Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Algorithms 52 (2004) 1–7
Journal of Algorithms
www.elsevier.com/locate/jalgor

Efficient communication in an ad-hoc network

Abraham Flaxman,[a,1] Alan Frieze,[a,*,2] and Eli Upfal[b,3]

[a] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[b] Department of Computer Science, Brown University, Providence, RI 02912, USA

Received 28 May 2003

Available online 16 March 2004

Abstract

We model the dynamics of an ad-hoc sensor network as a continuum percolation model, and prove that a simple local-flooding technique yields an efficient communication protocol in that setting.

1. Introduction

We consider a setting in which a large number of low power sensors are distributed in a large geographic area. The locations of the sensors are random and to save battery life only a random subset of the sensors are active at any given time. Furthermore, individual sensors are not reliable and a random subset of the sensors might be malfunctioning. A low power sensor can communicate directly only with neighbors within a relatively small radius. A communication to a distant unit must propagate through a large number of intermediate units. This setting poses especially challenging communication problems since individual sensors do not have a full knowledge of the network and the set of active nodes is dynamically changing in time [1,2,11]. A number of recent papers address basic static questions related to this setting, such as the density of sensors that guarantees connectivity of the system, exploring the relation between this setting and the percolation model studied

* Corresponding author.
E-mail addresses: abie@cmu.edu (A. Flaxman), alan@random.math.cmu.edu (A. Frieze), eli@cs.brown.edu (E. Upfal).
1 Supported in part by NSF VIGRE Grant DMS-9819950.
2 Supported in part by NSF grant CCR-0200945.
3 Supported in part by NSF grants CCR-0121154 and NSF-IIS 0325838.

0196-6774/$ – see front matter
doi:10.1016/j.jalgor.2004.01.004

in probability theory [4,6,7]. Very little has been done in developing provably efficient algorithms for communication and data exchange in this setting, which is the focus of our work.

Consider, for example, a large number of heat or fire sensors distributed in a large area. Even if only a random subset of the sensors is active at any given time, a sufficiently dense distribution of sensors would guarantee that a local fire triggers some number of sensors. The triggered sensors need now to propagate the signal to some data collecting units. Assume that data collection is done by units located in the perimeter of the controlled area. How can the signal be propagated as fast as possible to the data collecting units, and without using the battery power of too many intermediate units?

Since the topology of the network is unknown to the individual units, an obvious approach is to execute a sequence of local message flooding steps in order to push the message in the required direction, without involving an unnecessarily large number of units. The efficiency of this protocol depends on choosing the appropriate parameters for the flooding steps. If the flooding is too shallow there might be no path for the message. Too deep a flooding would waste battery power of too many units. Building on results from percolation theory we design an adaptive algorithm that transmits a message from an arbitrary location in the system to a boundary unit in time and number of hops that is close to linear in the length of the shortest path to the boundary. The total number of units involved in the communication is also close to linear in that length.

Another important communication problem is sending a message between two remote zones in the area covered by the sensor network. In the sensor setting there is no point in sending a message to an individual sensor since the set of active sensors is dynamically changing in time. Instead, we are interesting in a protocol for sending a message from some sensor in a given zone to some sensor in another zone. Our local flooding protocol provides an efficient solution to this problem as well.

2. Probabilistic model and main result

Place n points $X_i = (a_i, b_i)$, $i = 1, 2, \ldots, n$, uniformly at random in the unit square $S = [0, 1]^2$. Let D_i, $i = 1, 2, \ldots, n$, denote the disc with center X_i and radius $\rho = r/n^{1/2}$. The disc D_i represents the transmission area of a sensor located at X_i.

Consider the graph $G = (V, E)$ where $V = \{X_1, \ldots, X_n\}$ and $E = \{(X_i, X_j)\colon D_i \cap D_j \neq \emptyset\}$. A theorem of (continuum) percolation theory (see Penrose and Pisztora [9]) says that there is a critical radius r_c such that for $r > r_c$, there is a *giant component* C_g of G such that **whp** no point of S is at distance greater than $d_\alpha = \alpha \log n/n^{1/2}$, $\alpha = \alpha(r)$ from the set $S_g = \bigcup_{X_i \in C_g} D_i$. Thus, in the context of fire prevention, if a sensor can sense a fire at distance d_α then, for any location of fire in the area, at least one sensor in the giant connected component will be triggered by the fire. This sensor will be able to pass a message to the eastern boundary $\partial_E = \{(1, x)\colon 0 \leqslant x \leqslant 1\}$ since **whp** this boundary cuts S_g (else we would get a contradiction from considering a 1×2 rectangle). We present a simple algorithm which **whp** finds a path from any $X_i \in C_g$ to ∂_E of length $O(n^{1/2} \log n)$, which maintains a stack of depth $O((\log n)^2)$ and visits $O(n^{1/2} \log n)$ discs altogether. The algorithm LDFS is a *limited depth-first search*. In summary,

Theorem 1. *There is an algorithm LDFS which, assuming r is a sufficiently large constant,* **whp** *finds a path from any $X_i \in C_g$ to ∂_E of length $O(n^{1/2}\log n)$, which maintains a stack of depth $O((\log n)^2)$ and visits $O(n^{1/2}\log n)$ sensors altogether.*

The same technique can be used to communicate efficiently between different zones of the covered area.

Theorem 2. *Assume that the area $S = [0,1]^2$ is partitioned into sub-squares of side $\Theta(r(\log n)/n^{1/2})$. Fix an arbitrary pair of sub-squares, there is an algorithm which* **whp** *finds a path from some sensor in one sub-square to some sensor in the other square. The path is of length $O(n^{1/2}\log n)$, and the algorithm maintains a stack of depth $O((\log n)^2)$ and visits $O(n^{1/2}\log n)$ sensors altogether.*

3. The algorithm

Let c be a large constant and ε be a small constant, to be specified later. For any point (x, y), let $(\tilde{x}, \tilde{y})$ be the point with

$$\tilde{x} = \left\lfloor \frac{n^{1/2}}{\varepsilon r} x \right\rfloor \Big/ \left(\frac{n^{1/2}}{\varepsilon r}\right) \quad \text{and} \quad \tilde{y} = \left\lfloor \frac{n^{1/2}}{\varepsilon r} x \right\rfloor \Big/ \left(\frac{n^{1/2}}{\varepsilon r}\right).$$

Intuitively, $(\tilde{x}, \tilde{y})$ is the point representing (x, y) on the grid $\mathbb{Z}^2/(n^{1/2}/(\varepsilon r))$. Let $B_\ell(x_0, y_0)$ be a *quantized box* of radius ℓ centered at $(\tilde{x}_0, \tilde{y}_0)$. Formally,

$$B_\ell(x_0, y_0) = \big\{(x, y)\colon\ |x - \tilde{x}_0| \leqslant \ell \text{ and } |y - \tilde{y}_0| \leqslant \ell\big\}.$$

For simplicity, assume that a fire breaks out at the point $(0, 1/2)$ and the aim is to send a message to a sensor on ∂_E. We break the algorithm into *steps*.

Step 0. *Getting started*. The fire will set off a number of sensors, at least one will be in C_g. Sensors not in C_g will **whp** be in components of size $O(\log n)$—see [10]. We follow the progress of the message from one point $(x_1, y_1) \in C_g$. Furthermore, **whp** at most $O((\log n)^2)$ sensors will initiate message processes.

Step i. Let (x_i, y_i) be the position of the message at the start of Step i. Let G_i be the graph induced by the discs with centers in $B_{\ell_1}(x_i, y_i)$ for $\ell_1 = \varepsilon c r(\log n)/n^{1/2}$. Perform a depth-first search on G_i starting from (x_i, y_i) until the message reaches a point (x_{i+1}, y_{i+1}) with $x_{i+1} \geqslant x_i + \varepsilon r(\log n)/n^{1/2}$. If no such point is found, fail.

If the algorithm does not fail, then the message travels from left to right in at most $n^{1/2}/(\varepsilon r \log n)$ steps. In the next section, we will prove that **whp** each step visits at most $2c(1+\varepsilon)(\log n)^2$ discs **whp** and **whp** no step fails. So the message crosses the square visiting at most $2c(\varepsilon^{-1}+1)(n^{1/2}\log n)/r$ discs and the stack never grows beyond depth $2c(1+\varepsilon)(\log n)^2$ **whp**. We observe that the number of discs visited is within $O(\log n)$ of optimality.

The proof below that no step fails requires the radius of connectivity r to be larger than some constant r_0, which is larger than the percolation threshold r_c. It would be interesting to see if the same or similar result can be proved for any $r > r_c$.

3.1. Analysis of the algorithm

3.1.1. Proof that **whp** each step visits at most $(1+\varepsilon)(2c\log n)^2$ discs

Note that step i visits at most *all* the discs in $B_{\ell_1}(x_i, y_i)$. We will prove a statement stronger than needed: every quantized box $B_{\ell_1}(x, y)$ contains at most $(1+\varepsilon)(2\varepsilon rc\log n)^2$ disc centers **whp**. Let $\mathcal{E}_{x,y}$ be the event that $B_{\ell_1}(x, y)$ contains more than $(1+\varepsilon)\times(2\varepsilon rc\log n)^2$ discs. Since each point X_j is placed independently, uniformly at random in the unit square,

$$\Pr(\mathcal{E}_{x,y}) \leqslant \Pr\bigl(Bin\bigl(n, 4\ell_1^2\bigr) \geqslant 4(1+\varepsilon)\ell_1^2\bigr) \leqslant \exp\bigl\{-4\varepsilon^2\ell_1^2 n/3\bigr\}.$$

Since $B_{\ell_1}(x, y)$ is quantized, the set $\{\mathcal{E}_{x,y}\colon (x, y) \in [0, 1]^n\}$ has at most $(n/(\varepsilon r))^2$ distinct elements. So

$$\Pr\biggl(\bigcup_{(x,y)\in[0,1]^n} \mathcal{E}_{x,y}\biggr) \leqslant \biggl(\frac{n}{\varepsilon r}\biggr)^2 \Pr(\mathcal{E}_{x,y}) \leqslant \biggl(\frac{n}{\varepsilon r}\biggr)^2 \exp\bigl\{-4\varepsilon^2\ell_1^2 n/3\bigr\} = o(1). \qquad \Box$$

3.1.2. Proof that **whp** no step fails

We take an approach similar to Penrose [8] and Grimmet [3]. Let T be the graph with $V(T) = \{(\tilde{x}, \tilde{y})\colon (x, y) \in [0, 1]^2\}$ and edges connecting all nearest neighbors in $V(T)$, i.e. T is a 2-dimensional grid graph in the unit square with side length $\lambda = \varepsilon r/n^{1/2}$. We will show that if a steps fails there must exist a large *animal* in T which contains no disc.

Suppose at Step i the message is at point $(x_i, y_i) \in C_g$. Let $B = B_{\ell_2}(x_i, y_i)$ with $\ell_2 = \ell_1 - \rho$, $\rho = \frac{3}{4}\lambda$. Let G' be the subgraph of G with vertices in B, and let T' be the subgraph of T with vertices in B.

Let C be the component of G' containing (x_i, y_i). Let W denote the union of balls of radius ρ centered at the vertices of C. (We use balls to distinguish from the discs of radius λ). Note that W is path connected and we can assume that our paths do not go through T. Let U be the set of $z \in T'$ such that $z = (\tilde{x}, \tilde{y})$ for some $(x, y) \in W$.

We claim that the subgraph T'_U of T' induced by U is a connected graph. We prove this by induction. From the definition of U, we have that for any $u \in U$ there is some $w \in W$ with $\widetilde{w} = u$. And for any $w_1, w_2 \in W$, since W is path connected, we can find a path (in W) from w_1 to w_2 that intersects a minimum number of cells of T'. Let $d(u_1, u_2)$ denote this minimum. If $d(u_1, u_2) = 1$, then $u_1 = u_2$. Now, suppose that u_1 and u_2 are connected in T'_U for every pair u_1, u_2 with $d(u_1, u_2) \leqslant d$. Let now $u_1, u_2 \in U$ be vertices with $d(u_1, u_2) = d + 1$. Let $\widetilde{w}_1 = u_1$, $\widetilde{w}_2 = u_2$ and P be a path in W joining w_1, w_2 and let $A_1, A_2, \ldots, A_{d+1}$ be the successive cells of T visited by P. Then let w_3 be a point on P in the interior of A_2 so that $\widetilde{w}_3 \neq u_1$. Since $d(\widetilde{w}_3, u_2) \leqslant d$ (just use the path suggested by P), we know that $\widetilde{w}_3$ and u_2 are connected in $T'[U]$. And $u_1 \sim_{T'[U]} \widetilde{w}_3$, so u_1 and u_2 are connected in $T'[U]$ as well.

Let ∂U denote the boundary of U, i.e. the set of $z \in T' \setminus U$ such that z has a neighbor in U.

For each $z \in \partial U$, any (x, y) with $(\tilde{x}, \tilde{y}) = z$ is at most distance $\sqrt{5}\lambda + \rho$ from C. Choosing ε so that $\varepsilon \leqslant 1/(4\sqrt{5})$ we know $\{(x, y)\colon (\tilde{x}, \tilde{y}) = z\}$ contains no discs.

A subset A of T is called a *lattice animal* if for each $z, z' \in A$ there is a path from z to z' with $\|z_j - z_{j+1}\|_\infty = \lambda$ for each z_j in the path.

The next claim is slightly stronger than needed immediately, but will be useful for the proof of Theorem 2. (The condition $y' \leqslant y_i$ not needed for the proof of Theorem 1.)

Claim 1. *Let (x_i, y_i) be the disc at the start of the ith step. Suppose G' contains no path to a disc (x, y) with $x' \geqslant x_i + \ell_1/2$ and $y' \leqslant y_i$. Then ∂U contains a lattice animal with cardinality at least $\frac{1}{2}c \log n$.*

Proof. Let R be the set $(x, y) \in T'$ with $x' \geqslant x_i + \ell_1/2 + \rho$ and $y' \leqslant y_i$. Note that $R \cap U = \emptyset$. So the cut $(U, \overline{U})$ separates T' into components $A_1, A_2, \ldots, A_k$. Assume without loss of generality that $R \subseteq A_1$ and $U \subset A_2$. Let $U' = \bigcup_{i=2}^{k} A_i$. Then $(U', \overline{U'})$ separates T' into 2 components, so it is a minimal cut set (a *bond*).

By planar graph duality, the edges of the dual graph corresponding to $(U', \overline{U'})$ form a path or cycle. So the vertices of $\overline{U'}$ corresponding to dual vertices of this cycle are $*$-connected (i.e., form an animal). But these are all vertices of ∂U, so they form an animal free from discs.

It remains to lower bound the size of this animal. We do this by lower bounding the size of the cut $(U', \overline{U'})$. This is accomplished by exhibiting a collection of $\frac{1}{2}c \log n$ edge-disjoint paths from U to R. Since $U \subseteq U'$ and $R \subseteq \overline{U'}$ any cut separating U and U' must contain an edge from each of these paths.

We construct the flow as follows. Since (x_i, y_i) is path connected (in W) to a point on the boundary of B, U contains a path from $(\tilde{x}_i, \tilde{y}_i)$ to the boundary of T'. Suppose first that the path goes to the west side of the boundary, and let $z_1, \ldots, z_k$ be points on the path, with $z_j = (x_j, y_j)$ where $x_j = \tilde{x}_i - j\lambda$ and $k = \frac{1}{2}c \log n$. Now we route 1 unit of flow in an 'L', from $z_j = (x_j, y_j)$ to the point $(x_j, \tilde{y}_i - j\lambda)$ and from there to $(\tilde{x}_i + \frac{1}{2}\ell_1, \tilde{y}_i - j\lambda)$. This is illustrated in Fig. 1.

If the path goes to the north side, south side, or north half of the east side, we construct a similar set of $\frac{1}{2}c \log n$ edge-disjoint paths.

Any cut separating U from U' must have a different edge to break each of these paths, so any cut separating U from R has value at least $\frac{1}{2}c \log n$. □

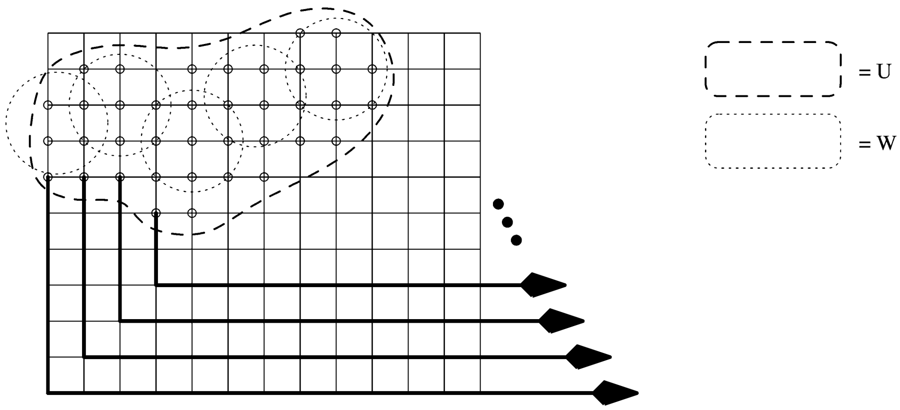

Fig. 1. A set of $\frac{1}{2}c \log n$ edge-disjoint paths.

Let $\mathcal{A}_m$ denote the set of lattice animals $A \subset T$ of size m. There is a constant γ such that the number of lattice animals of size m containing a given point is bounded above by $e^{\gamma m}$ (see [5]). Therefore the total number of lattice animals in T of size m in T is bounded by $(n^{1/2}/(\varepsilon r))^2 e^{\gamma m}$.

By the arguments above, the probability that there exists a step which fails is less than the probability that there exists a lattice animal A with size at least $\frac{1}{2}c \log n$ for which $\{(x, y)\colon (\tilde{x}, \tilde{y}) \in A\}$ contains no discs. We bound this probability as

$$\begin{aligned}
&\Pr(\text{exists a step which fails}) \\
&\quad \leqslant \sum_{m \geqslant \frac{c}{2}\log n} \Pr\Bigl(\text{exists } A \in \mathcal{A}_m\colon \bigl\{(x, y)\colon (\tilde{x}, \tilde{y}) \in A\bigr\} \cap V(G) = \emptyset\Bigr) \\
&\quad \leqslant \sum_{m \geqslant \frac{c}{2}\log n} \left(\frac{n}{\varepsilon r}\right)^2 e^{\gamma m} \left(1 - m\left(\frac{\varepsilon r}{n^{1/2}}\right)^2\right)^n \\
&\quad \leqslant \sum_{m \geqslant \frac{c}{2}\log n} (\varepsilon r)^{-2} e^{2\log n - (\varepsilon^2 r^2 - \gamma)m}.
\end{aligned}$$

Taking r large enough that $\varepsilon^2 r^2 - \gamma \geqslant 1$ and taking $c = 6$, we have

$$\Pr(\text{exists a step which fails}) \leqslant (\varepsilon r)^{-2} e^{-\log n} \frac{1}{1 - e^{-1}} = O(1/n). \qquad \square$$

Assume that $[0, 1]^2$ is partitioned into sub-squares of side $2\ell_1$. Then we see that the algorithm can efficiently find a path from any one square to any other, by traveling (almost) horizontally and then (almost) vertically. In fact, using the extra condition $y' \leqslant y_i$ (or $y' \geqslant y_i$ if needed) we can first make the message stay in the band $\{(x, y)\colon |y - y_1| \leqslant \ell_1\}$. Then we can make the message move vertically, proving Theorem 2.

References

[1] A. Cerpa, D. Estrin, ASCENT: Adaptive selfconfiguring sensor network topologies, in: Twenty First International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), June, 2002.

[2] D. Ganesan, R. Govindan, S. Shenker, D. Estrin, Highly resilient, energy efficient multipath routing in wireless sensor networks, Mobile Comput. Commun. Rev. 1 (2002).

[3] G. Grimmett, Percolation, 2nd ed., Springer, 1999.

[4] E. Jennings, C. Okino, On the diameter of sensors networks, in: Proceedings of the IEEE Aerospace Conference, Montana, 2002.

[5] D.A. Klarner, Cell growth problems, Canad. J. Math. 19 (1967) 851–863.

[6] S. Meguerdichian, F. Koushanfar, M. Potkonjak, M. Srivastava, Coverage problems in wireless ad-hoc sensor networks, in: Proceedings of INFOCOM, 2001, pp. 1380–1387.

[7] S. Muthukrishnan, G. Pandurangan, The bin-covering technique for thresholding random geometric graph properties, DIMACS Technical Report 2003-39, November 2003.

[8] M.D. Penrose, The longest edge of the random minimal spanning tree, Ann. Appl. Probab. 7 (1997) 340–361.

[9] M. Penrose, A. Pisztora, Large deviations for discrete and continuous percolation, Adv. Appl. Probab. 28 (1996) 29–52.
[10] M. Penrose, Random Geometric Graphs, Oxford University Press, 2003.
[11] A. Tsirigos, Z. Haas, S. Tabrizi, Multipath routing in mobile adhoc networks or how to route in the presence of topological changes, in: IEEE MILCOM, 2001.