Combinatorics, Probability and Computing (2009) 18, 145–163.
doi:10.1017/S0963548308009334 Printed in the United Kingdom

Line-of-Sight Networks

ALAN FRIEZE[1,†] JON KLEINBERG[2,‡]
R. RAVI[3§] and WARREN DEBANY[4 ¶]

[1]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, USA
(e-mail: alan@random.math.cmu.edu)

[2]Department of Computer Science, Cornell University, Ithaca NY 14853, USA
(e-mail: kleinber@cs.cornell.edu)

[3]Tepper School of Business, Carnegie Mellon University, Pittsburgh PA 15213, USA
(e-mail: ravi+@andrew.cmu.edu)

[4]Information Grid Division, Air Force Research Laboratory / RIG,
525 Brooks Road, Rome, NY 13441-4505, USA
(e-mail: warren.debany@rl.af.mil)

Received 25 February 2007; revised 28 January 2008; first published online 12 August 2008

Random geometric graphs have been one of the fundamental models for reasoning about wireless networks: one places n points at random in a region of the plane (typically a square or circle), and then connects pairs of points by an edge if they are within a fixed distance of one another. In addition to giving rise to a range of basic theoretical questions, this class of random graphs has been a central analytical tool in the wireless networking community.

For many of the primary applications of wireless networks, however, the underlying environment has a large number of obstacles, and communication can only take place among nodes when they are close in space *and* when they have line-of-sight access to one another – consider, for example, urban settings or large indoor environments. In such domains, the standard model of random geometric graphs is not a good approximation of the true constraints, since it is not designed to capture the line-of-sight restrictions.

Here we propose a random-graph model incorporating both range limitations and line-of-sight constraints, and we prove asymptotically tight results for k-connectivity. Specifically, we consider points placed randomly on a grid (or torus), such that each node can see up to a fixed distance along the row and column it belongs to. (We think of the rows and columns as 'streets' and 'avenues'

† Supported in part by NSF grant CCF-0502793.
‡ Supported by a David and Lucile Packard Foundation Fellowship, a John D. and Catherine T. MacArthur Foundation Fellowship, and NSF grants CCR-0122581, CCF-0325453, IIS-0329064, CNS-0403340, and BCS-0537606; work done in part while on sabbatical leave at Carnegie Mellon University.
§ Supported by NSF ITR grant CCR-0122581 (The ALADDIN project) and NSF grant CCF-043075.
¶ The fourth author's institution has designated this paper as approved for public release; distribution unlimited.

among a regularly spaced array of obstructions.) Further, we show that when the probability of node placement is a constant factor larger than the threshold for connectivity, near-shortest paths between pairs of nodes can be found, with high probability, by an algorithm using only local information. In addition to analysing connectivity and k-connectivity, we also study the emergence of a giant component, as well an approximation question, in which we seek to connect a set of given nodes in such an environment by adding a small set of additional 'relay' nodes.

1. Introduction

Most of today's approaches to wireless computing and communications are built on architectures where base stations connect the wireless devices to a supporting infrastructure. However, since the overwhelming trend is to transmit information in packets, over standard protocols, a dominant focus in the wireless research community is on more decentralized approaches where nodes cooperate to relay packets on behalf of other nodes. This focus is at the heart of current work on mobile *ad hoc* networks (MANETs) [19, 20].

Such networks can be viewed as consisting of a collection of nodes, representing wireless devices, positioned at various points in some physical region. The (wireless) 'links' of the network, joining pairs of nodes that can directly communicate with one another, are predominantly short-range and constrained by line of sight; this is an inevitable result of the scarcity of radio frequency (RF) spectrum and physical constraints on the propagation of RF and optical signals. The ways in which these physical limits on direct communication affect the overall performance of the network is a fundamental issue that motivates much of the theoretical work in this area.

Random geometric graphs. Given this framework, random geometric graphs have emerged as a dominant model for theoretical analysis of distributed wireless networks. One places n points uniformly at random in a geometric region (typically a disc or a square), and then, for a *range parameter* r, one connects each pair of nodes that are within distance r of one another. This model is the subject of a recent book by Penrose [22], and we refer the reader there for extensive background; we also note that the enormously influential work of Gupta and Kumar on the capacity of wireless networks is framed in this model as well [15, 16].

One of the most basic questions is to determine how the probability of connectivity of a random geometric graph depends on the number of nodes n and the range parameter r. A canonical result here is the following theorem of Penrose [21]. If we place n points uniformly at random in a unit square, and then continuously increase the range parameter r, with high probability the resulting geometric graph becomes k-connected at the smallest value of r for which there are no nodes of degree $< k$. In other words, the graph becomes k-connected at the moment that all trivial obstacles to k-connectivity (*i.e.*, low-degree nodes) disappear. An analogous type of result is familiar from the theory of classical Erdős–Rényi random-graph models, *e.g.*, Bollobás [4]. (For further results and discussion concerning thresholds for properties in random geometric graphs, see Goel, Rai and Krishnamachari [13].)

For modelling distributed wireless networks, the assumption of random node placement has proved to be a reasonable abstraction for the lack of structure in node locations, given that most frameworks for *ad hoc* networks assume some arbitrary initial 'scattering' of nodes, or that nodes reach their positions as a result of arbitrary mobility. More problematic is the fact that the analysis takes place in regions with no obstructions – in other words, that a node can communicate with *all* other nodes within distance r. This is at odds with the underlying constraints in many applications of distributed wireless networks, where there can generally be a large number of obstructions limiting communication between nearby nodes due to a lack of direct line-of-sight contact.

In other words, while random geometric graphs model wireless networks in open spaces, we lack a corresponding model for wireless networks in some of their most common domains: urban settings, large indoor environments, or any other context in which there are obstacles limiting visibility. With such a model would come the ability to address a range of basic theoretical problems. In particular, we are guided by the following genre of question:

How do connectivity and other structural properties of random geometric graphs change once we introduce line-of-sight constraints?

An understanding of such issues could help provide a framework for reasoning more generally about the performance of distributed wireless networks in obstructed environments.

The present work: Connectivity in line-of-sight networks. In this paper, we propose a random-graph model incorporating both range limitations and line-of-sight constraints, and we prove asymptotically tight results for k-connectivity. We also consider related structural questions, including the emergence of a giant component, as well as some of the algorithmic issues raised by the model.

To motivate the model, consider a stylized abstraction of limited-range wireless communication in an urban environment: there are n streets running east-west, n avenues running north-south, and wireless nodes can be placed at intersections of streets and avenues. Each node has *range* ω – it can see up to ω blocks north and south along the avenue it lies on, and up to ω blocks east and west along the street it lies on.

More concretely, we have an underlying set T of lattice points $\{(x, y) : x, y \in \{1, 2, \ldots, n\}\}$. We measure distance using the ℓ_1 metric, though to prevent complications arising from boundary effects in this presentation, we define the distance between points as though they form a torus:

$$d((x, y), (x', y')) = \min(|x - x'|, n - |x - x'|) + \min(|y - y'|, n - |y - y'|).$$

For a specified *range parameter* ω, we say that two points are *mutually visible* if they are in the same row or the same column of the torus, and if they are within distance at most ω from one another. We view the range ω as implicitly being a function of n, and in this paper we will make the assumption that ω is asymptotically bounded below by $\ln n$ and above by some polynomial in n; specifically, we assume $\ln n = o(\omega)$ and that $\omega = O(n^{\delta})$ for a value of $\delta < 1$ to be specified below.

We now study the random graph G that results if, for some *placement probability* $p > 0$, we locate a node at each point of T independently with probability p, and then connect those pairs of nodes that are mutually visible. As p increases, the torus becomes more crowded with nodes, and the resulting graph G is more likely to be connected. Our main result states, roughly, that the smallest value of p at which G becomes k-connected with high probability is asymptotically the same as the smallest value of p at which the minimum degree in G is k with high probability.

More concretely, for a critical value of the placement probability $p^* = O(\frac{\ln n}{\omega})$, we find that in an interval of width $O(\frac{1}{\omega})$ around p^*, the random graph G goes from being k-connected with arbitrarily small probability to being k-connected with probability arbitrarily close to 1. Moreover, the probability that G has no nodes of degree $< k$ undergoes a comparable transition in a corresponding interval around p^*. We state this theorem about k-connectivity as follows.

First, we write $\omega = n^{\delta}$ where we assume that $\omega \gg \ln n$ and $\delta < \frac{4}{8k+7}$. Note that we *do not* preclude the case where $\delta = o(1)$.

Theorem 1.1. *Let $k \geqslant 1$ be a fixed positive integer and let $p = \frac{(1-\frac{1}{2}\delta)\ln n + \frac{k}{2}\ln\ln n + c_n}{2\omega}$. Then*

$$\lim_{n\to\infty} \Pr(G \textit{ is } k\textit{-connected}) = \begin{cases} 0 & c_n \to -\infty, \\ e^{-\lambda_k} & c_n \to c, \\ 1 & c_n \to \infty, \end{cases}$$

where

$$\lambda_k = \frac{2^{k-2}\left(1-\frac{1}{2}\delta\right)^k e^{-2c}}{(k-1)!}.$$

The proof of this result, which occupies Section 2 of the paper, requires techniques quite different from the analysis of standard geometric random graphs, due to the line-of-sight constraints. One way to appreciate why this appears necessary is to consider that, as we vary ω, the resulting model interpolates between two well-known but qualitatively different random-graph models. When $\omega = 1$, so that a node can only see neighbouring points, we have site percolation on a lattice, a well-studied problem that is still not completely well understood. At the other extreme, when $\omega = n$ and nodes can see all points in their row and column, it is easy to see that the model is equivalent to a purely graph-theoretic one in which we start with the complete bipartite graph $K_{n,n}$ and keep each edge with probability p. Note that our bounds on ω preclude either of these exact extremes, but our analysis for the 'middle region' of ω that we consider involves ingredients from both extremes, combining techniques from 'classical' random-graph analysis with the combinatorics of the underlying grid of points.

Remarks. **(a)** The reader might wonder if the constraint $\omega \gg \ln n$ is really necessary. Suppose for example that $\omega = o(\ln n)$. The expected number of isolated vertices $X_0 = n^2 p(1-p)^{4\omega}$. If $p = o(1)$ then $X_0 \to 0$ only when $n^2 p \to 0$. If p is bounded from below, then unless $p = 1 - o(1)$, $X_0 = n^{2-o(1)}$ and one can show that w.h.p. $X_0 \neq 0$. Thus the threshold for connectivity is very close to one when $\omega = o(\ln n)$ and perhaps therefore somewhat less interesting.

(b) Theorem 1.1 could be strengthened to give a hitting time version where we add random vertices one at a time. Then w.h.p. the first vertex that makes the graph have minimum degree k will also make the graph k-connected. We will make some remarks on this version at the end of the proof of Theorem 1.1.

The present work: Further results. We consider the emergence of a giant component in our model. Note here that since G itself has $O(n^2 p)$ vertices, a giant component is one with $\Omega(n^2 p)$ vertices.

Theorem 1.2.

(a) *If $p = \frac{c}{\omega}$ where $c > 1$ and $\omega \to \infty$ then w.h.p. G contains a component with at least $(2\rho_c - \rho_c^2 - o(1))cn^2/\omega$ vertices, where ρ_c is the unique solution in $(0,1)$ of $1 - x = e^{-cx}$.*

(b) *If $p = \frac{c}{\omega}$ where $c < 1/(4e)$ and $\omega \to \infty$ then w.h.p. the largest component in G has size $O(\ln n)$.*

We also consider the problem of how nodes in such a random graph can construct paths between each other, possessing knowledge of their own coordinates but otherwise having only local information. We show that when p exceeds the threshold for connectivity by a fixed (relatively small) constant factor – *i.e.*, $p = C \ln n/\omega$ – then a simple decentralized algorithm allows a given pair of nodes at ℓ_1-distance d to construct, with high probability, a path of $O(d/\omega + \ln n)$ edges while involving only $O(d/\omega + \omega \ln n)$ nodes in the computation. This is nearly optimal, even with global information, since $\Omega(d/\omega)$ is a simple lower bound on the length of any path between nodes at ℓ_1-distance d (and hence also a lower bound on the number of nodes who need to participate in the construction of the path).

Theorem 1.3. *Let $p = C \ln n/\omega$ for a sufficiently large constant C and $\omega \geqslant C \ln n$. There is a decentralized algorithm that, given s and t, with high probability constructs an s–t path with $O(d(s,t)/\omega + \ln n)$ edges while involving $O(d(s,t)/\omega + \omega \ln n)$ nodes in the computation.*

Finally, we consider a basic algorithmic problem in a non-random version of the line-of-sight model: given an input set of nodes, we would like to add a small set of additional nodes so that the full set becomes connected. More concretely, suppose we are given a set of nodes at points $X \subset T$, such that the graph on X (defined by visibility with respect to the range parameter ω) is not connected. We would like to add further nodes, at a set $Y \subset T$, where Y should be as small as possible subject to the constraint that the graph on $X \cup Y$ should be connected. We think of the additional nodes Y as 'relays' that connect the original nodes in X under line-of-sight constraints; as a result, we refer to this as the *relay placement* problem.

By considering the graph of mutual visibility, and viewing the nodes in Y as Steiner nodes, an instance of relay placement can be easily cast as an instance of the *node-weighted Steiner tree* problem. The general node-weighted Steiner tree problem is inapproximable to within a factor of $\Omega(\ln n)$ (Klein and Ravi [18]). For the class of line-of-sight networks that we study here, however, we show how to exploit the underlying visibility structure to obtain a constant-factor approximation. In particular, we make use of a graph-theoretic notion that we call *cohesiveness*, which suggests some combinatorial questions of independent interest.

Theorem 1.4. *There is a polynomial-time algorithm that produces a Steiner set whose total cost is within a factor of 6.2 of optimal.*

Relay placement is clearly related to certain algorithmic *art gallery* problems (see, *e.g.*, Efrat and Har-Peled [8] and Efrat, Har-Peled and Mitchell [9], and the VC-dimension results in Kalai and Matousek [17] and Valtr [25]), since there too one is placing nodes in a region subject to visibility constraints. However, the problems considered in the literature on art gallery problems have a different focus, as they are concerned with placing nodes so as to see the entire region, as opposed to adding Steiner nodes so as to create a connected visibility graph, as we do here.

A preliminary version of this paper appeared in [11]. Also, since the appearance of the preliminary version, Bollobás, Janson and Riordan [5] have tightened Theorem 1.2 and shown that the threshold for the appearance of a giant component is at $p = \frac{\log(3/2)}{\omega}$.

2. Connectivity

This section is devoted to the proof of Theorem 1.1. We will concentrate first on the case where $c_n \to c$ and to avoid trivialities we will assume that $c_n = c$. Thus, until further notice, we will assume that

$$p = \frac{(1 - \frac{1}{2}\delta)\ln n + \frac{k}{2}\ln\ln n + c}{2\omega}.$$

The overall outline of the proof is as follows. We start by first studying the distribution of the minimum degree of G. We then imagine adding nodes in two stages – most of the nodes in the first stage, and a few final nodes in the second stage. Now, suppose the graph H formed by nodes added in the first stage can be disconnected by the deletion of some set S of fewer than k nodes. We argue that, with high probability, any two components J and K of $H - S$ come 'close' to one another at many disjoint locations on the torus T – in particular, at each of these locations, there is some point of the torus that sees nodes in both J and K. When we then add nodes in the second stage, it is enough that a node is placed at one of these points that can see both components; and we argue that there are enough such points that this happens with high probability. We also check that w.h.p. the new vertices do not create any small cuts.

2.1. Minimum degree computation

Proposition 2.1. $\lim_{n\to\infty} \Pr(G \text{ contains a vertex of degree} < k) = 1 - e^{-\lambda_k}$. □

Proof. Let X_l denote the number of vertices of degree $0 \leqslant l < k$. Then observe first that

$$\begin{aligned}
\mathbf{E}[X_l] &= n^2 p \binom{4\omega}{l} p^l (1-p)^{4\omega - l} \\
&\sim n^2 p^{l+1} \frac{4^l \omega^l}{l!} e^{-4\omega p} \\
&\sim n^2 \left(\frac{\left(1 - \frac{1}{2}\delta\right)\ln n}{2\omega} \right)^{l+1} \frac{4^l \omega^l}{l!} \frac{n^\delta e^{-2c}}{n^2 (\ln n)^k} \\
&\sim \begin{cases} 0 & l \leqslant k-2, \\ \lambda_k & l = k-1. \end{cases}
\end{aligned}$$

Thus the expected number of vertices of degree less than k is asymptotically λ_k. The rest of the proof is quite standard: see, for example, Bollobás [4]. Let S_k denote the set of vertices of degree less than k in G and let $X = |S_k|$. Let X'' denote the number of pairs of vertices $v, w \in S_k$ such that v, w are within ℓ_1 distance 2ω of each other. Let X' denote the number of vertices in S_k which are at ℓ_1 distance greater than 2ω from any other vertex in S_k. Then

$$X' \leqslant X \leqslant X' + X''.$$

Now

$$\mathbf{E}[X''] \leqslant 16\omega^2 n^2 p^2 \binom{8\omega}{2k}(1-p)^{6\omega-2k} = o(1),$$

using our upper bound on δ. Thus $X = X'$ with high probability.

Now fix a positive integer t. Then, where $(a)_t = a(a-1)\cdots(a-t+1)$, we compute

$$\left((n^2 - 16t\omega^2)p\sum_{i=0}^{k-1} p^i(1-p)^{4\omega-i}\right)^t \leqslant \mathbf{E}[(X')_t] \leqslant \left(n^2 p\sum_{i=0}^{k-1} p^i(1-p)^{4\omega-i}\right)^t,$$

which implies that

$$\lim_{n\to\infty} \mathbf{E}[(X')_t] = \lambda_k^t,$$

and so X' is asymptotically Poisson with mean λ_k, which implies the lemma. □

2.2. Probabilistic part of proof

We imagine placing nodes at random according to the following two-stage process. We place a node at each point with probability p_1 in the first stage. We then independently place a node at each point with probability p_2 in the second stage. We choose

$$p_1 = \frac{(1-\frac{1}{2}\delta)\ln n + \frac{k}{2}\ln\ln n + c - (\ln n)^{-1}}{2\omega} \geqslant \frac{\ln n}{3\omega}$$

and p_2 so that this is equivalent to the original placement process with probability p, in which case

$$p_2 \sim \frac{1}{2\omega \ln n}.$$

For ease of terminology, we say that a node is *red* if it was placed in the first stage, and we say that it is *blue* if it is placed in the second stage at a point not hit by the first stage. Let H denote the subgraph of G consisting only of red nodes.

For each point in T, we define its four *arms* to be the four sets of ω points that are visible from it in a single direction (north, south, east and west). We further partition each arm α of point x into 10 consecutive segments $\alpha_1, \alpha_2, \ldots, \alpha_{10}$ of length $\omega/10$. A segment is said to be *weak* if it contains fewer than $\frac{1}{50}\ln n$ red nodes. Otherwise we say that is *strong*. An arm is said to be *mighty* if all its segments are strong.

Let $\mathcal{E}_1$ be the event that there exists a red vertex which has an arm α on which we can find 1000 red vertices, each having an arm orthogonal to α which is not mighty.

Lemma 2.2. $\Pr(\mathcal{E}_1) = o(1)$.

Proof. For a fixed vertex x and arm α, the probability that the arm contains a weak segment can be bounded by

$$10\Pr\left(\mathrm{Bin}(\omega/10, p_1) \leqslant \frac{1}{50}\ln n\right) \leqslant e^{-(\ln n)/400} = n^{-1/400}.$$

So the probability that there is a red node giving rise to $\mathcal{E}_1$ is bounded by

$$8n^2\binom{\omega}{1000}p_1^{1000}n^{-1000/400} = o(1). \qquad \square$$

Now let $\mathcal{E}_2$ be the event that there exists a red vertex v of degree less than $\ln\ln n$ that has a red neighbour w such that w has an arm orthogonal to vw which is not mighty.

Lemma 2.3. $\Pr(\mathcal{E}_2) = o(1)$.

Proof. The probability that H contains such a pair v, w is bounded by

$$\begin{aligned} n^2 p_1 \sum_{t=1}^{\ln\ln n}\binom{4\omega}{t}p_1^t(1-p_1)^{4\omega-t}(2n^{-1/400}) &\\ \leqslant 2n^{-1/400}\sum_{t=1}^{\ln\ln n}\left(\frac{(4+o(1))e\ln n}{t}\right)^t e^{-2c+o(1)} &\\ = o(1). & \qquad \square \end{aligned}$$

Now let $\mathcal{E}_3$ be the event that there exists a red vertex with at most $k-1$ red neighbours and at least one blue neighbour.

Lemma 2.4. $\Pr(\mathcal{E}_3) = o(1)$.

Proof. The probability that H contains such a vertex v is bounded by

$$n^2 p_1 \sum_{t=0}^{k-1}\binom{4\omega}{t}p^t(1-p_1)^{4\omega-t}(4\omega p_2) \sim 4\lambda_k\omega p_2 = o(1). \qquad \square$$

Now let $\mathcal{E}_4$ be the event that there exists a blue vertex with fewer than k red neighbours.

Lemma 2.5. $\Pr(\mathcal{E}_4) = o(1)$.

Proof. The probability that G contains such a vertex v is bounded by

$$n^2 p_2 \sum_{t=0}^{k-1}\binom{4\omega}{t}p_1^t(1-p_1)^{4\omega-t} \sim \frac{\lambda_k p_2}{p_1} = o(1). \qquad \square$$

For the non-probabilistic part of this argument, we will assume that none of $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4$ hold.

2.3. Non-probabilistic part of proof

For the next part, we assume that $\delta(G) \geqslant k$.

Recall that H is the subgraph of G consisting only of the red nodes. Let S be an arbitrary set of $k-1$ red vertices, and let $H_S = H - S$. Our main goal is to show that if H_S has multiple connected components, then with high probability they will all be linked up by the addition of the blue nodes.

Let L be the set of points in T with coordinates (i, j), where each of i and j is a multiple of 3ω. For each connected component K of H_S, and for each point $x \in L$, let v_{Kx} denote the node in K that is closest to x in ℓ_1 distance. We make the following claim.

Lemma 2.6. *v_{Kx} lies within the $\omega \times \omega$ box B_x centred at x.*

Proof. Let a red node be *pink* if it is not in S. Assume without loss of generality that the point x is located at the origin of the torus, which we denote $x = 0$. Suppose that $v = v_{K0} = (a, b)$ is NE of 0 and that it does not lie in B_0. Then v has at least one arm containing a pink node w. This follows from the non-occurrence of $\mathcal{E}_3$. If the degree of v is less than $\ln\ln n$ then we can use the non-occurrence of $\mathcal{E}_2$ to argue that the two arms of w orthogonal to vw are mighty. If the degree of v is greater than $\ln\ln n$ then we can use the non-occurrence of $\mathcal{E}_1$ to argue that there is a choice of $\ln\ln n - 2000$ ws such that the two arms of w orthogonal to vw are mighty. Let α denote the arm of v containing a w with mighty arms. Note that every segment of a mighty arm contains at least $\frac{1}{51}\ln n$ pink nodes.

Case 1: α is the south arm of v.
If $a \leqslant \omega/2$ then any pink node on α is either in B_0 or closer to 0 than v_{K0}. Similarly, if $b > \omega/2$ then any pink node on α is closer to 0 than v_{K0}. So we can assume that $a > \omega/2 \geqslant b$. Also, if $(a, b') \in \alpha$ then we must have $0 > b' = -b''$ where we can assume that $b \leqslant b'' \leqslant \omega - b$.

Choose such a pink node $(a, -b'')$ with a mighty west arm β. Now choose a pink node $w = (a', -b'') \in \beta$ such that (i) $a - a' \in [0.4\omega, 0.5\omega]$ and (ii) the north arm γ of w is mighty. Now choose a pink node $(a', c) \in \gamma$ such that $|c - b| \leqslant 0.1\omega$. We can make these choices because of the non-occurrence of $\mathcal{E}_1$ and the fact that $\frac{1}{51}\ln n > 1000 + k$. It follows that $|a'| + |c| \leqslant a + b + 0.1\omega - 0.4\omega$, a contradiction.

Case 2a: α is the north arm of v and $a \geqslant \omega/2$.
Choose a pink node $(a, b') \in \alpha$ with a mighty west arm β. Then choose a pink node $w = (a', b') \in \beta$ such that (i) $a - a' \in [0.4\omega, 0.5\omega]$ and (ii) the south arm γ of w is mighty. Now choose a pink node $(a', b'') \in \gamma$ such that $|b'' - b| \leqslant 0.1\omega$. It follows that $|a'| + |b''| \leqslant a + b + 0.1\omega - 0.4\omega$, a contradiction.

Case 2b: α is the north arm of v and $a < \omega/2$.
We must have $b > \omega/2$, else $v_{K0} \in B_0$. Choose a pink node $(a, b') \in \alpha$ with a mighty west arm β. Then choose a pink node $w = (a', b') \in \beta$ such that (i) $|a - a'| \leqslant 0.1\omega$ and (ii) the south arm γ of w is mighty.

If $|b - b'| \leqslant 0.7\omega$ then choose a pink node $(a', b'') \in \gamma$ such that $|b'' - b| \in [0.9\omega, \omega]$. It follows that $|a'| + |b''| \leqslant a + b + 0.1\omega + 0.7\omega - 0.9\omega$, a contradiction. Otherwise, $|b - b'| > 0.7\omega$. We can choose a pink node $y = (a', b'') \in \gamma$ such that the west arm δ of y is mighty and $|b' - b''| \geqslant 0.9\omega$. Choose a pink node $z = (a'', b'') \in \delta$ such that $|a'' - a'| \leqslant 0.1\omega$ and its south arm ε is mighty. Finally, we note that there exists a pink node $w = (a'', b''') \in \varepsilon$ such that $|b'' - b'''| \in [0.5\omega, 0.6\omega]$. Then we have $|a''| + |b'''| \leqslant a + b + \omega + 0.1\omega - 0.9\omega + 0.1\omega - 0.5\omega$, a contradiction.

The case where α is the west arm is dealt with as in Case 1 and the case where α is the east arm is dealt with as in Case 2. $\square$

Now, let J and K be two distinct components of H_S. Since v_{Jx} and v_{Kx} both lie in the $\omega \times \omega$ box around x, there is some point $z(J,K,x)$ that is visible from both of them. We make the following observation.

Lemma 2.7. *The points $z(J,K,x)$ and $z(J,K,y)$ are distinct, for distinct points $x, y \in L$.*

Proof. $z(J,K,x)$ lies in the $\omega \times \omega$ box around x, and $z(J,K,y)$ lies in the $\omega \times \omega$ box around y, and these boxes are disjoint, since x and y are at least 3ω apart. □

2.4. Finishing the proof

Note that if a node is placed at $z(J,K,x)$, then it will be a neighbour both of a point in J and K, and hence J and K will belong to the same component in G. In the second stage of node placement, a blue node will be placed at each point $z(J,K,x)$ with probability p_2. We should not, however, forget that we have conditioned on the events $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4$ not occurring and that $\delta(G) \geqslant k$. This accounts for the $(e^{\lambda_k} + o(1))$ factor in (2.1) below. By Lemma 2.7, there are $\frac{n^2}{9\omega^2}$ such points for a fixed pair of components J, K, and so the conditional probability that no blue point is placed at any of them is bounded by

$$(e^{\lambda_k} + o(1))(1 - p_2)^{n^2/(9\omega^2)} \leqslant e^{-n^2/(20\omega^3 \ln n)} \leqslant e^{-n^{2-3\delta}/(20 \ln n)}. \tag{2.1}$$

There are at most ω^2 components, since for any fixed point $x \in L$, each component has a node in the $\omega \times \omega$ box around x. Thus, the probability that there exists a set S of size at most $k-1$ and components J, K of H_S, which are not connected in G by a blue vertex, is at most $\omega^4 e^{-n^{2-3\delta}/(20 \ln n)} n^{2k-2} = o(1)$. Thus, conditional on there being no vertices of degree $k-1$ or less, if we remove any set S of $k-1$ red vertices, then w.h.p. the graph H_S is connected. Deleting blue vertices does not affect H_S, and so if we remove any set S of $k-1$ vertices, then w.h.p. the graph H_S. It follows from the non-occurrence of $\mathcal{E}_4$ that $G - S$ will also be connected.

This finishes the case $c_n \to c$. If $c_n \to -\infty$ then one uses the Chebyshev inequality to show that with high probability there are vertices of degree less than k. If $c_n \to \infty$ then with high probability there are no vertices of degree less than k (the expected number tends to zero), and the argument for $c_n \to c$ implies that G will be k-connected with high probability.

This completes the proof of Theorem 1.1.

Remark. One can prove a hitting time version by modifying the proof as follows.

(1) Let $m_\pm = n^2\left(\frac{(1-\frac{1}{2}\delta)\ln n + \frac{k}{2}\ln\ln n \pm \ln\ln\ln n}{2\omega}\right)$. If we put in m_- random vertices, then w.h.p. the resulting graph G_0 has minimum degree $k-1$ and $O((\ln\ln n)^2)$ vertices of degree $k-1$.
(2) The vertices of degree $k-1$ are all more than 10ω apart in G_0.
(3) The graph G_1 obtained by deleting the vertices of $k-1$ will be k-connected. This follows from a minor change to the proof of Theorem 1.1.
(4) Adding $m_+ - m_-$ random vertices results in a graph G_2 which is k-connected and none of the new vertices have degree less than $k+1$ when they are added.

These four properties imply that the graph is k-connected at the time the graph has minimum degree k.

3. The existence of a giant component: Proof of Theorem 1.2

We now consider the existence of a *giant* component in our model of line-of-sight networks.

(a) To prove this part of the theorem, we first require a lemma about the existence of a giant component in the random graph $H = B_{m,m,q}$ where $q = d/m$. Here we create H by including each edge of the complete bipartite graph $K_{m,m}$ independently with probability q.

Lemma 3.1. *If $d > 1$ then w.h.p. H contains a component C_g with $(1 - o(1))\rho_d m$ vertices on each side of the partition, where ρ_d is the unique solution in $(0, 1)$ of $1 - x = e^{-dx}$. Furthermore C_g contains $(1 - o(1))(2\rho_d - \rho_d^2 - o(1))dm$ edges.*

Proof. We follow the proof of the existence of a giant component via branching processes as elaborated in Chapters 10.4 and 10.5 of Alon and Spencer [1]. Note that the degree of a vertex of H has a distribution which is asymptotically Poisson with mean d and the proof in [1] can easily be adapted to H. This will show that C_g has $\sim (1 - x_d)m$ vertices on each side. To get the number of edges, imagine the model where we fix the number of edges as $\mu \sim dm$. Suppose now we put in $\mu - 1$ random edges and obtain a giant component C_g' with $(1 - o(1))(1 - x_d)m$ vertices on each side. Now put in the μth random edge. We see that the probability it is not part of the giant component C_g is $\sim x_d^2$. This shows that $|E(C_g)| \sim (1 - x_d^2)m$ in expectation. By adding two random edges we can estimate the variance and then use the Chebyshev inequality. □

Now divide the torus T into $N = n^2/\omega^2$ sub-squares $S_1, S_2, \ldots, S_N$ of size $\omega \times \omega$. Fix a particular sub-square S_i, and consider the bipartite graph H_i with $\omega + \omega$ vertices $R_i \cup C_i$ (rows/ columns) where there is an edge $(x, y) \in R_i \times C_i$ if the gridpoint of T corresponding to (x, y) is occupied by a node of G. Applying Lemma 3.1 with $m = \omega$ and $d = c$, we see that, with probability $(1 - o(1))$, H_i contains a giant component Γ_i with $(1 - o(1))(1 - x_c)\omega$ vertices on each side and $(1 - o(1))(1 - x_c^2)\omega^2$ edges. When translated into a subgraph of G, we see that Γ_i induces a connected subgraph G_i with $(1 - o(1))(1 - x_c^2)\omega^2$ vertices. This is because each edge of H_i corresponds to a vertex of G.

We divide each sub-square S_i further into 16 $\omega/4 \times \omega/4$ sub-squares. We choose 4 *special* sub-squares $S_{i,1}, \ldots, S_{i,4}$. These will either be at $(1, 2), (2, 1), (3, 4), (4, 3)$ or at $(1, 3), (2, 4), (3, 1), (4, 2)$, where (i, j) denotes the sub-square in row i, column j, $1 \leqslant i, j \leqslant 4$. We then have these two sorts of sub-square alternate along the rows and columns of T as in Figure 1.

Each special sub-square is associated with a direction. If $i = 1$ then the direction is north. If $i = 4$ then the direction is south. If $j = 1$ then the direction is west and if $j = 4$ then the direction is east.

Now w.h.p. each of the 4 special sub-squares will contain $\sim(1 - x_c)\omega/4$ *useable* columns (north or south sub-squares) or rows (east or west sub-squares) that correspond to vertices of a giant component of the corresponding H_i. We say that a square S_i is *good* if H_i contains a giant component with $\sim(1 - x_c^2)\omega^2$ edges and each special sub-square has $\sim(1 - x_c)\omega/4$ useable rows or columns, depending on its direction.

If S_i is good then we choose $(1 - x_c)\omega/5$ random rows or columns from the useable rows or columns of each of the four special sub-squares. Suppose that $X_{i,j}$ is the set of rows or columns

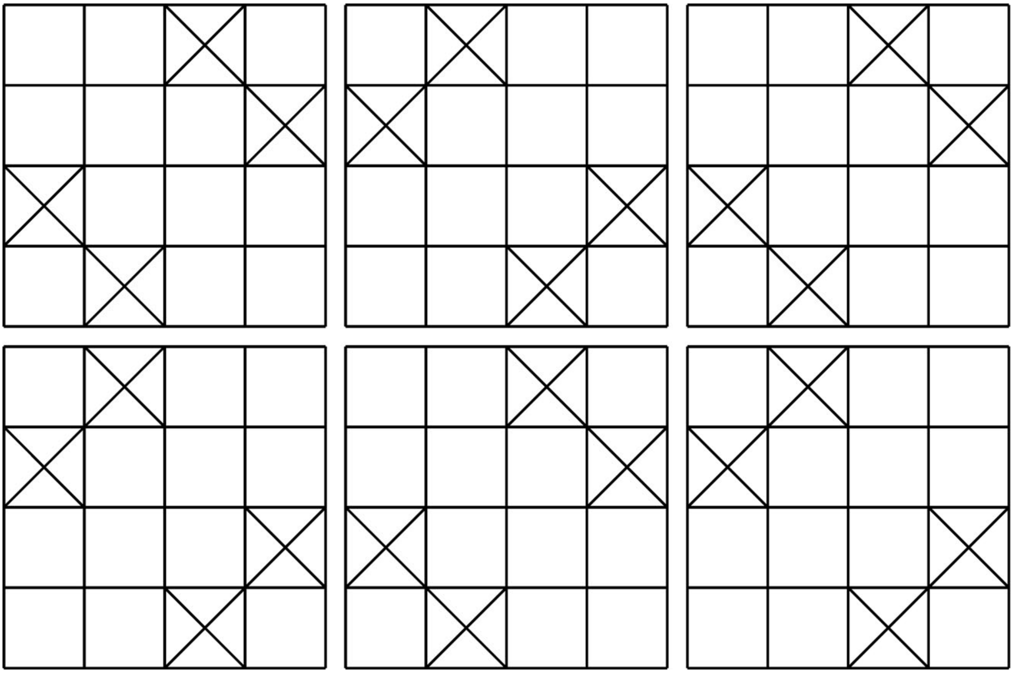

Figure 1.

chosen from $S_{i,j}$. We observe that conditional on S_i being good, the sets $X_{i,j}$ are uniformly random and independent of each other.

We are now in a position to use *mixed percolation*. Let $\mathcal{L}$ denote the $N \times N$, $N = n/\omega$ lattice $\mathcal{L}$ with site percolation $p_V = 1 - o(1)$ and bond percolation $p_E = 1 - o(1)$. Here we place a vertex at site i if the square S_i is good. If two adjacent sites H_i, H_{i+1}, say, are good then we join them by an edge in the lattice if the following holds. Let the adjacent special squares be $S_{i,r}$ and $S_{i+1,s}$. We add the edge if $X_{i,r} \cap X_{i+1,s} \neq \emptyset$. If this occurs then there is a pair of nodes of G, $u \in G_i$, $v \in G_{i+1}$ such that u, v are in the same row or column and are at distance $\leqslant \omega$ apart. Hence G_i and G_{i+1} will form part of the same component in G.

In this model of percolation the giant cluster will contain almost all of the points: see the argument below. We had the good fortune to get outline proofs, for the case of an $n \times n$ grid, from two experts on percolation.[†,‡]

For completeness, we include an elementary proof of this fact.

† Agoston Pisztora: That there is a cluster of this size follows from a simple generalization of Theorem 1.1 of Deuschel and Pisztora [7].

‡ Geoffrey Grimmett: Let the density $p \in (0, 1)$ of open sites be fixed. Let I_n be the set of vertices of B_n (the $n \times n$ box containing the origin) that lie in infinite open paths. With probability $1 - o(1)$, the largest open cluster C_n of B_n satisfies $C_n \oplus I_n \subseteq B_n \setminus B_{n-m}$ where $m = (\ln n)^2$. See the proof of Lemma 2 of [10], particularly (3.5)–(3.6). Therefore, the density of the largest cluster of B_n is very close to the percolation probability $\theta(p)$. The claim then follows once we have that $\theta(p) \to 1$ as $p \to 1$. This holds by (1.18) of Grimmett [14].

Size of the giant. We now argue that if, in the $N \times N$ toroidal lattice $\mathcal{L}_N$, we have $p_V = p_E = p = 1 - 1/\gamma$ where $\gamma = \gamma(N) \to \infty$ as $N \to \infty$, then w.h.p. the size of the giant cluster in the associated random sub-graph $\mathcal{L}_p$ is $(1 - o(1))N^2$. We can assume that $\gamma \leqslant N^2 \ln N$, for otherwise $\mathcal{L}_p = \mathcal{L}_N$, w.h.p.

Let $\lambda = \gamma^{1/3}$ and partition $\mathcal{L}_N$ into N^2/λ^2 sub-squares of side λ. Fix such a sub-square S and consider the $\lambda - 2$ vertex-disjoint paths that cross S in the E–W direction and are not part of the perimeter of S. Each such path occurs in $\mathcal{L}_p$ with probability $p^{2\lambda-1} \geqslant 1 - 2\lambda/\gamma$. Thus the number Z of E–W paths that *do not* appear in $\mathcal{L}_p$ is dominated by the binomial $B(\lambda, 2\lambda/\gamma)$. Thus

$$\Pr(Z \geqslant \lambda^{1/2}) \leqslant \left(\frac{2e\lambda^{3/2}}{\gamma}\right)^{\lambda^{1/2}} \leqslant \frac{1}{2}e^{-\lambda^{1/2}}.$$

Now consider mixed percolation on an $(N/\lambda) \times (N/\lambda)$ lattice where each site corresponds to a $\lambda \times \lambda$ sub-square and each site and bond are open with probability $p_1 = 1 - e^{-\gamma^{1/6}}$. We denote this random graph by $\mathcal{L}'$. We go from $\mathcal{L}_p$ to $\mathcal{L}'$ by having a site open for each sub-square in which there are at least $\lambda - 2\lambda^{1/2} - 2$ complete paths in both the E–W and N–S directions. We open a bond between 2 sites in $\mathcal{L}'$ if there is an open bond in $\mathcal{L}_p$ which connects 2 completely open paths. So if both of the two ends of a bond in $\mathcal{L}'$ are open, the bond itself is open with probability $\geqslant 1 - (1-p)^{\lambda - 2\lambda^{1/2} - 4} \geqslant p_1$. Observe also that if $\mathcal{L}'$ contains a component with M sites, then $\mathcal{L}_p$ has a component with $\geqslant M\lambda(\lambda - 2\lambda^{1/2} - 2)$ sites.

By repeating this argument, we obtain a sequence of mixed percolation models $\mathcal{M}_i$ on an $N_i \times N_i$, $N_i = N\prod_{j=0}^{i-1} \lambda_j^{-1}$, $i = 0, 1, \ldots,$ lattice with bond and site probabilities p_i, where $p_0 = p$ and $1 - p_{i+1} = \exp\{-(1-p_i)^{-1/6}\}$ and $\lambda_i = \min\{N_i, (1-p_i)^{-1/3}\}$. Furthermore, the models are coupled so that if $\mathcal{M}_{i+1}$ has a component with M_{i+1} sites then $\mathcal{M}_i$ has a component with $M_{i+1}\lambda_i(\lambda_i - 2\lambda_i^{1/2} - 2)$ sites.

Let $i_0 = \min\{i : N_i^2 p_i \leqslant 1/\gamma\}$. We see that with probability $1 - O(1/\gamma)$ the final model $\mathcal{M}_{i_0+1}$ has all its sites and bonds open. Hence, with this probability, $\mathcal{L}_p$ has a component with at least

$$N^2 \prod_{i=0}^{i_0}\left(1 - \frac{2\lambda_i^{1/2} + 2}{\lambda_i}\right) \geqslant N^2\left(1 - \sum_{i=0}^{i_0} \frac{3}{\lambda_i^{1/2}}\right) = N^2(1 - O(\gamma^{-1/6}))$$

sites.

It follows that w.h.p. almost all of the giants G_i will be part of the same component of G. This completes the proof of part (a) of Theorem 1.2.

(b) We first note that an r-regular, N-vertex graph contains at most $N(er)^{k-1}$ trees with k vertices. This is proved, for example, in Claim 1 of [12]. Thus the expected number of k-vertex trees in G is bounded by

$$n^2(4e\omega p)^{k-1} = n^2(4ec)^{k-1} = o(1)$$

if $k \geqslant A \ln n$ and A is sufficiently large.

Remark. Examining the above proof, we see that if ω is a sufficiently large constant, then there will w.h.p. be a component of size $\Omega(n^2)$, although it will not be so straightforward to put a lower bound on its size.

4. Finding paths between nodes: Proof of Theorem 1.3

Thus far, we have considered the existence of paths between nodes in random line-of-sight networks. In terms of the motivating applications, it is also interesting to consider the algorithmic problem faced by a pair of nodes s and t trying to construct a path between them in such a network. We consider a decentralized model in which each node knows only its own coordinates and those of its neighbours in G; given the coordinates of t, the node s must pass a message to t by forwarding it through a sequence of intermediate nodes. We consider the standard goal in wireless *ad hoc* routing: we wish to construct an s–t path with a small number of edges, while consulting a small number of intermediate nodes [24].

To begin the proof of Theorem 1.3, let $N = n^2$ and let $S_1, S_2, \ldots, S_N$ be the collection of all $\omega \times \omega$ sub-squares obtained by choosing ω consecutive rows and columns. Let $G_i, H_i,\ i = 1, 2, \ldots, N$ be defined as in Section 3. We first make the following observation.

Lemma 4.1.

(a) *With high probability, $G_1, G_2, \ldots, G_N$ are all connected.*

(b) *With high probability the diameter of G_i is at most $D \ln n$, $i = 1, 2, \ldots, N$, where D is some absolute constant.*

Proof. **(a)** G_1 is connected if and only if H_1 is connected. If H_1 is not connected then then there exist non-empty subsets $K \subseteq R_1$, $L \subseteq C_1$, $|K| + |L| \leqslant \omega$ such that $K \cup L$ induces a connected component of H_1. The probability that such a pair exists is at most

$$\sum_{2\leqslant k+\ell\leqslant\omega} \binom{\omega}{k}\binom{\omega}{\ell}\binom{k\ell}{k+\ell-1} p^{k+\ell-1}(1-p)^{k(\omega-\ell)+\ell(\omega-k)}$$
$$\leqslant \frac{1}{p}\sum_{2\leqslant k+\ell\leqslant\omega} \left(\frac{\omega e}{k}\right)^{k}\left(\frac{\omega e}{\ell}\right)^{\ell}\left(\frac{k\ell e}{k+\ell}\right)^{k+\ell} p^{k+\ell} e^{-((k+\ell)\omega-2k\ell)p}$$
$$\leqslant \frac{1}{p}\sum_{2\leqslant k+\ell\leqslant\omega} \left(\frac{e^2 C\ln n}{\exp\{C\ln n(1-\frac{2k\ell}{\omega(k+\ell)})\}}\right)^{k+\ell}$$
$$\leqslant \frac{1}{p}\sum_{2\leqslant k+\ell\leqslant\omega} \left(\frac{e^2 C\ln n}{n^{C/2}}\right)^{k+\ell} = O((\ln n)^2 n^{-C}).$$

So if $C \geqslant 3$ we can inflate this latter estimate by n^2/ω to account for all of $G_1, G_2, \ldots, G_N$.

(b) We concentrate on H_1. Fix $x \in R_1$ and $y \in C_1$.

Case 1: $C \ln n \leqslant \omega^{1/100}$.
Let S_k, $k = 0, 1, \ldots, t_0 = \lceil \frac{2}{3}\log_{\omega p}\omega \rceil$ be the set of vertices at distance k from x in H_1. Given $S_0, S_1, \ldots, S_{2k-1}$, the size of S_{2k} is distributed as the Binomial $B(\omega - |S_0| - |S_2| - \cdots - |S_{2k-2}|, 1-(1-p)^{|S_{2k-1}|})$. Given $S_0, S_1, \ldots, S_{2k}$ the size of S_{2k+1} is distributed as the Binomial $B(\omega - |S_1| - |S_3| - \cdots - |S_{2k-1}|, 1-(1-p)^{|S_{2k}|})$. Suppose we define the events

$$\mathcal{E}_i = \{|S_i| \in [(\omega p/2)^i, (2\omega p)^i]\}, \quad 1 \leqslant i \leqslant t_0.$$

Note that $\mathcal{E}_i$ implies that

$$|S_i| = (Ce^{o(1)} \ln n)^i \leqslant \omega^{7/10}.$$

Then the Chernoff bounds imply that $\Pr(\neg\mathcal{E}_1) \leqslant n^{-C/10}$, and for $i > 1$,

$$\Pr(\neg\mathcal{E}_i \mid \mathcal{E}_j, j < i) \leqslant e^{-C|S_{i-1}|\ln n/10}.$$

It now follows easily that

$$\Pr(\exists i \leqslant t_0 : \neg\mathcal{E}_i) = O(n^{-C/11}).$$

Thus, with probability $1 - O(n^{-C/11})$, we have that

$$|S_{t_0}| = \omega^{2/3+o(1)}.$$

Defining T_k, $k = 0, 1, \dots, t_0 = \lceil \frac{2}{3} \log_{\omega p} \omega \rceil$ to be the set of vertices at distance k from y in H_1, we see that with probability $1 - O(n^{-C/11})$ we have

$$|T_{t_0}| = \omega^{2/3+o(1)}.$$

After the construction of S_{t_0}, T_{t_0}, edges between $S_{t_0} \setminus \bigcup_{i<t_0} S_i$ and $T_{t_0} \setminus \bigcup_{i<t_0} T_i$ are unconditioned. Thus,

$$\Pr(\nexists \text{ an edge } (u,v) \in S_{t_0} \times T_{t_0}) = O(n^{-C/11}) + (1-p)^{\omega^{4/3+o(1)}} = O(n^{-C/11}).$$

Thus, with probability $1 - O(n^{-C/11})$, the distance from x to y in H_1 is $o(\ln n)$. For $C > 44$ we can inflate the failure probability by n^4 to deal with all pairs of vertices in all H_i.

Case 2: $C \ln n = \omega^\alpha$ for constant $1/100 < \alpha \leqslant 1/2$.
The same argument as above with $t_0 = \lceil 2/(3\alpha) \rceil$.

Case 3: $C \ln n = \omega^\alpha$ for constant $1/2 < \alpha \leqslant 1$.
The same argument as above with $t_0 = 1$. □

The next thing we observe is that we can now assume that all arms of all vertices are mighty. This is again a simple calculation, similar to that given for the proof of (2.2). This also allows us to specify the value of C in the expression $p = C \ln n/\omega$: it should be large enough for Lemma 4.1 to hold and for all arms of all nodes to be mighty. (In fact, as will be clear from the subsequent discussion, we will need only a weak variant of mightiness in the analysis.)

We now describe the decentralized algorithm to pass a message from a node s to a node t (thereby constructing an s–t path). The algorithm consists of two stages. First, starting at s, the message is passed between nodes on the row of s, moving the 'short way' around the torus toward the column of t. Each node passes the message to its farthest neighbour on the arm in the correct direction; since all arms are mighty, the message travels an ℓ_1-distance of at least $\omega/2$ in each step. This process stops, at a node u, when the message is about to 'overshoot' the column of t. At this point, the message is then passed between nodes in the column of u, according to the same rule. This process stops when the message is about to overshoot the row of t.

The second stage now begins, with the message at a node v that belongs to a subset B of size $\omega \times \omega$, such that B also contains t. The message is now propagated by breadth-first search to all nodes within $D \ln n$ steps, but only including nodes that belong to the set B. Here D is the constant from Lemma 4.1. (Note that by our assumption that nodes know the coordinates of themselves and their neighbours, a node can determine which subset of its neighbours lies in B and hence should be included in the breadth-first search.) By Lemma 4.1, the node t will be reached by this breadth-first search, since the subgraph of G restricted to B is connected and with appropriately short paths.

The bound on the number of edges in the resulting s–t path follows directly from the definition of the two stages. To bound the number of nodes involved in the computation, we observe that $O(d(s,t)/\omega)$ nodes are involved in the first stage, and the second stage involves at most the total number of nodes in B, which is $O(p\omega^2) = O(\omega \ln n)$ with high probability.

5. Relay placement: an approximation algorithm; Proof of Theorem 1.4

Finally, we discuss an approximation result for the relay placement problem: given a set of nodes on a grid, we would like to add a small number of additional nodes so that the full set becomes connected. As before, we are given an $n \times n$ torus of points T. Let $K = (T, E)$ be the graph defined on the points of T, in which we join two points by an edge if they can see one another. Also, we are given a cost c_x for each point $x \in T$, and for a set $X \subseteq T$ we define $c(X) = \sum_{x \in X} c_x$.

Let $X = \{x_1, x_2, \ldots, x_k\}$ be a given set of points in T. We consider the problem of choosing a set of additional points $Y = \{y_1, \ldots, y_s\}$ such that $K[X \cup Y]$ is a connected. We call Y a *Steiner set* for X; nodes placed at Y can act as 'relays' for an initial set of terminal nodes placed at X. Our goal is to find a Steiner set whose total cost is as small as possible.

This is an instance of the *node-weighted Steiner tree* problem in the graph K, with X as the set of given terminals and Y as the set of additional Steiner nodes whose total cost we want to minimize. In general, there is an $\Omega(\ln n)$ hardness of approximation for this problem [18] (and this is matched in [18] by a corresponding upper bound). However, the special structure of the graph K allows us to efficiently find a Steiner set whose cost is within a constant factor of minimum. This is the content of the following theorem, which we prove in the remainder of the section.

The crucial combinatorial property of K that we use is captured by the following definition. We say that a graph H is *d-cohesive* if every connected subset of H has a spanning tree of maximum degree d. That is, given any connected subset S of $V(H)$, we can choose a set F of edges, each with both ends in S, such that (S, F) is a tree of maximum degree at most d.

We note that it is easy to construct graphs that are not d-cohesive for any specified d; for example, any graph containing an induced $K_{1,d+1}$ is not d-cohesive. In fact, although it is not crucial for our purposes here, we note that cohesiveness is a combinatorial property of G that is almost entirely characterized by this particular type of obstruction; if we let $\kappa(G)$ denote the minimum d for which G is d-cohesive and we let $\varphi(G)$ denote the maximum t for which G contains an induced $K_{1,t}$, then we can prove the following.

Proposition 5.1. $\varphi(G) \leqslant \kappa(G) \leqslant \varphi(G) + 1$.

Proof. The proof is a direct generalization of a result of Chrobak, Naor and Novick [6], who proved the special case that every claw-free graph contains a spanning tree of maximum degree three.

The simple argument for the lower bound was observed above: if S is a set of nodes for which $G[S]$ is isomorphic to $K_{1,t}$, then S is clearly a connected subset of G with no spanning tree of maximum degree $\leqslant t-1$.

For the upper bound, let us suppose that G does not contain an induced $K_{1,t+1}$; we show how to construct a spanning tree of maximum degree $t+1$ for the subgraph induced on any connected subset S of G. To do this, we choose an arbitrary root node $v \in S$, and let $\Lambda = (S, F)$ be a spanning tree of S. We claim that Λ has maximum degree $t+1$. Indeed, suppose that some node $w \in S$ has degree greater than $t+1$ in Λ; we will suppose $w \neq v$, as the case in which $w = v$ is strictly analogous. Now, if we consider w in the depth-first traversal of Λ rooted at v, then this means that w has a parent and at least $t+1$ children. By the properties of depth-first search traversal, there can be no edges among the children of w; hence w together with ($t+1$ of) its children form an induced $K_{1,t+1}$ in G, a contradiction. It follows that Λ has maximum degree $t+1$, as desired. □

Returning to the line-of-sight graph K, a direct application of Proposition 5.1 implies that K is 5-cohesive. With somewhat more care, we can show the following.

Lemma 5.2. *The graph K is* 4*-cohesive.*

Proof. A direct application of Proposition 5.1 implies that K is 5-cohesive, but we can do better via the following argument. For each edge of K, define its *length* to be the number of rows or columns of T that separate its ends. Now, consider an arbitrary connected subset S of K, and let (S, F) be a spanning tree of S whose total edge length is minimum.

We claim that the maximum degree of (S, F) is four. For suppose not; then some node $u \in S$ has degree at least five, and hence there are two nodes $v, w \in S$ that lie on the same arm of u, and for which (u, v) and (u, w) are both edges in F. In other words, u, v, w lie in the same row or column of T, in this order, and u and w are close enough to see one another. It follows that (v, w) is also an edge of K. But now $(S, F \cup \{(v, w)\} - \{(u, w)\}$ is a spanning tree of S whose total length is strictly less than that of (S, F), a contradiction. □

We now describe the approximation algorithm and its analysis. We first define weights on the edges of K as follows. First, we say that the *X-reduced cost* c_v^X of a node v is equal to 0 if $v \in X$, and equal to c_v otherwise. We define $c^X(Y) = \sum_{y \in Y} c_y^X$. For each edge $e = (v, w)$ of K, we define its *weight* w_e to be $\max(c_v^X, c_w^X)$. For a subgraph Λ of K, let $w(\Lambda)$ denote its total edge weight.

Now, let Y^* be a Steiner set for X of minimum cost, and let Λ^* be a Steiner tree for X of minimum total edge weight. (Note that the Steiner nodes of Λ^* may be different from Y^*.) The 4-cohesiveness of K implies a corresponding gap of 4 between the cost of the optimal Steiner set Y^* and the weight of the optimal Steiner tree Λ^*.

Lemma 5.3. $w(\Lambda^*) \leqslant 4c(Y^*)$.

Proof. Since $X \cup Y^*$ is a connected subset of K, Lemma 5.2 implies that it has a spanning tree $\overline{\Lambda}$ of maximum degree four. By the definition of the edge weights, each edge $e = (v, w)$ of $\overline{\Lambda}$ has the property that at least one of its ends has an X-reduced cost that is at least as large as w_e. We charge the weight of e to this end.

Each node in $X \cup Y^*$ is charged for the cost of at most four edges, and hence $w(\overline{\Lambda}) \leqslant 4c^X(X \cup Y^*) = 4c(Y^*)$. Since $\overline{\Lambda}$ is a Steiner tree for X, and Λ^* is the Steiner tree for X of minimum total edge weight, we also have $w(\Lambda^*) \leqslant w(\overline{\Lambda})$, completing the proof. $\square$

A Steiner tree whose edge weight is within a constant factor $\gamma \leqslant 1.55$ of optimal can be computed in polynomial time via an algorithm from Robins and Zelikovsky [23]. Let Λ' be a Steiner tree for X computed using this algorithm. Let Y' be the Steiner nodes of Λ'. By charging the costs of nodes in Y' to the weights of distinct incident edges in Λ', we have the following.

Lemma 5.4. $c(Y') \leqslant w(\Lambda')$.

Proof. We root Λ' at a node in X, and we charge the cost of each node in Y' to the incident edge leading toward the root in the rooted version of Λ'. The cost of each $y \in Y'$ is thus charged to a distinct edge $e(y)$ in Λ', and by the definition of the edge weights, we have $c_y \leqslant w_{e(y)}$. $\square$

Finally, we use Y' as our Steiner set for X. Using Lemma 5.3 and Lemma 5.4, together with the approximation guarantee for the edge weight of Λ', we obtain a bound of $4\gamma \leqslant 6.2$ on $c(Y')$ relative to the optimum $c(Y^*)$:

$$c(Y') \leqslant w(\Lambda') \leqslant \gamma w(\Lambda^*) \leqslant 4\gamma c(Y^*).$$

Acknowledgement

We thank Geoffrey Grimmett and Agoston Pisztora for their advice on mixed percolation. We further thank a referee for a very thorough and helpful review.

References

[1] Alon, N. and Spencer, J. H. (2000) *The Probabilistic Method*, 2nd edn, Wiley.

[2] Bettstetter, C. (2002) On the minimum node degree and connectivity of a wireless multihop network. In *Proc. 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing* (MobiHoc '02), pp. 80–91.

[3] Bettstetter, C. (2004) On the connectivity of ad hoc networks. *The Computer Journal* (special issue on Mobile and Pervasive Computing) **47** 432–447.

[4] Bollobás, B. (2001) *Random Graphs*, 2nd edn, Cambridge University Press.

[5] Bollobás, B., Janson, S. and Riordan, O. Line-of-sight percolation. *Combin. Probab. Comput.*, this issue.

[6] Chrobak, M., Naor, J. and Novick, M. Using bounded degree spanning trees in the design of efficient algorithms on claw-free graphs. In *Proc. Workshop on Algorithms and Data Structures*, Vol. 382 of *Lecture Notes in Computer Science*, pp. 147–162.

[7] Deuschel, J. and Pisztora, A. (1996) Surface order large deviations for high-density percolation. *Probab. Theory Rel. Fields* **104** 467–482.

[8] Efrat, A. and Har-Peled, S. (2002) Locating guards in art galleries. In *2nd IFIP International Conference on Theoretical Computer Science*, pp. 181–192.

[9] Efrat, A., Har-Peled, S. and Mitchell, J. S. B. (2005) Approximation algorithms for two optimal location problems in sensor networks. In *Broadnets 2005*, pp. 767–776.

[10] Falconer, K. J. and Grimmett, G. R. (1992) On the geometry of random Cantor sets and fractal percolation. *J. Theoret. Probab.* **5** 465–485,

[11] Frieze, A. M., Kleinberg, J., Ravi, R. and Debany, W. (2007) Line-of-sight networks. In *Proc. 18th ACM–SIAM Symposium on Discrete Algorithms*, pp. 968–977.

[12] Frieze, A. M. and Molloy, M. (1999) Splitting an expander graph. *J. Algorithms* **33** 166–172.

[13] Goel, A., Rai, S. and Krishnamachari, B. (2004) Sharp thresholds for monotone properties in random geometric graphs. In *Proc. ACM Symposium on Theory of Computing*, pp. 580–586.

[14] Grimmett, G. (1999) *Percolation*, 2nd edn, Springer.

[15] Gupta, P. and Kumar, P. (1998) Critical power for asymptotic connectivity in wireless networks. In *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming* (W. M. McEneaney, G. Yin and Q. Zhang, eds), Birkhäuser, Boston, pp. 547–566.

[16] Gupta, P. and Kumar, P. (2000) The capacity of wireless networks. In *IEEE Trans. Inform. Theory* **46** 388–404. Corrigendum: A correction to the proof of a lemma in 'The capacity of wireless networks', *IEEE Trans. Inform. Theory* **49** 3117.

[17] Kalai, G. and Matousek, J. (1997) Guarding galleries where every point sees a large area. *Israel J. Math.* **101** 125–139.

[18] Klein, P. N. and Ravi, R. (1995) A nearly best-possible approximation algorithm for node-weighted Steiner trees. *J. Algorithms* **19** 104–115.

[19] Mauve, M., Widmer, J. and Hartenstein, H. (2001) A survey on position-based routing in mobile ad hoc networks. *IEEE Network Magazine* **15** (6) 30–39.

[20] Mobile ad-hoc networks (MANET) charter.
`http://www.ietf.org/html.charters/manet-charter.html`

[21] Penrose, M. D. (1999) On k-connectivity for a geometric random graph. *Random Struct. Alg.* **15** 145–164.

[22] Penrose, M. D. (2003) *Random Geometric Graphs*, Vol. 5 of *Oxford Studies in Probability*, Oxford University Press.

[23] Robins, G. and Zelikovsky, A. (2000) Improved Steiner tree approximation in graphs. In *Proc. 10th Annual ACM–SIAM Symposium on Discrete Algorithms*, pp. 770–779.

[24] Royer, E. and Toh, C.-K. (1999) A review of current routing protocols for ad-hoc mobile wireless networks. *IEEE Personal Communications*, April 1999.

[25] Valtr, P. (1998) Guarding galleries where no point sees a small area. *Israel J. Math.* **104** 1–16.