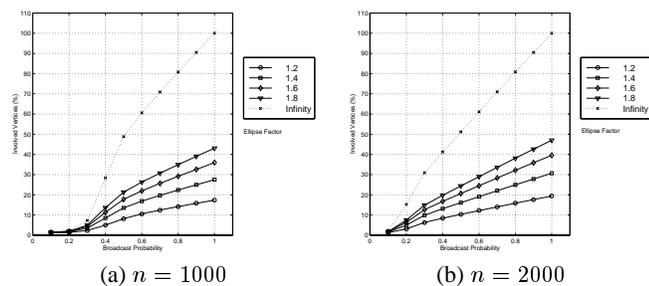


(a) $n = 1000$        (b) $n = 2000$

Fig. 2. Nodes involved v.s. relay probability. Here $r = 1$.

Not only the delivery rate is important, but also the number of vertices involved in the message delivery is important for the network life since the wireless devices are often powered by the batteries only. The challenge is to find an *ellipse factor* and a *relay probability* such that not only the delivery rate is high (close to $100\%$) but also the number of vertices involved in the message delivery is as small as possible. Actually they work against each other: a bigger ellipse factor implies higher delivery rate, but more vertices involved. As can be seen in Figure 2, the relation between the number of vertices involved and the relay probability with respect to ellipse factors is close to linear.

Assume that we want to have the delivery rate more than $99\%$ while minimize the number of involved nodes, denoted by $V_I$. We built the left table of Table I as follows: for each ellipse factor, we can find the needed relay probability to guarantee the message delivery above $99\%$ from Figure 1, and then by knowing the values of ellipse factor and the relay probability we can find the percentage of vertices that are involved from Figure 2. To guarantee the message delivery rate above $99\%$, we could involve only $15\%$ of vertices when $\ell = 1.2$, and $p = 0.9$ when $n = 1000$.

TABLE I

PERCENTAGE OF THE VERTICES INVOLVED IN MESSAGE DELIVERY

| $\ell$ | $p$ | $V_I$ % | $\ell$ | $p$ | $V_I$ % |
|---|---|---|---|---|---|
| 1.2 | 0.9 | 15% | 1.2 | 0.3 | 8% |
| 1.4 | 0.8 | 22% | 1.4 | 0.24 | 11% |
| 1.6 | 0.7 | 25% | 1.6 | 0.22 | 13% |
| 1.8 | 0.7 | 30% | 1.8 | 0.20 | 14% |
| infinity | 0.7 | 70% | infinity | 0.20 | 15% |

$n = 1000$                        $n = 2000$

We then study the effect of transmission range. In Figure 3 the transmission range is 2 and $n = 1000$. Obviously the larger the transmission range is, the denser the graph will be and that causes the jump to occur earlier. Since the jump happens earlier and quicker, we plot the figures using relay probability range $[0, 0.3]$ for $r = 2$, instead of $[0, 1]$ for $r = 1$. We built the right table of Table I for $n = 1000$, $r = 2$, the same way as we built the left one. It shows that we could involve only $8\%$ of vertices to guarantee the message delivery rate above $99\%$ for $n = 1000$, $r = 2$. For networks of 2000 nodes with $r = 2$, then only $6\%$ of vertices will be involved (figures are not shown here).
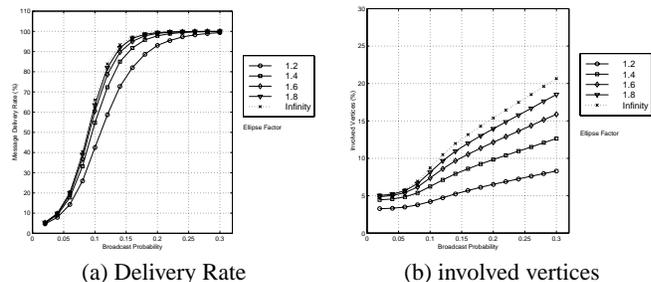


(a) Delivery Rate        (b) involved vertices

Fig. 3. Delivery rate and involved nodes v.s. relay probability

## IV. FUTURE WORK

We had assumed that two nodes can always communicate if their distance is no more than the transmission range. However, practically, some pair of nodes cannot communicate at all even they are close. We can model this by assigning another link probability $p_l$: a link exist with probability $p_l$. Here probability $p_l$ could be uniform or dependent on the distance between the pair of nodes. We need study the effect of link probability on the performance of regional gossiping. We can also use non-uniform ellipse factors: a bigger ellipse factor when the source and target are close.

## REFERENCES

[1] Z. Haas, J. Halpern, and L. Li, "Gossip-based ad hoc routing," in *IEEE INFOCOM*, 2002.
[2] R. Meester and R. Roy, *Continuum Percolation*, Cambridge University Press, 1996.
[3] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming, W. M. McEneaney, G. Yin, and Q. Zhang (Eds.)*, 1998.
[4] Peng-Jun Wan, Chih-Wei Yi, Xiang-Yang Li, Yu Wang, and Ophir Frieder, "Asymptotic distribution of critical transmission range for $k$-connectivity in wireless ad hoc networks," Submitted for publication, 2003.